5 The Impact of Anger Expressions on Leadership Evaluations

The analysis of the experimental data is divided into three subchapters and followed by a discussion of the experimental findings. The first subchapter deals with the average treatment effects that occurred after participants received the experimental treatment. The second subchapter looks at moderating effects that shaped how respondents reacted to the video clips. Characteristics on the individual level such as party identification and personality traits are considered in this subchapter. Lastly, the third subchapter considers the relevance of the findings in light of some broader considerations, such as the longevity of the average treatment effects, potential spillover effects, and the response time of participants. By analyzing the response time as an indication of underlying mechanisms further insights into the occurrence of treatment effects can be discerned. Spillover effects in regard to the evaluation of the respective political parties are investigated in order to determine the broader relevance of anger expressions. Experimental studies often focus solely on immediate short-term effects. To counteract this shortcoming, the third subchapter focuses on the longevity of the treatment effects by analyzing evaluations of political candidates in the third wave of the panel design. The following subchapter deals with the average treatment effects and begins with a brief introduction into the statistical analysis that is chosen to analyze the experimental data, after which the analyses of the results are presented.

5.1 The Average Treatment Effects

Several strategies to analyze experimental data are discussed within the literature on experimental design. Since the 1960s and Lord's paradox (Lord 1967), statisticians have debated over the correct approach to experimental data. Lord's paradox revolves around the question whether to analyze experimental data based on a gain score analysis in a simple regression framework or as a multiple regression with the post-test measures as dependent variable, the treatment as independent variable, and the initial score of the dependent variable as covariate. A gain score analysis focuses on the occurring change between conditions by subtracting the initial pre-test score from the final post-test score after the treatment was administered.

Both variations can lead to varying conclusions about the statistical association. It has been pointed out that both approaches are suitable to answer slightly different research questions. In the context of experimental data, the choice between a gain score analysis and an analysis with covariates depends on the underlying research question. According to Hand (1994) and Wright (2006), both approaches answer different questions. The gain score analysis answers the question, "whether the average gain in score is different for the two groups" (Wright 2006: 666); the analysis of covariance shifts the focus from differences between the experimental groups to individuals with the same baseline values, by asking the question, "whether the average gain, partialling out pre-scores, is different between the two groups" (Wright 2006: 666). Focusing on the research question of this study, it seems most appropriate to start with a gain score analysis without any covariates to focus on changes that occurred between the experimental groups.

Imbens and Rubin (2015) explain the suitability of regression methods for analyzing completely randomized experiments (Imbens & Rubin 2015: 113–140). It can be implemented to estimate the (local) average treatment effect and they can easily be extended to include control and moderating variables. In this subchapter, the experimental results are presented without any covariates as a first step; in the following analysis of moderating effects, the previously discussed control variables will be used in addition to the moderating factors of interest.

Differences between several experimental and control groups can be determined by fitting either a regression model or an analysis of variance. In order to analyze experimental data, research in psychology commonly uses an analysis of variance, while regression models are traditionally more frequently applied in economics. The distinction and separation between both approaches is the result of varying research traditions within the fields of economics and psychology. Previously, researchers have stressed the differences between both approaches instead of emphasizing their great similarities. The analysis of variance can even be interpreted as a special case of a regression analysis (Field 2009: 349).[72]

---

72 In a simple regression model, which only includes the treatment variable as independent variable, the F-test is identical to the F-test of a one-way ANOVA. If the experimental groups differ in their means, the F-test is statistically significant and accounting for the means of each experimental group leads to a better overall fit to the dependent variable rather than taking the overall mean. The one-way ANOVA usually requires an additional investigation of the group means or further post hoc tests for comparing the significant differences. The regression analy-

176

Like post hoc tests after significant ANOVA tests, regression models can be adjusted to control for multiple testing. A Bonferroni adjustment can be incorporated into the framework of a regression analysis. Considering the experimental design that involves groups of male and female politicians, coefficients of four contrasts or dummy variables are estimated. Consequently, a Bonferroni adjustment for four simultaneous comparisons due to the four dummy variables results in an alpha level that should be four times smaller than the initial 5 percent threshold. Therefore, the adjusted threshold to determine significance for any of the four comparisons should be $p < 0.0125$ $(0.05/4 = 0.0125)$. If the significance is estimated based on the Bonferroni adjustment, the statistical significance for each dummy-variable should fall below this adjusted threshold of $p < 0.0125$, which equals a type I or alpha error rate of 1.25 percent for each comparison against the reference category. By doing so, the initial significance level of 5 percent remains as threshold for the analysis, i.e., a treatment effect of the experimental video clips. Likewise, the Bonferroni adjustment for three comparisons – used in models that focus on Merkel and Gysi – results in an adjusted threshold of $p < 0.0167$ $(0.05/3 = 0.0167)$ for each comparison, and the models which analyze Merkel and Gabriel should have a Bonferroni adjustment of $p < 0.025$ $(0.05/2 = 0.025)$. In multiple comparisons, the Bonferroni adjustment is a common procedure to adjust for multiple testing and is known as a family-wise error rate.[73] In this analysis the alpha level is adjusted accordingly.[74]

Dummy codings are a widespread convention for treating categorical variables in regression models across the social sciences and therefore, make the results easily interpretable compared to the use of user-specified contrasts. Using the experimental treatment that showed negative-active emotions as a reference category and therefore as a constant (intercept) in the

---

sis typically provides the estimated coefficients and a dummy coding is used as default for investigating experimental treatments, whereby one variable is set as a baseline category to which the other variables are compared. However, different contrasts such as Helmert contrasts or user-defined contrasts can be also conducted within a regression model. In addition, the reference category can also be changed to facilitate the interpretation of the model and allow for multiple comparisons.

73 As the critical t-values for the corrected p-values depend on the degrees of freedom, the t-values are adjusted to the respective sample sizes of the experimental subtypes 1,2, and 3.

74 Alternatively, the observed p-values could have been multiplied by the number of comparisons while adhering to the 5-percent level.

regression models illustrates the results, since all four comparisons of interest are conducted simultaneously. Although it might be less conventional to use the variable of interest as a reference category, this approach offers an additional advantage. By using the experimental treatment of anger expressions as a reference category, not only is the coefficient easily interpretable as the estimated mean for the negative-active experimental group, but from gauging its significance it can be easily seen whether an overall change between pre- and post-test measures occurred within the anger conditions.[75] If the constant is not significant, this can be interpreted as an indication that no substantial change occurred within the group itself. The overall F-test of the regression models shows whether the group means differ from the overall mean – or to put it differently, whether the experimental treatment groups improve the prediction of individual scores, compared to the overall mean as a predictor for each single score (Field 2009: 201–204).

The following section focuses on the average treatment effect of negative-active emotions. Dependent variables such as overall evaluations, candidate orientation, and semantic differentials are analyzed separately for each sub-experiment and politician. The variables under consideration are largely suited for the analysis. The gain scores of these Likert-scale variables as dependent variables approximately follow a normal distribution. Since the experimental treatments are the main independent variables, and only a few moderating and control variables are considered at a later stage, multicollinearity among the independent variables is not a major concern of this analysis. The observations are statistically independent of each other. The underlying assumptions of regression analyses are largely met for the following analysis.[76] It begins with the overall opinion of each politician and politicians in general; more specific items of candidate orientations follow, and it continues by focusing on the evaluations according to semantic differentials. The last section of Subchapter 5.1 summarizes and discusses the average treatment effects.[77]

---

75 In the following, the experimental treatment groups (conditions) in which participants were exposed to politicians' anger expressions are referred to as anger conditions (for comparison see Valentino et al. 2008; Ryan 2012), or negative-active conditions.

76 This includes the statistical independence of the observations as well as the homoscedasticity and approximated normal distribution of the residuals.

77 All statistical models were fitted using the statistical software R Version 3.4.1 (R Core Team 2017). In addition, the following R packages were used: "ggthemes" (Arnold et al. 2017), "gridExtra" (Auguie & Antonov 2017), "questionr" (Barnier

5.1.1 The Evaluation of Overall Favorability Ratings

Table 22 shows the changes in the evaluation of overall opinions for those participants who saw several male or female politicians in sub-experiment *Type 1* (Models 1 and 2), Angela Merkel and Gregor Gysi in sub-experiment *Type 2* (Models 3 and 4), as well as those participants who saw video clips of Angela Merkel and Sigmar Gabriel in sub-experiment *Type 3* (Models 5 and 6).

Evaluations in Sub-Experiment Type 1 (Politicians in General)

Starting with the two models that focused on the evaluation of politicians in general (Models 1 and 2), the overall F-test of both models indicate significant differences between the experimental groups. Focusing on the experimental groups in which male politicians were seen in the experimental treatment, it can be seen that the constant – and therefore negative-active emotions – have no significant effect on the evaluation of overall opinions about politicians. The significant effects have occurred because participants in other experimental groups rate politicians as more favorably in the post-test condition. This significant change also occurs in the control group without video treatment, which is contrary to the initial assumptions of the experimental design. However, in this light, the absence of improved ratings only within the negative-active treatment group could imply that negative-active emotions might not necessarily lead to lower ratings on an individual level, but they could have prevented an increase in overall favorability. The adjusted R^2 of this model indicates that only 0.4

2017), "GPArotation" (Bernaards & Jennrich 2015), "pwr" (Champely et al. 2017), "igraph" (Csardi 2017), "xlsx" (Dragulescu 2015), "qgraph" (Epskamp et al. 2017), "tm" (Feinerer et al. 2017), "wordcloud" (Fellows 2013), "car" (Fox et al. 2017a), "effects" (Fox et al. 2017b), "Hmisc" (Harrell 2017), "stargazer" (Hlavac 2015), "quantreg" (Koenker et al. 2017), "statnet.common" (Krivitsky et al. 2017), "gtrendsR" (Massicotte & Eddelbuettel 2017), "vegan" (Oksanen et al. 2017), "psych" (Revelle 2017), "XML" (Temple Lang & CRAN-Team 2017), "IsingFit" (van Borkulo et al. 2016); "NetworkComparisonTest" (van Borkulo 2016), "corrplot" (Wei 2016), "reshape2" (Wickham 2017), "ggplot2" (Wickham et al. 2016), "haven" (Wickham et al. 2017a), "dplyr" (Wickham et al. 2017b), "httr" (Wickham & RStudio 2017a), "stringr" (Wickham & RStudio 2017b), "devtools" (Wickham et al. 2017c).

percent of the variance within the dependent variable can be explained by the experimental treatment. Therefore, the effect size of the experimental treatment is very small. Furthermore, it has to be noted that the difference for the effect of negative-active emotions compared to the neutral control group and the positive experimental group is only significant at the 5-percent level.

The next model analyzes the impact displays of negative-active emotions shown by female politicians have on the overall ratings of politicians (Model 2). The effect of negative-active emotional displays is statistically significant only at a 5-percent level in comparison to the group without video treatment as well as in comparison to the group that received negative-passive emotional expressions of female politicians. As before, compared to the negative-active treatment, the overall rating of politicians increased slightly in the no video condition and negative-passive experimental group. No significant difference occurs between the groups that showed angry and neutral expressions of female politicians, while a significant difference at a 0.1-percent level occurred when comparing the effect of expressions of anger to a positive video treatment. In this case, participants in the anger condition did not judge politicians less favorably over time per se, as no significant change occurred within the experimental group. Only in comparison to positive emotions are they rated 0.5 points lower on average. According to the adjusted R^2 of this model, only 0.7 percent of variance within the dependent variable can be explained by the model – slightly higher than the previous model but still very low.

Evaluations in Sub-Experiment *Type 2* (Merkel and Gysi)

The next two models are based on the sub-experiment *Type 2*, in which participants were exposed to Angela Merkel and Gregor Gysi (Models 3 and 4). Only once participants saw both videos did they evaluate Angela Merkel and Gregor Gysi. The model which measures changes in Angela Merkel's overall assessment can explain 1.4 percent of variance within the dependent variable. In this model, two comparisons are significant at the 1-percent level. Participants evaluated Angela Merkel on average -0.35 points lower, if they received her negative-active emotional displays as the experimental treatment. Compared to the control group without video treatment, participants evaluated her on average 0.33 points lower than before they received the video treatment. In addition, participants who were exposed to the anger condition also rated her on average 0.48 points lower

than participants who were exposed to her positive expressions. There is, however, no significant difference between the effects, which her anger and neutral emotional expressions have on participants' overall evaluations of her. Both conditions led to lower ratings in this sub-experiment.

Looking at the evaluation of Gregor Gysi, participants' overall opinions of him changed more drastically. In this model, almost 5 percent of variance can be explained within the dependent variable of Gregor Gysi's overall ratings (adj. R^2 = 0.049). While this model still does not explain much of the variance, it is a noticeable increase compared to the previous models which ranged from less than half a percent to slightly more than the one percent of the explained variance. This increase in the effect size is also reflected in larger coefficients within this model. All comparisons are highly significant at the 0.1-percent level. The strongest effect occurs when comparing changes in Gregor Gysi's overall ratings between the control group without video treatment and anger. Compared to participants who saw no video, participants in the anger condition evaluated Gysi on average 1.16 points more favorably. This positive effect of Gysi's displays of anger is also significantly different from those groups that were exposed to other emotional expressions. Although the effects are cut in half, participants in the anger condition still rated Gysi more favorably after having been exposed to anger compared to neutral expressions, by 0.57 points on average; and even by 0.68 points on average in comparison to those who were exposed to Gysi's positive emotional expressions. While this analysis indicates that participants formed a more positive impression of Gregor Gysi, once they were exposed to him in either video treatment, displays of anger differed significantly from effects of his other appearances and resulted in the highest positive changes of his overall evaluation. This finding contradicts the occurrence of a mere exposure effect due to the visual exposure. Instead it supports the importance of emotional communication because his anger expressions make an additional difference in his overall evaluation.

Evaluations in Sub-Experiment *Type 3* (Gabriel and Merkel)

Sigmar Gabriel's displays of anger led to a decrease of his overall impression ratings, both comparisons are significant at the 5-percent level and also significant according to the Bonferroni adjusted significance level of 2.5 percent, since both t-values exceed the critical t-value of 2.24 (-0.309/0.134 = 2.31 and -0.331/0.137 = 2.42). On average, participants in the anger condition evaluated Sigmar Gabriel -0.31 points lower than

those who received no video; compared to participants who saw Sigmar Gabriel's neutral emotions, it was on average -0.33 points lower. No significant change occurred within the experimental group that saw his negative-active emotional displays (see Table 22). The adjusted R^2 indicates that a variance of only one half a percent (adj. $R^2 = 0.005$) within the dependent variable can be explained by the experimental treatment (Model 5).
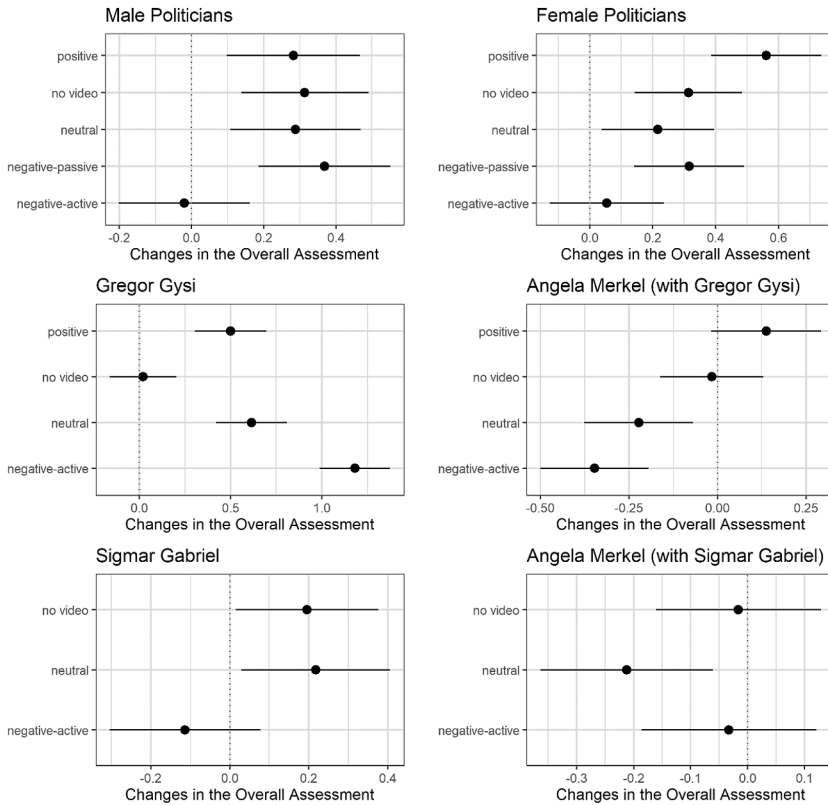
An additional model measured Merkel's overall assessment when participants were exposed to her and Sigmar Gabriel instead of Gregor Gysi (Model 6). Although the sub-experiment *Type 2* found significant differences for the overall ratings of Merkel when comparing participants who saw anger with those who received either no video or positive emotions in a video treatment, the findings could not be replicated when participants were also exposed to Sigmar Gabriel. In this instance, the overall F-test is not significant and significant differences cannot be obtained between the experimental treatments. When participants were also exposed to Sigmar Gabriel, Angela Merkel's expressions of anger did not have a negative impact on her evaluation, as they had when Gysi was seen (Model 3). This is an indication of a contrast effect that might have occurred within both sub-experiments (see the middle and bottom row in Figure 17).

Table 22: Changes in the Overall Evaluation of Politicians

	Male Politicians (1)	Female Politicians (2)	Merkel (E2) (3)	Gysi (E2) (4)	Gabriel (E3) (5)	Merkel (E3) (6)
			Dependent Variable: Scalometer			
Negative-Passive	0.389** (0.131)	0.262* (0.128)				
Neutral	0.308* (0.131)	0.162 (0.130)	0.125 (0.110)	-0.567*** (0.139)	0.331b (0.137)	-0.179 (0.110)
No Video	0.334** (0.129)	0.260* (0.127)	0.331** (0.107)	-1.161*** (0.136)	0.309b (0.134)	0.017 (0.108)
Positive	0.302* (0.132)	0.508*** (0.129)	0.484*** (0.111)	-0.682*** (0.140)		
Constant	-0.020 (0.093)	0.054 (0.093)	-0.347*** (0.078)	1.182*** (0.098)	-0.113 (0.097)	-0.033 (0.078)
Observations	1,764	1,768	1,374	1,374	1,052	1,051
Adjusted R^2	0.004	0.007	0.014	0.049	0.005	0.002
F Statistic	2.726*	4.131**	7.580***	24.696***	3.692*	2.018

Note: *p < 0.05; b p < Bonferroni adjusted; **p < 0.01; ***p < 0.001. The reference category is the negative-active condition. Cells display OLS-estimates and standard errors in parentheses.

Figure 17: Changes in the Overall Evaluations of Politicians

To sum up, the effects on the scalometer ratings were very small or negligible for male and female politicians as well as for Sigmar Gabriel – and partially also for Angela Merkel. However, the impact of Gregor Gysi's anger expressions differed in two regards: they induced positive evaluations and had a stronger impact on the overall evaluation. In addition, the exposure to Gregor Gysi potentially affected Angela Merkel's ratings. Participants rated her on average more negatively after they had seen negative emotional displays by her and Gregor Gysi, whereas a third experiment in which participants were exposed to Angela Merkel and Sigmar Gabriel failed to replicate these negative effects on Merkel's evaluations. Therefore, it seems that the evaluation of politicians depends on the context of other politicians and their emotional expressions. Politicians might have been evaluat-

ed in relation to other politicians and a *contrast effect* could have occurred (Kuklinski et al. 1997: 328). This finding could also indicate the use of an anchor or sequential heuristic (Tversky & Kahneman 1974: 1128), if the evaluation was dependent on the sequence in which the participants saw the videos. Hence, the sequential order will be considered later as a moderating factor.

5.1.2 The Evaluation of Warmth

Next, the dimension of warmth is considered as the dependent variable. According to theoretical assumptions, two items – likeability ratings and evaluations of trustworthiness – can be considered as representing the dimension of warmth. While an overall rating is often considered as a summary score and most closely connected to the dimension of warmth (Laustsen & Bor 2017), both indicators give a more precise measure of two more distinct aspects. The following section considers effects of anger that appear when likeability ratings are chosen as the dependent variable (see Table 23 and Figure 18).

5.1.2.1 The Evaluation of Likeability Ratings

Treatment: Sub-Experiment Type 1 (Politicians in General)

When considering the impact anger expressions of male politicians have on the likeability ratings of politicians in general, the overall F-test of the model reveals that no significant effects between the experimental groups can be determined (Model 1). Hence, politicians' likeability ratings in general were not affected by negative emotional expressions – or by any emotional displays from male politicians (see Figure 18 for illustration).

When participants saw female politicians in the experimental condition, anger again did not lead to a significant change in the evaluation (neither positive nor negative). However, one individual predictor indicates that negative-active displays differ significantly at a 5-percent level from positive emotional expression of female politicians, which led to higher likeability ratings.

184

Treatment: Sub-Experiment Type 2 (Gysi and Merkel)

When analyzing likeability ratings, only the models for Angela Merkel and Gregor Gysi show significant results. Angela Merkel's negative-active emotional displays resulted in lower likeability ratings when compared with the impact of her positive expressions and the no video condition. Her angry and neutral emotional expressions did not differ significantly. Her negative emotional displays decreased her likeability ratings on average by 0.17 points, while no substantial change occurred for the control group without video treatment (-0.01 = -0.168 + 0.162) and a slight positive change in likeability ratings can be observed, for the experimental group which received positive emotional stimuli (0.16 = -0.168 + 0.327). The adjusted R^2 indicates that in the case of Angela Merkel only 1.9 percent of variance can be explained; for Gysi, the amount of explained variance is twice as much – but still only a mere 3.8 percent."

Comparing these effects to the impact of Gregor Gysi's displays of anger, a different pattern emerged. The mean of the experimental groups which saw his displays of anger and were not exposed to any emotional expressions show a highly significant difference of their mean likeability ratings at a 1-percent level. His displays of anger result on average in an increase of 0.4 points in his likeability ratings; the group without video treatment showed no substantial change between the evaluation of Gysi's likeability between the first and second panel wave (-0.08 = 0.400 - 0.478). Furthermore, Gysi's displays of anger differ only at a 5-percent level from his neutral expressions (failing to meet the Bonferroni adjusted p-value), whereby his negative-active emotions led to more favorable ratings compared to his neutral emotions. His anger expressions did not differ significantly from his positive emotional expressions, which also have a positive impact on his ratings. Overall, this pattern could indicate an overall exposure effect, whereby Gregor Gysi's likeability seems to be rated more favorably regardless of the emotions he expresses. His negative-active emotional displays were significantly more beneficial than his neutral expressions for his likeability ratings at the 5-percent level, however, this difference falls slightly short of the Bonferroni adjusted significance level.

Treatment: Sub-Experiment Type 3 (Gabriel and Merkel)

Regarding the sub-experiment that showed video clips of Sigmar Gabriel and Angela Merkel, no significant differences occurred between the experi-
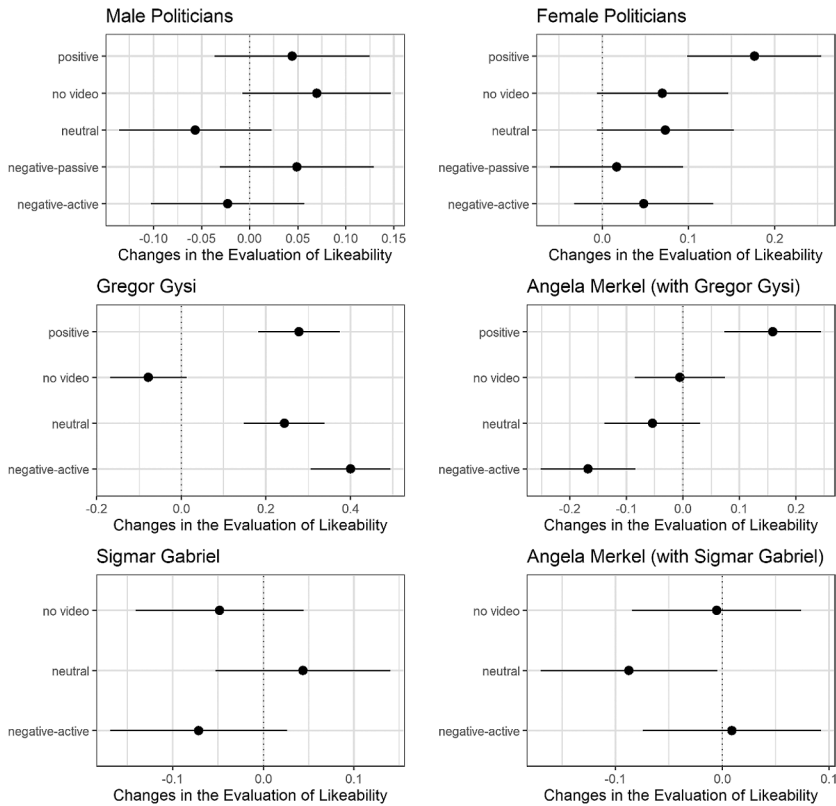
mental groups, when participants were asked to evaluate Sigmar Gabriel's and Angela Merkel's likeability. Once again, this finding could indicate that Angela Merkel's likeability was at least partially affected by the second video that respondents saw, leading to significant effects of her anger expressions in one instance, while they did not occur in another setting.

Table 23: Changes in the Evaluation of Likeability

	Dependent Variable: Likeability					
	Male Politicians (1)	Female Politicians (2)	Merkel (E2) (3)	Gysi (E2) (4)	Gabriel (E3) (5)	Merkel (E3) (6)
Negative-Passive	0.072 (0.058)	-0.031 (0.057)				
Neutral	-0.034 (0.057)	0.025 (0.058)	0.114 (0.061)	-0.157* (0.068)	0.115 (0.070)	-0.096 (0.060)
No video	0.093 (0.057)	0.022 (0.057)	0.162** (0.059)	-0.478*** (0.067)	0.023 (0.069)	-0.014 (0.059)
Positive	0.067 (0.058)	0.128* (0.057)	0.327*** (0.061)	-0.122 (0.069)		
Constant	-0.023 (0.041)	0.048 (0.041)	-0.168*** (0.043)	0.400*** (0.048)	-0.071 (0.050)	0.009 (0.043)
Observations	1,764	1,766	1,373	1,372	1,051	1,053
Adjusted R^2	0.002	0.003	0.019	0.038	0.001	0.001
F Statistic	1.780	2.290	9.851***	19.027***	1.535	1.526

Note: *p < 0.05; ^{b}p < Bonferroni adjusted; **p < 0.01; ***p < 0.001. The reference category is the negative-active condition. Cells display OLS-estimates and standard errors in parentheses.

Figure 18: Changes in the Evaluation of Likeability

Note: Figure displays coefficient plots with 95-percent confidence intervals.

5.1.2.2 The Evaluation of Trustworthiness

While trustworthiness is considered to belong to the dimension of warmth by some scholars (e.g., Fiske et al. 2002; Cuddy et al. 2008), others argue that trustworthiness represents a unique quality of integrity (e.g., Ohr & Oscarsson 2011). Compared to the previously analyzed likeability ratings, the evaluation of trustworthiness follows a similar pattern. Only the regression models for the sub-experiment that showed Angela Merkel and Gregor Gysi are statistically significant (Models 3 and 4 in Table 24; Figure 19). According to the adjusted R^2-values, the amount of variance in their trustworthiness ratings that can be explained by the experimental condi-

187

tions is very small. In the case of Angela Merkel, only 0.5 percent of variability can be attributed to the experimental treatment, while in Gregor Gysi's case, 3.6 percent of variability in the dependent variable (changes in the evaluation of trustworthiness) can be accounted for by the experimental treatment.

When analyzing Angela Merkel's trustworthiness, displays of negative-active emotions led on average to lower ratings of trustworthiness by 0.130 points. In comparison to the control group which received no video treatment, this difference is highly significant at a 1-percent level, while the mean for the control group is 0.03 (0.03 = -0.126 + 0.156). Compared to the positive treatment condition in which no change occurred, the difference is significant at a 5-percent level. Compared to neutral emotional expressions, no significant difference can be observed.
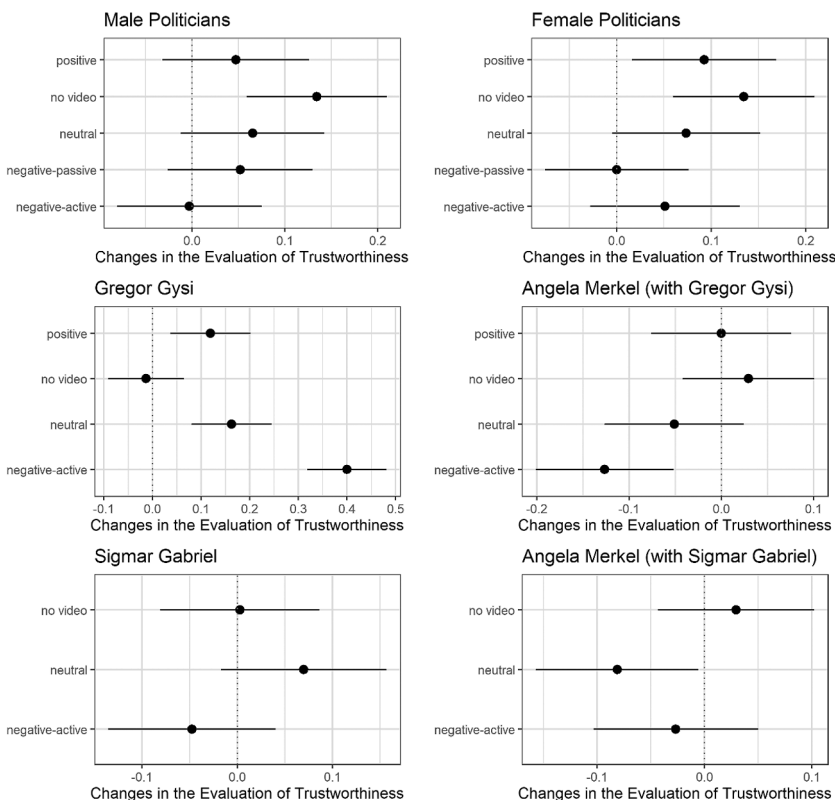
Regarding the evaluation of Gregor Gysi's trustworthiness (Model 4), it can be seen that Gysi's anger expressions had a positive impact and resulted in an average increase of his trustworthiness ratings by 0.4 points. The mean of the negative-active experimental condition differed significantly from any other treatment and control group at a 0.1 percent level. Anger had the strongest positive effect on Gysi's trustworthiness ratings compared to all other emotional expressions. While it is possible that an exposure effect occurred, the anger condition differed significantly from the neutral control group that has a mean of 0.16 (0.16 = 0.400 - 0.237), which is 0.24 points lower than the mean for the experimental group that was exposed not only to Gregor Gysi's appearance but to his negative-active emotional displays as well. Likewise, participants who saw his anger rated him on average 0.28 points more trustworthy than those participants who saw his positive emotional expressions. Participants within the positive treatment condition evaluated him on average only 0.12 points more trustworthy when receiving the positive experimental treatment.

188

Table 24: Changes in the Evaluation of Trustworthiness

	Male Politicians (1)	Female Politicians (2)	Merkel (E2) (3)	Gysi (E2) (4)	Gabriel (E3) (5)	Merkel (E3) (6)
	Dependent Variable: Trustworthiness					
Negative-Passive	0.055 (0.056)	-0.051 (0.056)				
Neutral	0.068 (0.056)	0.022 (0.057)	0.075 (0.054)	-0.237*** (0.059)	0.117 (0.063)	-0.055 (0.055)
No Video	0.137* (0.055)	0.083 (0.056)	0.156** (0.053)	-0.413*** (0.057)	0.050 (0.062)	0.056 (0.054)
Positive	0.050 (0.057)	0.041 (0.056)	0.126* (0.054)	-0.281*** (0.059)		
Constant	-0.003 (0.040)	0.051 (0.040)	-0.126*** (0.038)	0.400*** (0.041)	-0.048 (0.045)	-0.027 (0.039)
Observations	1,762	1,764	1,374	1,372	1,052	1,053
Adjusted R^2	0.001	0.002	0.005	0.036	0.001	0.002
F Statistic	1.599	1.665	3.307*	17.848***	1.748	2.145

Note: *$p < 0.05$; [b]$p <$ Bonferroni adjusted; **$p < 0.01$; ***$p < 0.001$. The reference category is the negative-active condition. Cells display OLS-estimates and standard errors in parentheses.

Figure 19: Changes in the Evaluation of Trustworthiness

Note: Figure displays coefficient plots with 95-percent confidence intervals.

5.1.3 The Evaluation of Competence

The dimension of competence includes individual traits that are related to the concepts of performance and capability of fulfilling the requirements of a political office – problem-solving and leadership skills. A politician's capability to solve political problems could take several forms and shapes, as problems could be solved by different paths and negotiating styles. Hence, it is worthwhile to distinguish between problem-solving and leadership skills. Leadership skills more strongly require a dominant behavior than problem-solving skills. The reliability of both items as a measure of competence is merely adequate (see Table 16). Measuring the underlying

structure with correlation and factor analysis as well as network models showed that leadership skills were less strongly correlated with items belonging to the warmth dimension than were problem-solving skills (see Subchapter 4.4.1).

5.1.3.1 Problem-Solving Skills

Similarly to the evaluation of politicians' warmth, negative-active emotional displays have only significant effects on the evaluation of problem solving-skills for the sub-experiment *Type 2*, in which participants saw videos of Angela Merkel and Gregor Gysi. However, the overall F-test is only significant for the evaluation of Gregor Gysi (see Table 25).

The overall model for Angela Merkel is not significant at a 5-percent level in the sub-experiment *Type 2*, only at a 10-percent level (F(3, 1370) = 2.44, p = 0.063), indicating that the means between the experimental groups hardly differ from each other. Keeping this limitation in mind, participants who received Angela Merkel's negative-active emotional expressions rated her problem-solving skills on average 0.11 points lower after having seen the video treatment. This effect is significant at a 1-percent level, and compared to those participants who received no video treatment significant at a 5-percent level – failing, however, to meet the Bonferroni adjusted p-value. No other significant differences could be obtained for Angela Merkel's problem-solving skills, highlighting the overall absence of differences between the groups. Therefore, the adjusted R^2 of this model is very small, with a mere 0.3 percent of explained variance in the dependent variable. The potential negative effect of anger expressions by Angela Merkel cannot be replicated when participants were also exposed to Sigmar Gabriel.

For Gregor Gysi the effect is still very small but slightly stronger and significant. In Model 4, at least 1.3 percent of variance in the dependent variable can be explained by the model (see Table 25). This occurring pattern is similar to the previous impressions based on models which analyzed Gysi's character traits. In this instance, the evaluation of his problem-solving skills also increased after participants saw a video of his anger, which is a highly significant effect at the 1-percent level compared to the no video and neutral conditions. In addition, in this instance negative-active emotional expressions increased the evaluation of problem-solving skills by 0.32 points on average. This effect is significant even compared to the effect of positive emotional displays according to the Bonferroni adjusted p-
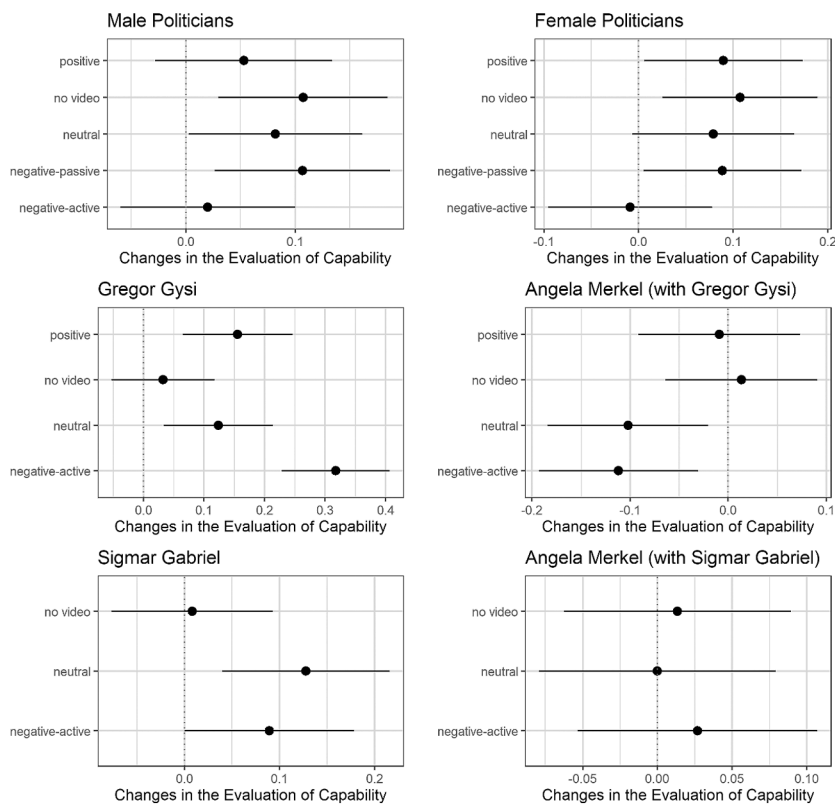
value (t = 2.5). Hence, this positive effect cannot be simply described as a mere exposure effect due to his physical appearance or general attributes (see also Figure 20). The next section focuses on the assessment of leadership skills after the experimental treatment occurred.

Table 25: *Changes in the Evaluation of Problem-Solving Skills*

	Male Politicians (1)	Female Politicians (2)	Merkel (E2) (3)	Gysi (E2) (4)	Gabriel (E3) (5)	Merkel (E3) (6)
			Dependent Variable: Problem-Solving Skills			
Negative-	0.087	0.098				
Passive	(0.058)	(0.061)				
Neutral	0.062	0.088	0.010	-0.194**	0.039	-0.027
	(0.057)	(0.062)	(0.059)	(0.065)	(0.064)	(0.057)
No Video	0.087	0.116	0.125*	-0.285***	-0.081	-0.013
	(0.057)	(0.061)	(0.057)	(0.063)	(0.063)	(0.056)
Positive	0.033	0.099	0.103	-0.162^b		
	(0.058)	(0.062)	(0.059)	(0.065)		
Constant	0.020	-0.009	-0.112**	0.318***	0.089*	0.027
	(0.041)	(0.044)	(0.041)	(0.045)	(0.045)	(0.041)
Observations	1,764	1,766	1,374	1,372	1,052	1,053
Adjusted R^2	-0.0003	0.0002	0.003	0.013	0.002	-0.002
F Statistic	0.849)	1.101	2.436	7.089***	1.940	0.109

Note: $^*p < 0.05$; $^bp <$ Bonferroni adjusted; $^{**}p < 0.01$; $^{***}p < 0.001$. The reference category is the negative-active condition. Cells display OLS-estimates and standard errors in parentheses.

Figure 20: Changes in the Evaluation of Problem-Solving Skills

Note: Figure displays coefficient plots with 95-percent confidence intervals.

5.1.3.2 Leadership Skills

In line with the previous models, no significant effect occurred for the evaluation of politicians in general, regardless of whether participants saw several female or male politicians (see Table 26 and Figure 21). In contrast to the previous findings regarding the dimension of warmth, no significant effect occurred when participants evaluated Angela Merkel's leadership skills. The F-tests of Model 3 and Model 6 are not significant, as the experimental conditions did not deviate from the overall mean. For the evaluation of leadership skills, the experimental treatment has only a significant impact for the two leading male politicians – Gregor Gysi and Sig-

193

mar Gabriel (Models 4 and 5). The amount of explained variance is very small in both instances, a change of 0.8 percent in the evaluation of Gregor Gysi's leadership skills and only a change of 0.4 percent in the evaluation of Sigmar Gabriel's leadership skills can be attributed to the experimental treatment.

Both models indicate a positive effect of anger displays on the evaluation of leadership skills for the respective male politician (Models 4 and 5). In Gregor Gysi's case, anger had a positive effect on his leadership skill evaluations, which is only significant at the 5-percent level in comparison to participants who received no video treatment. The experimental group that received anger displays did not differ significantly from the other experimental and control conditions of positive and neutral emotional expressions.

Sigmar Gabriel's displays of anger have a positive impact on his leadership skill evaluations according to the Bonferroni adjusted p-value. Compared to the evaluation of his other character traits, such a positive effect had not occurred regarding any other of his evaluations. In this instance, displays of anger increased his leadership ratings on average by 0.12 points and differed significantly from the control group, which received no video treatment and has an overall mean of -0.05 (-0.048 = 0.119 - 0.167). Sigmar Gabriel's leadership skill ratings did not differ significantly between the experimental condition that received angry displays and the control group which received neutral emotional expressions. Hence, the positive impact might be linked to his general appearance and conversation style.

Table 26: Changes in the Evaluation of Leadership Skills

	Male Politicians (1)	Female Politicians (2)	Merkel (E2) (3)	Gysi (E2) (4)	Gabriel (E3) (5)	Merkel (E3) (6)
	Dependent Variable: Leadership Skills					
Negative-Passive	0.090 (0.063)	0.057 (0.065)				
Neutral	0.084 (0.063)	0.092 (0.066)	-0.015 (0.057)	0.046 (0.069)	-0.046 (0.068)	-0.003 (0.058)
No Video	-0.005 (0.062)	-0.008 (0.064)	-0.079 (0.055)	-0.147* (0.068)	-0.167^b (0.067)	-0.112* (0.057)
Positive	0.092 (0.064)	0.134* (0.065)	0.039 (0.057)	0.095 (0.070)		
Constant	0.043 (0.045)	0.045 (0.047)	-0.012 (0.040)	0.182*** (0.049)	0.119^b (0.049)	0.021 (0.041)
Observations	1,763	1,766	1,374	1,373	1,052	1,053
Adjusted R^2	0.001	0.002	0.001	0.008	0.004	0.003
F Statistic	1.287	1.775	1.595	4.811**	3.358*	2.592

Note: $^*p < 0.05$; $^bp <$ Bonferroni adjusted; $^{**}p < 0.01$; $^{***}p < 0.001$. The reference category is the negative-active condition. Cells display OLS-estimates and standard errors in parentheses.

Figure 21: Changes in the Evaluation of Leadership Skills

Note: Figure displays coefficient plots with 95-percent confidence intervals.

5.1.4 The Evaluations of Specific Candidate Impressions

In addition to these four classic items of candidate orientation, a range of individual attributes was measured by five semantic differential scales. These semantic differentials are also analyzed based on their gain scores, the differences between the first and second panel wave. These measurements of impression formations are potentially stronger related to the experimental treatment, as similar measures have been used as treatment checks for displays of incivility (Mutz & Reeves 2005: 14). The following section describes the impact the experimental treatment has on the impression formations according to the range of semantic differentials.

5.1.4.1 Evaluations of Emotionality

A significant effect occurred for all politicians and groups of politicians (see Table 27, see also Figure A.5 in the online appendix). The strongest effects occurred for Gysi and Gabriel with approximately 5 percent of explained variance for each model (adjusted R^2 = 0.044 and adjusted R^2 = 0.053). All politicians were assessed as being more emotional than rational by the participants after the experimental treatment. This effect is highly significant in almost all instances, regardless of whether the other participants saw no video or videos with neutral and positive emotional content – with the exception of Angela Merkel. Her negative-active emotional expressions differed only significantly from the condition without video treatment and not her neutral or positive expressions in sub-experiment *Type 2*.

Table 27: Changes in the Evaluation of Emotionality

	Male Politicians (1)	Female Politicians (2)	Merkel (E2) (3)	Gysi (E2) (4)	Gabriel (E3) (5)	Merkel (E3) (6)
	Dependent Variable: Emotional–Rational					
Negative-Passive	0.886*** (0.106)	0.357*** (0.106)				
Neutral	0.730*** (0.105)	0.358*** (0.108)	0.161 (0.109)	0.852*** (0.124)	0.852*** (0.110)	0.273[b] (0.116)
No Video	0.941*** (0.104)	0.657*** (0.105)	0.415*** (0.106)	0.859*** (0.120)	0.431*** (0.108)	0.606*** (0.114)
Positive	0.439*** (0.106)	0.356*** (0.106)	-0.048 (0.077)	0.632*** (0.124)		
Constant	-0.718*** (0.075)	-0.434*** (0.077)	-0.351*** (0.077)	-0.864*** (0.087)	-0.423*** (0.078)	-0.542*** (0.083)
Observations	1,761	1,766	1,372	1,370	1,049	1,052
Adjusted R^2	0.056	0.019	0.015	0.044	0.053	0.025
F Statistic	27.317***	9.712***	7.816***	21.885***	30.116***	14.296***

Note: *p < 0.05; [b]p < Bonferroni adjusted; **p < 0.01; ***p < 0.001. The reference category is the negative-active condition. Cells display OLS-estimates and standard errors in parentheses.

5.1.4.2 Evaluations of Politeness

Politicians' emotional expressions have an impact on the assessment of their politeness, since all F-tests indicate significant differences between the experimental conditions (Table 28, see also Figure A.6 in the online ap-

pendix). Compared to the previous measurements on candidate orientation, this analysis shows that a slightly larger portion of the variance in the dependent variable can be explained by the experimental treatment. Overall, the amount of explained variance is still very small and ranges in most instances from 0.4 percent (for male politicians) to 1.1 percent for Gregor Gysi, with one exception: The adjusted R^2 for Sigmar Gabriel explains 7.7 percent of variability in the change of his politeness score. This is in line with the expectations, since his incivility should have affected evaluations of politeness more strongly than mere displays of negative-active emotions that are not necessarily linked to impolite behavior towards others. Nevertheless, politicians in general as well as Sigmar Gabriel and Angela Merkel in particular were assessed as being more impolite after participants saw their negative-active emotional displays. Only the participants who saw Gregor Gysi's negative-active expressions did not evaluate him as being more impolite after having seen his displays of anger and indignation. The mean for this experimental group did not differ significantly from zero or the no video condition. However, a difference in means occurred compared to the negative-active treatment group, as he was evaluated as being more polite by participants who were in the neutral and positive treatment conditions.

Table 28: Changes in the Evaluation of Politeness

| | *Dependent Variable:* Impolite–Polite | | | | | |
	Male Politicians (1)	Female Politicians (2)	Merkel (E2) (3)	Gysi (E2) (4)	Gabriel (E3) (5)	Merkel (E3) (6)
Negative-Passive	0.181	0.286**				
	(0.096)	(0.098)				
Neutral	0.189*	0.230*	0.062	0.237**	0.819***	0.207[b]
	(0.095)	(0.099)	(0.079)	(0.090)	(0.093)	(0.081)
No Video	0.296**	0.346***	0.181*	-0.075	0.686***	0.260**
	(0.094)	(0.097)	(0.076)	(0.088)	(0.092)	(0.080)
Positive	0.204*	0.412***	0.195[b]	0.220[b]		
	(0.096)	(0.098)	(0.079)	(0.090)		
Constant	-0.178**	-0.228**	-0.195***	0.086	-0.676***	-0.274***
	(0.068)	(0.071)	(0.055)	(0.063)	(0.066)	(0.058)
Observations	1,762	1,765	1,372	1,370	1,049	1,052
Adjusted R^2	0.004	0.009	0.004	0.011	0.077	0.009
F Statistic	2.586*	5.134***	2.937*	6.252***	44.519***	5.842**

Note: *p < 0.05; [b]p < Bonferroni adjusted; **p < 0.01; ***p < 0.001. The reference category is the negative-active condition. Cells display OLS-estimates and standard errors in parentheses.

5.1.4.3 The Evaluation of Agitation

Again, all models are highly significant (Table 29, see also Figure A.7 in the online appendix). All politicians are rated as being more agitated after participants saw their negative-active emotional expressions. This effect is highly significant and occurred in all instances in comparison with the no video-condition, except for Sigmar Gabriel. In all instances the effect also occurred in comparison to the experimental groups that received neutral emotional expressions. Throughout the models the adjusted R^2 values are very small ranging from an explained variance of 0.7 percent for female politicians to 3.8 percent for Gregor Gysi.

Table 29: Changes in the Evaluation of Agitation

	Dependent Variable: Calm–Agitated					
	Male Politicians (1)	Female Politicians (2)	Merkel (E2) (3)	Gysi (E2) (4)	Gabriel (E3) (5)	Merkel (E3) (6)
Negative-Passive	-0.517*** (0.102)	-0.149 (0.103)				
Neutral	-0.271** (0.102)	-0.323** (0.104)	-0.473*** (0.094)	-0.699*** (0.112)	-0.674*** (0.111)	-0.466*** (0.098)
No Video	-0.577*** (0.100)	-0.379*** (0.102)	-0.580*** (0.091)	-0.542*** (0.109)	-0.129 (0.109)	-0.570*** (0.096)
Positive	-0.297** (0.103)	-0.237* (0.103)	-0.441*** (0.094)	-0.759*** (0.112)		
Constant	0.486*** (0.072)	0.287*** (0.074)	0.563*** (0.066)	0.566*** (0.079)	0.226** (0.079)	0.554*** (0.070)
Observations	1,761	1,766	1,372	1,370	1,049	1,052
Adjusted R^2	0.021	0.007	0.030	0.038	0.037	0.034
F Statistic	10.258***	4.248**	15.256***	19.161***	21.000***	19.441***

Note: *p < 0.05; [b]p < Bonferroni adjusted; **p < 0.01; ***p < 0.001. The reference category is the negative-active condition. Cells display OLS-estimates and standard errors in parentheses.

5.1.4.4 The Evaluation of Aggressiveness

An overall pattern emerged when looking at the impact the experimental treatment of anger has on the evaluations of politicians as being aggressive versus peaceful (Table 30, see also Figure A.8 in the online appendix). All models show significant effects in the same direction: Politicians in general, Angela Merkel, Sigmar Gabriel and Gregor Gysi are rated as being more aggressive and less peaceful when participants received their negative-ac-

tive emotional displays. Model 5 deals with Sigmar Gabriel and shows the largest R^2 with explained variance of 6.1 percent, followed by the models for Angela Merkel which have an R^2 of 4.2 percent when participants also see Sigmar Gabriel and 2.2 percent when participants also see Gregor Gysi. The models regarding politicians in general can hardly explain any variance. The evaluation of Gregor Gysi can be seen as a slight exception to this overall pattern. The effect of his negative-active emotions is not significantly different from the no video condition, it only differs significantly from his neutral and positive emotional displays, in which conditions he was rated as more peaceful.

Table 30: Changes in the Evaluation of Aggressiveness

	Dependent Variable: Aggressive–Peaceful					
	Male Politicians (1)	Female Politicians (2)	Merkel (E2) (3)	Gysi (E2) (4)	Gabriel (E3) (5)	Merkel (E3) (6)
Negative-Passive	0.243[b] (0.097)	0.157 (0.097)				
Neutral	0.325*** (0.096)	0.179 (0.099)	0.244** (0.084)	0.400*** (0.095)	0.753*** (0.091)	0.553*** (0.087)
No Video	0.325*** (0.095)	0.223* (0.097)	0.406*** (0.082)	0.159 (0.093)	0.477*** (0.089)	0.484*** (0.085)
Positive	0.213* (0.097)	0.240* (0.098)	0.391*** (0.084)	0.495*** (0.096)		
Constant	-0.390*** (0.068)	-0.287*** (0.070)	-0.425*** (0.059)	-0.189** (0.067)	-0.565*** (0.064)	-0.503*** (0.062)
Observations	1,763	1,766	1,372	1,370	1,049	1,052
Adjusted R^2	0.006	0.002	0.020	0.022	0.061	0.042
F Statistic	3.851**	1.886	10.252***	11.214***	35.228***	24.264***

Note: *p < 0.05; [b]p < Bonferroni adjusted; **p < 0.01; ***p < 0.001. The reference category is the negative-active condition. Cells display OLS-estimates and standard errors in parentheses.

5.1.4.5 The Evaluation of Arrogance

A similar overall trend occurred in the six models that analyze changes in the evaluation of arrogance versus modesty (Table 31, see also Figure A.9 in the online appendix). The overall model in which participants saw female politicians is not significant (Model 2). When participants were exposed to male politicians a treatment effect of anger expressions took place; participants rated politicians in general on average as being more modest in comparison to the control group without video treatment

(Model 1). A negative effect did occur however in the models for Angela Merkel and Sigmar Gabriel, whereby they were both evaluated as being more arrogant due to their anger expressions (Models 3,5, and 6).

In contrast to Angela Merkel and Sigmar Gabriel, Gregor Gysi was not evaluated as being more arrogant, since the negative emotional displays did not lead to an unfavorable evaluation in his case (Model 4). On the contrary, participants in the anger condition considered him even as more modest or less arrogant on average after they received the experimental treatment. This effect differs from the control group without video treatment and is also significantly different from his positive appearances; however, it does not differ significantly from his neutral appearances. Because his positive emotional displays did not lead to more modest evaluations, mere exposure effects cannot fully explain the occurrence of these positive effects.

Table 31: Changes in the Evaluation of Arrogance

	Dependent Variable: Arrogant–Modest					
	Male Politicians (1)	Female Politicians (2)	Merkel (E2) (3)	Gysi (E2) (4)	Gabriel (E3) (5)	Merkel (E3) (6)
Negative-Passive	0.184* (0.081)	-0.027 (0.078)				
Neutral	0.196* (0.080)	-0.067 (0.080)	0.202 [b] (0.082)	0.015 (0.083)	0.679*** (0.085)	0.238** (0.080)
No Video	0.221** (0.079)	0.007 (0.078)	0.259** (0.079)	-0.349*** (0.081)	0.489*** (0.083)	0.204** (0.079)
Positive	0.150 (0.081)	0.088 (0.079)	0.427*** (0.082)	-0.287*** (0.084)		
Constant	-0.106 (0.057)	0.108 (0.057)	-0.277*** (0.058)	0.268*** (0.059)	-0.527*** (0.060)	-0.223*** (0.057)
Observations	1,762	1,766	1,373	1,369	1,049	1,052
Adjusted R^2	0.003	0.0002	0.018	0.021	0.059	0.008
F Statistic	2.416*	1.067	9.234***	10.695***	34.036***	5.176**

Note: $^*p < 0.05$; $^b p <$ Bonferroni adjusted; $^{**}p < 0.01$; $^{***}p < 0.001$. The reference category is the negative-active condition. Cells display OLS-estimates and standard errors in parentheses.

5.1.4.6 The Evaluation of Gender Stereotypical Attributes

Two more semantic differentials are considered which measured attributes that are closely linked to gender stereotypical expectations. These items were only measured in the sub-experiment *Type 1*, which focused on the

evaluation of politicians as a social group. Hence, the following section only concerns politicians in general.

5.1.4.6.1 The Evaluation of Decisiveness

According to the nonsignificant F-tests of both models, no differences in means occurred across the experimental conditions regarding the dependent variable that measured changes in the evaluation of decisiveness (see Table 32). Emotional expressions of anger and indignation by female as well as male politicians did not affect the assessment of decisiveness for politicians in general in the experiment.

Table 32: Changes in the Evaluation of Decisiveness

	Dependent Variable: Decisive–Unassertive	
	Male Politicians	Female Politicians
	(1)	(2)
Negative- Passive	0.100	-0.018
	(0.102)	(0.101)
Neutral	0.113	-0.012
	(0.101)	(0.102)
No Video	0.022	-0.021
	(0.100)	(0.100)
Positive	-0.057	-0.044
	(0.103)	(0.101)
Constant	-0.172*	-0.129
	(0.072)	(0.073)
Observations	1,762	1,766
Adjusted R^2	-0.0001	-0.002
F Statistic	0.964	0.052

Note: *p < 0.05; ^{b}p < Bonferroni adjusted; **p < 0.01; ***p < 0.001. The reference category is the negative-active condition. Cells display OLS-estimates and standard errors in parentheses.

5.1.4.6.2 The Evaluation of Resilience

Looking at the second set of attributes that relate to gender stereotypes, the F-test of model 1 indicates a significant difference between the means of the experimental groups after participants saw the video treatments of male politicians (see Table 33). However, only the control group without video treatment differed from the group which showed displays of negative-active emotions significantly at a 1-percent level; this effect was ob-

202

tained because a positive change occurred within the control group, while no change was induced in the experimental group. Hence, this significant finding might not be substantially related to the display of negative emotions. Furthermore, the effect sizes were very small (adjusted $R^2 = 0.003$). To sum up, the effects emotional displays – more specifically anger – have on the evaluation of politicians in general are negligible in this instance. Assessments of politicians as a social group, such as their decisiveness and resilience, which are related to gender stereotypes of males and females were not affected substantially by politicians' anger expressions. A more direct effect might have occurred if instead of measuring assessments of politicians in general, participants had evaluated female and male politicians separately before and after the treatment occurred.

Table 33: Changes in the Evaluation of Resilience

	Dependent Variable: Resilient–Overstrained	
	Male Politicians (1)	Female Politicians (2)
Negative- Passive	0.184[*]	-0.027
	(0.081)	(0.078)
Neutral	0.196[*]	-0.067
	(0.080)	(0.080)
No Video	0.221[**]	0.007
	(0.079)	(0.078)
Positive	0.150	0.088
	(0.081)	(0.079)
Constant	-0.106	0.108
	(0.057)	(0.057)
Observations	1,762	1,766
Adjusted R^2	0.003	0.0002
F Statistic	2.416[*]	1.067

Note: [*]$p < 0.05$; [b]$p <$ Bonferroni adjusted; [**]$p < 0.01$; [***]$p < 0.001$. The reference category is the negative-active condition. Cells display OLS-estimates and standard errors in parentheses.

5.1.4.7 The Evaluation of External Efficacy

In addition, to these two semantic differentials, the concept of external efficacy was also measured for politicians in general. These items were linked to the warmth of politicians as the items regard how interested politicians seem in ordinary people, whether they care and seek close contact. The experimental treatment of negative-active emotional displays had no effect on the evaluation of politicians' external efficacy, as can be seen in

the following models in Table 34. Therefore, the construct of external efficacy is not influenced by anger expressions of groups of male and female politicians. The next section summarizes the overall effects.

Table 34: Changes in the Evaluation of External Efficacy

	Dependent Variable:			
	Contact Seeking		Caring	
	Male Politicians (1)	Female Politicians (2)	Male Politicians (3)	Female Politicians (4)
Negative-Passive	0.026 (0.060)	-0.014 (0.058)	0.063 (0.059)	-0.003 (0.056)
Neutral	0.040 (0.059)	0.071 (0.059)	0.012 (0.058)	0.023 (0.057)
No Video	-0.005 (0.059)	0.034 (0.058)	0.082 (0.058)	0.023 (0.056)
Positive	0.065 (0.060)	0.209*** (0.058)	0.002 (0.059)	0.053 (0.056)
Constant	0.000 (0.042)	-0.039 (0.042)	-0.023 (0.041)	0.036 (0.040)
Observations	1,762	1,767	1,760	1,766
Adjusted R^2	-0.001	0.009	-0.0003	-0.002
F Statistic	0.473	4.887***	0.867	0.333

Note: *p < 0.05; [b]p < Bonferroni adjusted. The reference category is the negative-active condition. Cells display OLS-estimates and standard errors in parentheses.

5.1.5 Summary of the Main Effects

The following Table 35 shows the nature of each significant effect that anger has on the evaluation of the respective politician(s). The items that measure the general favorability of a politician as well as more specific character traits measured by the candidate orientations are overall minimally affected by the experimental treatment. Only the experiment *Type 2*, in which participants saw Angela Merkel and Gregor Gysi, show consistent effects. Gregor Gysi was generally rated more favorably in every regard, while Angela Merkel was rated less favorably, when participants also saw Gregor Gysi's negative-active emotions. However, her negative emotional displays did not affect the evaluation of her leadership skills. In contrast, Sigmar Gabriel's leadership skills were the only characteristic that was positively affected by his displays of anger and incivility.

This finding supports the first hypothesis about varying effects of anger expressions according to their position within the political system and more specifically the hypothesis $H1_a$, whereby displays of anger can lead to

higher ratings for politicians of the opposition. The hypothesis H1$_b$ about negative emotional expressions of incumbents is only partially supported – less favorable evaluations of Angela Merkel only occurred when participants also saw Gregor Gysi.

The hypothesis H4 assumed that anger expressions had the strongest effects on the assessment of politicians in comparison to the other expressions; however, the empirical findings do not indicate a consistent pattern. Gregor Gysi's anger expressions had the strongest effects compared to positive or neutral appearances, in certain – but not all – instances, such as his overall assessment and ratings of trustworthiness. Hence, at least based on these two instances the hypothesis H4 about the strength of anger expressions is partially supported by the experiment.

Furthermore, the inspection of the average treatment effects presents insights about the evaluations of warmth and competence. The evaluation of warmth and competence are in fact affected differently as stated in hypothesis H2$_{(a-b)}$. The hypothesis H2$_a$ is supported by the findings for Gregor Gysi and Sigmar Gabriel, whereby their competence was assessed more positively after the experimental treatment was administered. However, the hypothesis H2$_b$ about negative effects on the politicians' warmth is hardly supported by the empirical evidence.

By focusing on the four typical items to assess candidate appearance effects, it can be seen that the perception of politicians as a social group was minimally affected by the experimental treatment of anger expressions. Hence, this finding supports the hypothesis H11, whereby politicians as a social group are less strongly affected by the experimental treatment than the evaluation of individual political leaders, although they are well-known public figures.

The candidate evaluations that are measured by semantic differentials and are therefore closely related to the experimental treatment show more consistent effects for all three sub-experiments. Displays of negative-active emotions appeared to impact the evaluation of politicians negatively according to the semantic differentials – except for Gregor Gysi. Unlike the other politicians and politicians in general, he is neither rated as being more arrogant nor as being more impolite by participants who were exposed to his negative-active emotions. On the contrary, he was seen as being more modest (and less arrogant) by those who saw his negative-active emotional displays.

These differences occurred according to the semantic differential scales and point to the fact that Gregor Gysi's emotional expressions can be seen as an exception. These positive evaluations highlight that the evaluation of

anger expressions might strongly depend on the context – and in this case the political message. So far, it strengthens the assumption that the positive side of anger as moral outrage can have positive effects on the evaluation of a politician, exemplified by Gregor Gysi's anger about social injustices. This finding is in contrast to the effects of uncivil behavior, as it was shown by Sigmar Gabriel. By comparing these two politicians as two case studies, hypothesis H3 about the political message as contextual factors gains support (see Subchapter 4.5.3 for further discussions on the political messages).

Table 35: Overview of Treatment Effects According to the Change in Each Variable from Pre- to Post-Test

	Male Politicians	Female Politicians	Angela Merkel (E2)	Gregor Gysi (E2)	Sigmar Gabriel (E3)	Angela Merkel (E3)
Scalometer	ns	ns	-	+	ns	ns
Likeability	ns	ns	-	+	ns	ns
Trustworthiness	ns	ns	-	+	ns	ns
Problem-Solving Skills	ns	ns	-	+	ns	ns
Leadership Skills	ns	ns	ns	+	+	ns
Emotional-Rational	-	-	-	-	-	-
Rude-Polite	-	-	-	ns	-	-
Calm-Agitated	+	+	+	+	+	+
Aggressive-Peaceful	-	-	-	(-)	-	-
Arrogant-Modest	(-)	ns	-	+	-	-

Note: The table shows significant effects at least at a 5-percent level, if the coefficient for negative-active emotions as well as the difference-in-difference comparison between the anger condition and the control condition without video treatment is significant. Brackets indicate that only one criterion is met at the 1-percent level, "-" indicates negative effects, "+" indicates positive effects, and the abbreviation "ns" stands for "not significant".

The effects on the additional items that were measured within the experiment *Type 1* which focused on the gender stereotypical evaluation of politicians in general are summarized in Table 36 below. The evaluation of politicians' decisiveness and resilience remained unaffected by the experimental treatment of negative-active emotions of male and female politicians.

The external efficacy was also not affected by the experimental treatment of negative-active emotions. It has to be noted that any of these effects would technically be spillover effects from the specific politicians in the

video material that participants were exposed to on politicians in general. It might be more likely to detect effects when it comes to the three political leaders, because they appeared within the video clips and their evaluations were measured directly.[78] Since the effects on the evaluation of specific politicians show larger effects overall, it might be worthwhile to analyze external efficacy in the future after anger expressions of specific politicians.

Table 36: Overview of Effects of Additional Variables for Politicians in General

	Female Politicians	Male Politicians
Resilient-Overstrained	ns	ns
Unassertive-Decisive	ns	ns
Caring	ns	ns
Contact Seeking	ns	ns

Note: The table shows significant effects at least at a 5-percent level, if the coefficient for negative-active emotions as well as the difference-in-difference comparison between the anger condition and the control condition without video treatment is significant. Brackets indicate that only one criterion is met at the 1-percent level, "-" indicates negative effects, "+" indicates positive effects, and the abbreviation "ns" stands for "not significant".

As previously mentioned in Subchapter 4.4.1, network structures of candidate evaluations can be used to assess attitudes towards candidates. Network structures can be compared between time points to detect changes in the network structures (Dalege et al. 2017). The largest treatment effects occurred in regard to the evaluation of Gregor Gysi, whereas the effects for the other two leading politicians and politicians in general had only a minor impact on attitudes towards them in a few instances. If these treatment effects led to changes in the relationship between the evaluations, the network structure could reflect such shifts between variables. Those changes in the network structure are most likely to have occurred for Gregor Gysi compared to the other cases, which hardly showed substantial variances in the evaluations before and after the treatment. Hence, the case of Gregor Gysi can be seen as a most-likely case for the analysis of changes in the network structure and was therefore analyzed.

A network comparison test was conducted in order to detect significant changes between the pre-test and post-test condition (Dalege et al. 2017;

---

78 The items for external efficacy were not measured for the three individual politicians due to the length of the questionnaire.

van Borkulo 2016). This test was based on 340 participants who took part in both panel waves, were exposed to Gysi's anger expressions and answered all 15 items under consideration. The network comparison test indicated that both models did not differ significantly regarding the structure in variance of the evaluations after conducting 1000 permutations (M = 1.58, p = 0.381). Since no invariance could be detected between the networks, a further test for invariances of edges between two nodes was not considered necessary (for comparison see: van Borkulo et al. 2017). In addition, significant changes in the global strength between network models did not occurred (S = 0.35, p = 0.921). Therefore, it can be concluded that Gregor Gysi was evaluated more positively across all characteristics without changing the structure between the attitudes towards him. Since no changes in the network structure could be obtained for Gregor Gysi, it appears unlikely that the network structure of other politicians would have changed as a result of the experimental treatment.

The average treatment effects of anger expressions showed varying effects for Angela Merkel in experiment *Type 2* and *Type 3*. These differences of contrasting videos could have been further impacted by the order in which politicians were seen. Hence, the order of the video clips can be considered in the following subchapter. In addition to the order of the video clips, individual predispositions can be tested in regard to their moderating effect on the evaluation.

5.2 Overview of Potential Moderating Factors

This subchapter presents an overview of the moderating effects that were tested for the experimental treatment. The corresponding tables and figures can be found in the online appendix. The additional consideration of moderating factors showed only minor differences for varying subgroups compared to the average treatment effect. An overview of the effects is presented in Table 37. As dependent variables, gain scores of the scalometer ratings were used as overall summary scores as well as warmth and competence ratings. In each model, effects of age, gender, party identification, and political interest were controlled for as possible confounding factors, while an interaction between the following factors and the experimental treatment was also included in the regression models.

The Order of the Experimental Treatment

First, the order in which the political leaders were seen had a small impact on the evaluation of the three politicians, whereby it was generally more favorable – or at least not disadvantageous – for politicians to been seen first. This finding is in accordance with a priming effect and thus supporting the hypothesis H12 regarding the order of the video treatment and anchoring effects of politicians who appeared first.

Regarding changes in Angela Merkel's scalometer ratings, the model showed a positive treatment effect of Merkel's positive emotional expressions for those participants who saw her first. The reference category in these models is the control group that featured neutral expressions. Since anger as well as neutral expressions led to slightly lower overall evaluations of Angela Merkel, the coefficient for the anger condition is not statistically significant in comparison to the neutral control condition. Furthermore, the treatment order had no impact on her overall evaluation.

The evaluation of her warmth shows a significant effect of the treatment order, whereby her warmth is evaluated on average lower by participants who saw Merkel in the second video after Gregor Gysi's expressions. Hence, her anger expressions had a significant negative effect for those participants who saw her last.

The order of the video clips did not have a significant effect on the evaluation of Gregor Gysi; the treatment effects remained stable regardless of whether he is seen first or second. For Sigmar Gabriel, only the model that focused on the assessment of his warmth is statistically significant; however, the order of the experimental treatment is not significant as a main or interaction effect. His negative-active emotional displays led to lower ratings when he was seen after Merkel, while no effect occurred when he was seen first. The model for his competence evaluations is not significant because none of the coefficients differ significantly from each other, but it can be noted that it appears to be slightly more beneficial for him to be seen last, since those participants evaluated him more positively afterwards. Despite the lack of statistical significance, the overall tendency is consistent across the dependent variables, in the sense that it was more beneficial for Merkel and Gabriel to be seen first, and less beneficial to be seen second.

Party Identification

Second, the consideration of party identification showed that ceiling effects likely occurred for the evaluation of Angela Merkel for CDU supporters, and in part for the overall sample. Angela Merkel's evaluations were

209

the highest of the three politicians before the experimental treatment was administered; therefore, a further increase in positive evaluations was hardly achievable. For this subchapter, party identification was measured as a dichotomous variable indicating whether participants identified with a certain political party. For politicians in general, party identification was measured as whether participants identified with any political party (1) or no party at all (0). For the three political leaders – Angela Merkel, Gregor Gysi and Sigmar Gabriel – party identifiers were classified as such if they identified with the CDU/CSU for Angela Merkel, the Left for Gregor Gysi and the SPD for Sigmar Gabriel (1). Participants who did not identify with these parties where classified as non-identifiers (0). It was not further differentiated whether participants supported any other political party or no party at all.

For Gregor Gysi, the ceiling effect was limited to supporters of the Left, who showed high evaluations of Gysi before the treatment was administered, thereby creating little room for a more favorable evaluation. However, participants who did not support the Left were not affected by such a ceiling effect and showed on average a positive response to Gysi's anger expressions. The analysis of Sigmar Gabriel's incivility showed that supporters remained unaffected by his expressions of incivility, which had a negative effect on other voters. However, the absence of moderating effects of party identification could be explained by ceiling effects, at least for Angela Merkel and Gregor Gysi. The evaluation of politicians as social groups was not moderated by levels of individual party attachments; hence, the empirical analysis showed mixed results for the impact of party identifications. The hypothesis $H5_a$, whereby party supporters responded more favorably to anger expressions of their leaders, could not be supported. The hypothesis $H5_b$, whereby negative responses to anger expressions of opposing political leaders had a negative impact of non-supporters, was partially supported by the findings of Sigmar Gabriel's expressions on the evaluation of his warmth. His supporters' evaluations of him remained unaffected by his displays of incivility.

The Order of the Experimental Treatment and Party Identification
In addition to moderating effects for the order of the experimental treatment and party identification, three-way-interactions between the experimental treatments, its order and party identification were tested for the three politicians. Focusing on Merkel's overall assessment, it can be seen that those who did not identify with the Christian Democratic Union and were also first exposed to her video clip, evaluated her more positively (see

Figure 22 and Table A.28 in the online appendix). Furthermore, and more to the point regarding negative-active emotional expressions, those who did not identify with the CDU evaluated her less favorably after seeing her negative-active emotional expressions. This effect does not differ significantly from the control group with neutral expressions.

The model that focuses on her warmth also points towards a moderating effect of party identification and the order of the experimental treatment. Those participants who saw Angela Merkel before Gregor Gysi responded more favorably towards her positive expressions – if they did not identify with the CDU. On the contrary, those who did not identify with the CDU and saw her after they were exposed to Gregor Gysi's anger and indignation, evaluated her less favorably. This negative effect neither occurred for those who identified with the CDU nor for those non-identifiers that saw her anger expressions first.

Lastly, the models for Angela Merkel's competence ratings are not significant regardless of whether participants also saw Gregor Gysi or Sigmar Gabriel in addition to her. Nonetheless, the coefficient for those who saw her anger expressions after those of Gregor Gysi and who did not identify with the CDU showed a slightly lower evaluation of her competence. While this coefficient is only significant at a 10-percent level, it shows a potentially consistent negative effect of being seen after other politicians, as indicated by the previous two models. This effect is specific to the co-occurrence of politicians. When participants were also exposed to Sigmar Gabriel, no significant effects occur (see Figure A.15 in the online appendix).

Figure 22: *Interaction between the Experimental Treatment, Treatment Order and Party Identification for the Evaluation of Angela Merkel (With Exposure to Gregor Gysi)*

Note: Figure displays coefficient plots with 95-percent confidence intervals.

To summarize the effects of three-way-interactions between the experimental treatments, its order and party identification, it can be said that the order of the video clips had the strongest effect on the evaluation of Angela Merkel, at least when participants also saw Gregor Gysi. While Angela Merkel is not only evaluated in comparison to Gregor Gysi, it mattered in which order she appeared. Overall, it might be more favorable for her ratings to have appeared first, as they did not result in more negative evaluations. When she appeared after participants had already been exposed to

Gregor Gysi, her negative-active expressions resulted in less favorable evaluations of her, particularly her warmth.

For Sigmar Gabriel and Angela Merkel, it was more advantageous appearing first, while being seen second or last, resulted in less positive changes. Appearing second, might induce evaluations and judgments in relation to the previous politicians whose appearances might act as an anchor heuristic or priming effect. This subchapter analyzed party identification by using a dichotomous measurement of support. Since the effects on non-supporters were counter-intuitive, it seems worthwhile to test the individual ideological predispositions of participants further to shed more light on the evaluation of politicians and political leaders. The ideological disposition of participants is therefore taken into consideration. This might offer a more finely grained approach that is more suited to parliamentary systems in which citizens might react positively towards politicians of one ideological bloc – consisting of several political parties – rather than one party. Alongside ideological predispositions, further individual personality traits, such as neuroticism, might affect how participants respond to the experimental treatment

Personality Traits
Finally, the consideration of further individual predispositions such as self-positions on the socioeconomic left-right continuum and personality traits hardly enhanced the regression models. The ideological self-positioning was only significant for the evaluations of Gregor Gysi and Sigmar Gabriel. It is likely that another ceiling effect occurred for those who positioned themselves ideologically as leftists and evaluated Gregor Gysi, which resulted in a more positive evaluation with an increase in socioeconomic right-wing attitudes. The moderating effects for Sigmar Gabriel mitigated the negative effect of his anger expressions for those who showed more leftist socioeconomic positions. The hypothesis H6, which stated that the individual response to anger expressions is moderated by individual personality traits such as neuroticism, was not supported by the experimental findings. Likewise, individual levels of extraversion did not moderate the effect of anger expressions on subsequent candidate evaluations.

213

Table 37: Overview of Main Effects and Moderating Effects for Politicians' Anger Expressions

	Male Politicians	Female Politicians	Angela Merkel (E2)	Gregor Gysi (E2)	Sigmar Gabriel (E3)	Angela Merkel (E3)
The Order of the Treatment						
Overall Evaluation			✓	✗	✗	✗
Warmth			✓	✗	✗	✗
Competence			✗	✗	✗	✗
Party Identification						
Overall Evaluation	✗	✗	✗	✓	✗	✗
Warmth	✗	✗	✗	✗	✓	✗
Competence	✗	✗	✗	✗	✗	✗
The Treatment Order and Party Identification						
Overall Evaluation	✗	✗	✗	✗	✗	✗
Warmth	✗	✗	✓	✗	✗	✗
Competence	✗	✗	✓	✗	✗	✗
Socioeconomic Position						
Overall Evaluation	✗	✗	✗	✓	✓	✗
Warmth	✗	✗	✗	✓	✗	✗
Competence	✗	✗	✗	✓	✓	✗
Neuroticism						
Overall Evaluation	✗	✗	✗	✗	✗	✗
Warmth	✗	✗	✗	✗	✗	✗
Competence	✗	✗	✗	✗	✗	✗
Extraversion						
Overall Evaluation	✗	✗	✗	✗	✗	✗
Warmth	✗	✗	✗	✗	✗	✗
Competence	✗	✗	✗	✗	✗	✗

Note: Cells display whether the respective variable had a statistically significant impact on the assessment of politicians as a main effect or moderating effect of anger expressions. "✗" stands for no effect, while "✓" indicates an effect.

5.3 The Broader Implications of the Experimental Treatment
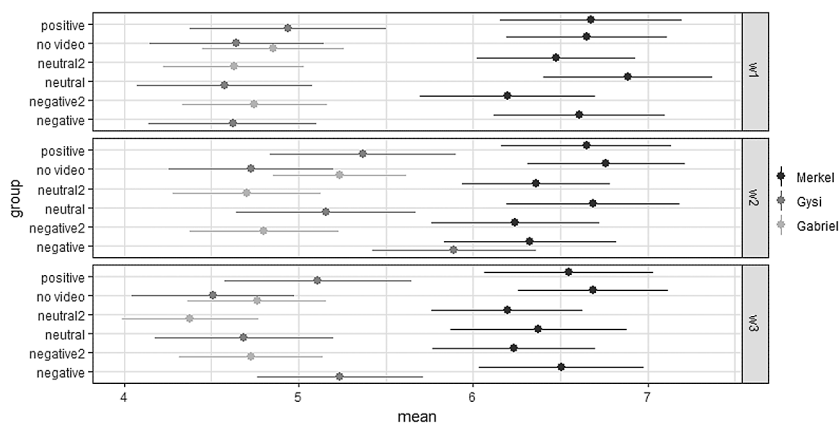
5.3.1 The Longevity of the Treatment Effects

This section tests hypothesis H9 about short-term effects by analyzing whether participants were affected by the experimental treatment, even one week after the experimental treatment was administered. The following Figure 23 displays the average overall evaluation of Angela Merkel,

Gregor Gysi and Sigmar Gabriel by those participants who took part in the survey experiment at all three time points (Waves 1–3).[79] Based on Figure 23, Angela Merkel was by far evaluated more positively than her male counterparts. Those who participated in all three questionnaires have a fairly high opinion of her, with a mean evaluation that averages approximately 6.58 in the first wave (SD = 2.72), 6.50 in the second wave (SD = 2.76), and 6.43 in the third wave (SD = 2.84).

Based on a graphical inspection, no major changes occurred for her evaluation over time. Furthermore, it can be seen that participants had a similar opinion of Gregor Gysi and Sigmar Gabriel during the pre-test (Wave 1) before the experimental treatment occurred. As previously analyzed, the experimental effect had the largest positive effect on the overall evaluation of Gregor Gysi, if participants saw his negative-active emotional expressions. During the second post-test this overall effect decreased slightly; however, participants still held a more favorable view of Gregor Gysi. This descriptive analysis was followed by a regression analysis to determine whether these changes are statistically significant. The dependent variable in this analysis was measured as a gain score between their overall evaluation in the third wave (t_3) and their overall evaluation during the pre-test (t_1).

---

79 It has to be noted that only half of the participants were invited to take part in the third questionnaire. The participants were randomly invited to the third wave.

Figure 23: Evaluation of Angela Merkel, Gregor Gysi, and Sigmar Gabriel in Three Panel Waves

Note: The number of observations: Gregor Gysi (N = 527), Sigmar Gabriel (N=410–411), Angela Merkel (N=790–791. "Neutral 2" and "Negative 2" indicate the groups in which participants saw Angela Merkel and Sigmar Gabriel.

As expected from the graphical display in Figure 23, the regression models for Angela Merkel and Sigmar Gabriel are not significant, since none of the groups vary significantly from each other. Only the model for Gregor Gysi is statistically significant ($F_{(3, 523)} = 2.65$, $p < 0.05$). This model shows that participants within the anger condition evaluated Gregor Gysi on average 0.75 points higher, even one week after the treatment occurred, in comparison to those who took part in the control group without video treatment. This evaluation indicates a decay compared to the initial treatment effect (see Model 2 in Table 38); nonetheless, a more positive evaluation has remained at t_3, one week after the initial exposure.

Table 38: *Longevity of Effects on the Overall Evaluation of Angela Merkel, Gregor Gysi and Sigmar Gabriel*

	Experimental Type	*Dependent Variable:* Scalometer		
		Merkel (1)	Gysi (2)	Gabriel (3)
Negative-Active	E2	-0.135	0.752***	
		(0.230)	(0.271)	
Positive	E2	-0.162	0.307	
		(0.240)	(0.283)	
Negative-Active 2	E3	-0.019		0.065
		(0.233)		-0.24
Neutral	E2	-0.546**	0.248	
		(0.236)	(0.278)	
Neutral 2	E3	-0.314		-0.155
		(0.232)		-0.239
Constant		0.034	-0.136	-0.095
		(0.160)	(0.189)	-0.165
Observations		790	527	410
Adjusted R^2		0.003	0.009	-0.003
F Statistic		1.468	2.645**	0.422

Note: *p < 0.1; **p < 0.05; ***p < 0.01. The reference category is the experimental group without video treatment. "Neutral 2" and "Negative 2" indicate the groups in which participants saw Angela Merkel and Sigmar Gabriel.

In addition to the overall evaluation of Gregor Gysi, the four items that measured the candidate orientation towards him are analyzed in more detail. Table 39 shows the longevity of the experimental treatment regarding Gregor Gysi's warmth and competence evaluations. Only two out of four models are statistically significant according to their overall F-tests – the models that measured Gregor Gysi's warmth. His competence ratings were not affected by his displays of anger and indignation one week after the treatment occurred. This includes the evaluation of his leadership skills ($F(3, 523) = 1.70$, $p = 0.166$), as well as his problem-solving skills ($F(3, 523) = 0.97$, $p = 0.406$). His likeability ratings were slightly affected by his displays of anger, since participants who saw his anger still rated him 0.32 points higher compared to those who saw no video, even one week after the experimental treatment was administered. Furthermore, his trustworthiness was also affected positively by his displays of anger and indignation by 0.20 points compared to those who saw no video. In this instance however, the coefficient is only significant at the 10 percent level, while the overall model is significant at a 5-percent level ($F(3, 523) = 2.94$, $p < 0.05$).

217

Table 39: Longevity of Effects on the Evaluation of Gregor Gysi's Leadership Qualities

	Dependent Variable:			
	Leadership (1)	Likeability (2)	Capability (3)	Trustworthiness (4)
Negative-Active	0.150	0.318[**]	0.156	0.202[*]
	(0.114)	(0.124)	(0.118)	(0.109)
Positive	0.267[**]	0.038	-0.027	-0.136
	(0.119)	(0.130)	(0.123)	(0.114)
Neutral	0.115	0.108	0.099	0.018
	(0.117)	(0.127)	(0.121)	(0.112)
Constant	-0.027	-0.020	0.061	0.102
	(0.080)	(0.086)	(0.082)	(0.076)
Observations	527	527	527	527
Adjusted R^2	0.004	0.009	-0.0002	0.011
F Statistic	1.700	2.525[*]	0.971	2.937[**]

Note: *p < 0.1; **p < 0.05; ***p < 0.01. The reference category is the experimental group without video treatment.

Contrary to the assumptions stated in hypothesis H9, this analysis partially supports the existence of prolonging effects – at least for Gregor Gysi. More long-lasting effects on the evaluation of Angela Merkel's and Sigmar Gabriel's warmth and competence evaluations were also analyzed; however, none of the models were statistically significant (see Table A.11 and Table A.12 in the online appendix). The varying effects are likely to be impacted by the varying media exposure of all three politicians and the negative predispositions held towards Gysi at the beginning of the survey experiment (Wave 1), as well as the quality of his anger expressions. In line with previous research (Lodge et al. 1995; Gerber et al. 2011), this finding shows a decay of the experimental treatment effects – even for the anger expressions of Gregor Gysi. Nonetheless, if emotional expressions leave a strong first impression, they can also influence the "online tally" of an already well-known politician. The enduring effects occurred for the evaluation of Gysi's warmth, which is in line with previous findings that the "online tally" of politicians is predominantly related to the evaluation of warmth (Laustsen & Bor 2017).

5.3.2 Spillover Effects on Political Parties

As hypothesis H10 states, the emotional expressions of key political figures could potentially influence how someone feels towards and evaluates the

respective political parties. Previous research has shown that not only do political parties hold certain trait ownerships of valence and positional issues such as the economy or security that can have an impact on the evaluation of a candidate in being more competent in those regards, but that the image of political parties can also be affected by party leaders and top candidates who shape the trait ownership of the party (e.g., Garzia 2017; Hayes 2005). Hence, this next section explores whether the experimental conditions showing Angela Merkel, Gregor Gysi and Sigmar Gabriel affected the evaluation of their respective political parties.

In order to test whether a spillover occurred, the scalometer for each party was used as a summary score to determine whether a more positive or negative evaluation has occurred after the experimental treatment was administered. In addition to the experimental treatment as the independent variable of interest, the models presented in the following controlled for individual factors such as age (mean-centered), gender, interest in politics (mean-centered) and party identification as control and moderating factors, since these factors have been shown to shape individuals' perceptions and stereotypical evaluations (Masters 1994; McDermott 1998; Pietraszewski 2016; West 2017). First, the empirical findings are presented for Angela Merkel as leader of the Christian Democratic Party, followed by the results for Gregor Gysi as leader of the Left, and Sigmar Gabriel as leader of the Social Democratic Party, at the time of the data collection.

The Evaluation of the Christian Democratic Party

A spillover effect based on Angela Merkel's negative emotional display could be tested with both sub-experiments in which participants saw either Gregor Gysi or Sigmar Gabriel. Both sub-experiments do not support an effect of Angela Merkel's negative expressions on the evaluation of the CDU (see Table 40). The experiment that also featured Gregor Gysi shows a small positive effect of the negative treatment condition with an increase of 0.22 points on an 11-point Likert scale; however, this effect is only significant at a 10 percent level. More clearly than her negative expressions, a significant spillover effect can be observed for her positive emotional expressions, which resulted in a more positive evaluation of the Christian Democratic Party by an increase of 0.56 points (Model 1).

In addition to the treatment effect, a moderating effect of party support was tested by analyzing potential differences for those participants who identified with the Christian Democratic Party. Such an interaction term did not result in significant changes in the evaluation of the CDU after seeing Angela Merkel's emotional expressions (Model 2).

219

When participants were additionally exposed to Sigmar Gabriel instead of Gregor Gysi, the experimental treatment conditions did not make a difference in the evaluation of the Christian Democratic Union. This failure to replicate the spillover effect of sub-experiment *Type 2* could indicate that not only are political leaders evaluated in relation to each other, but these contrasting effects also spillover to the evaluation of the respective political parties.

Table 40: Changes in the Evaluation of the CDU

	Dependent Variable: CDU Scalometer			
	E2	E2	E3	E3
	(1)	(2)	(3)	(4)
Negative-Active	0.216*	0.118	0.151	0.112
	(0.121)	(0.147)	(0.123)	(0.148)
Positive	0.558***	0.620***		
	(0.123)	(0.145)		
Neutral	-0.033	-0.142	0.126	0.101
	(0.122)	(0.149)	(0.122)	(0.147)
PID	-0.105	-0.214	-0.164	-0.231
	(0.094)	(0.183)	(0.108)	(0.185)
Female	0.362***	0.350***	0.108	0.108
	(0.093)	(0.093)	(0.107)	(0.107)
Age	-0.007**	-0.007**	-0.011***	-0.011***
	(0.003)	(0.003)	(0.003)	(0.003)
Political Interest	-0.009	-0.011	-0.096	-0.097
	(0.052)	(0.052)	(0.059)	(0.060)
Negative-Active*PID		0.303		0.124
		(0.261)		(0.265)
Positive*PID		-0.247		
		(0.272)		
Neutral*PID		0.323		0.081
		(0.260)		(0.264)
Constant	-0.292***	-0.254**	-0.151	-0.131
	(0.100)	(0.111)	(0.106)	(0.115)
Observations	1,371	1,371	1,051	1,051
Adjusted R^2	0.034	0.036	0.015	0.014
F Statistic	7.844***	6.092***	3.740***	2.828***

Note: $^{*}p < 0.1$; $^{**}p < 0.05$; $^{***}p < 0.01$. The reference category is the experimental group without video treatment. The abbreviation "PID" stands for "party identification".

The Evaluation of the Left
Similarly to the absence of spillover effects based on Angela Merkel's angry displays, the evaluation of the Left is not particularly affected by Gregor Gysi's anger expressions (see Table 41). While Gregor Gysi's anger expres-

sions led to a more positive evaluation of the Left, these effects also oc-
curred for conditions in which participants saw neutral as well as positive
expressions. Since the effect of his anger did not differ significantly from
the effects of his other emotional expressions, this finding further supports
the occurrence of an overall exposure effect rather than an effect caused by
specific emotional expressions, such as anger. When testing an interaction
effect between the experimental treatment and individual party identifica-
tion for those who identified themselves with the Left, exposure effects
were not significant; therefore, exposure to video material did not affect
the evaluation of the Left for their supporters. The only significant effect
remaining is merely significant at a 10-percent level, and indicating that re-
spondents rated the Left higher if they were in the neutral control condi-
tion and did not identify with the Left. Hence, this finding at best supports
a spillover effect that is due to exposure to key politicians rather than their
specific emotional expressions.

The Evaluation of the Social Democratic Party
Lastly, a spillover effect on the evaluation of the Social Democratic Party
after seeing Sigmar Gabriel's incivility was analyzed (Table 41). The regres-
sion model shows that a negative spillover occurred after participants saw
Sigmar Gabriel's negativity; however, this effect is only significant at the
10-percent level. His negative expressions resulted, on average, in a less fa-
vorable evaluation of the SPD: by -0.21 points. This effect is significantly
different from the control condition without video and the neutral control
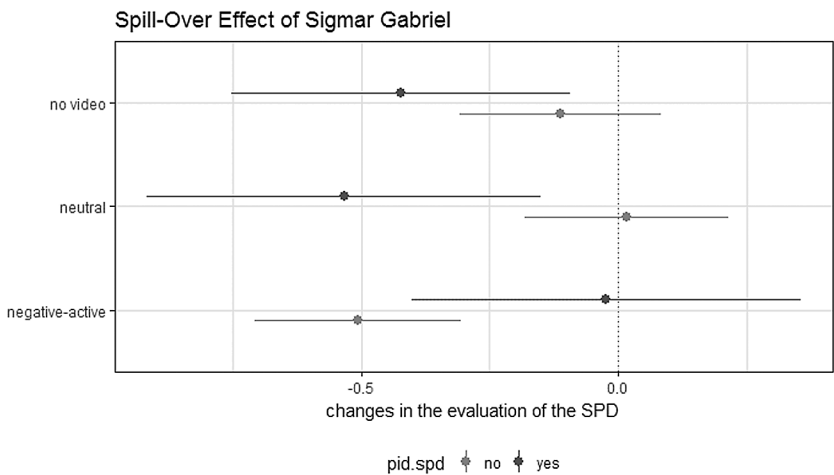condition, in which the SPD was not evaluated more negatively (Model 3).

By testing whether this spillover effect is moderated by party identifica-
tion, it can be seen that Sigmar Gabriel's anger had a negative impact on
the evaluation of the SPD only for participants who did not identify with
the Social Democratic Party (Model 4). Participants who supported the So-
cial Democratic Party remained unaffected by Sigmar Gabriel's expressions
in their support of the SPD (see Table 41 and Figure 24). This effect is sig-
nificant at the 1-percent level.

Table 41: Changes in the Evaluation of the Left and the SPD

	Dependent Variable: Scalometer			
	The Left		SPD	
	Gysi (1)	Gysi (2)	Gabriel (3)	Gabriel (4)
Negative-Active	0.457***	0.475***	-0.211*	-0.394***
	(0.144)	(0.151)	(0.125)	(0.143)
Positive	0.359**	0.338**		
	(0.145)	(0.152)		
Neutral	0.571***	0.636***	0.088	0.129
	(0.145)	(0.151)	(0.125)	(0.141)
PID	-0.481**	-0.325	-0.140	-0.311
	(0.192)	(0.334)	(0.122)	(0.196)
Female	0.140	0.138	-0.159	-0.185*
	(0.110)	(0.110)	(0.109)	(0.109)
Age	0.003	0.003	-0.003	-0.004
	(0.004)	(0.004)	(0.004)	(0.004)
Political Interest	-0.059	-0.059	-0.061	-0.062
	(0.062)	(0.062)	(0.061)	(0.060)
Negative-Active*PID		-0.183		0.794***
		(0.508)		(0.293)
Positive*PID		0.336		
		(0.524)		
Neutral*PID		-0.897*		-0.238
		(0.537)		(0.294)
Constant	0.188	0.173	-0.083	-0.025
	(0.115)	(0.117)	(0.106)	(0.113)
Observations	1,370	1,370	1,051	1,051
Adjusted R^2	0.016	0.017	0.005	0.015
F Statistic	4.201***	3.421***	1.880*	2.968***

Note: *p < 0.1; **p < 0.05; ***p < 0.01. The reference category is the experimental group without video treatment. The abbreviation "PID" stands for "party identification".

Figure 24: *Moderating Effects of Party Identification on the Evaluation of the SPD*

Spill-Over Effect of Sigmar Gabriel

Note: Figure displays coefficient plots with 95-percent confidence intervals.

The Longevity of the Spillover Effects

The longevity of these spillover effects can be tested by analyzing the difference between the pre-test values (Wave 1) and the second post-test (Wave 3). None of the treatment effects remained significant one week after the experiment was conducted according to the overall F-tests (see Table 42). This indicates weak and short-term spillover effects on the evaluation of political parties.

Table 42: Longevity of Spillover Effects on the Evaluation of Political Parties

	Dependent Variable: Party Scalometer			
	CDU		The Left	SPD
	E2	E3	E2	E3
	(1)	(2)	(3)	(4)
Negative-Active	0.236	-0.003	0.185	0.252
	(0.272)	(0.271)	(0.282)	(0.241)
Positive	0.413		0.013	
	(0.284)		(0.295)	
Neutral	0.052	0.179	0.034	0.158
	(0.279)	(0.271)	(0.290)	(0.242)
Constant	-0.156	-0.156	0.286	-0.517***
	(0.189)	(0.186)	(0.196)	(0.166)
Observations	527	410	527	410
Adjusted R^2	-0.001	-0.004	-0.005	-0.002
F Statistic	0.853	0.285	0.178	0.563

Note: *$p<0.1$; **$p<0.05$; ***$p<0.01$. The reference category is the experimental group without video treatment.

To conclude, the hypothesis H10 regarding the effect of politicians' emotional expressions on the evaluation of their respective parties has been partially supported. Overall, the effects were relatively weak and were not limited to anger expressions, but rather their appearances and expressions in general as the spillover effects for the Left were obtained after exposure to Gregor Gysi.

   A clear effect only occurred for Sigmar Gabriel's negative expressions and incivility, which resulted in less favorable evaluations of non-supporters. It can be noted that key politicians can potentially influence the imagine of their parties. Strong emotional expressions by party leaders, most likely in noticeable situations, have the potential to affect not only the evaluation of the politicians but also their parties.

5.3.3 The Response Time

The hypotheses H7 and H8 refer to participants' cognitive effort when political leaders are evaluated. The following section tests whether "high level" cognitive processes are activated by the experimental treatment, or whether automatic information processes are at work. By analyzing the response time of the candidate evaluation, some inferences about underlying mechanisms can be made in this regard.

   Cognitive theories of information processing often assume that two modes of processing can be distinguished in automatic processing and sys-

tematic processing. The use of heuristics falls into the former category. Hence, if a "high level" cognitive process is not necessary, the response time between experimental groups should not vary drastically – neither in comparison to the other experimental treatment groups nor to the group without video treatment.

Higher response times might be observed if the experimental treatment elicited emotional responses that lead to anxiety or irritation, which then triggered participants to think more carefully about the candidate evaluations. Participants who are enthusiastic about an opposing politician might feel irritated and therefore think more carefully about their responses.

Emotional expressions of anger could be used as heuristics that someone cares about something (Hochschild 2012), for example signaling that politicians care about political issues and voters. The use of heuristics does not require an enormous amount of cognitive effort and therefore, if no differences can be found between experimental and control groups, this could indicate that automatic processes might be at work.

In order to test the hypotheses H7 and H8, the response time is considered as the dependent variable for the analysis. The response time was measured only for sets of indicators throughout the questionnaire. When focusing on the evaluation of politicians, two sets of indicators measured response times and are suitable to gauge whether any changes in response time appeared: a response time variable that belongs to the candidates' evaluation of warmth and competence, and one that relates to the candidate evaluation according to semantic differentials. Both variables measured how long it took participants to answer the items of candidate orientation (warmth and competence) and semantic differentials for politicians in the second wave.

A quantile regression analysis is used to compare response time measures across experimental groups in order to examine the underlying processes. In a quantile regression any given quantile, for example the median as 0.5 quantile, can be estimated instead of the mean of the dependent variable (Koenker 2005). A quantile regression model is chosen due to its robustness against outliers. Compared to the mean, the median has the advantageous property that it is less strongly affected by outliers, such as those participants who were distracted throughout the online survey and

took longer to complete the questionnaire.[80] The following quantile regression models were conducted while controlling for age, gender, political interest and party identification. The quantile regressions were estimated based on a modified implementation of the Barrodale-Roberts (BR) algorithm (Koenker & d'Orey 1987) and standard errors were bootstrapped using the pwy-method as suggested by Parzen, Wei and Ying (1994), which has to be shown to perform well (Davino et al. 2013: 123–127).

The first table (Table 43) shows quantile regressions with the median response time of the candidate evaluation focusing on the items of candidate orientation as the dependent variables. Of particular interest in these models are the varying effects for Angela Merkel and Gregor Gysi, while no significant effects can be found for Sigmar Gabriel.[81] The constant in these models represents the median response time for those respondents who did not see a video, were male, of average age, with an average interest in politics, and without supporting the respective political party. At this baseline, without exposure to video clips, respondents took noticeably longer to assess Merkel's candidate qualities (Median = 19.5 seconds) than they did for Gregor Gysi (Median = 16.32 seconds), while both response times were longer than for politicians in general. Judgments about politicians in general were made fairly quickly, with a median of 23.62 seconds for groups with male politicians and 24.01 seconds for groups with female politicians. These were quick given that in these two instances the response time also included the response to twice as many items (or four additional indicators). Being exposed to video material is associated with faster response times for both of these models regarding the evaluation of politicians as a social group, except for the group of participants which saw negative-passive emotion of male politicians (Models 1 and 2).

If participants were exposed to videos clips, their median response time also decreased, when they evaluated Angela Merkel's warmth and competence – regardless of the experimental condition. Participants who saw Merkel's positive emotions were at the median almost two seconds faster in their judgment than those who did not see a video. This effect for An-

---

80 Similarly, the median is often used as a descriptive statistic in survey response time measures, and speeders are often identified by focusing on the median rather than the mean (Roßmann 2017: 100–101; Greszki et al. 2015).

81 The models for male and female politicians are not as relevant when focusing on candidate orientation due to an underlying measurement issue. The response time measure in both of these instances also includes four additional items regarding political efficacy and perceptions of politicians as a homogeneous social group.

gela Merkel can be replicated in the sub-experiment *Type 3*, in which participants also saw Sigmar Gabriel.

The exposure to video stimuli had the opposite effect for Gregor Gysi across experimental groups. Each video condition is associated with longer response times, whereby participants who saw his negative-active emotions were the fastest of those who saw a video clip with an increase in only 1.58 seconds, whereas those in the neutral control condition took 2.44 seconds longer at a median value. The response time for the evaluation of Sigmar Gabriel's qualities was not affected significantly by the experimental conditions (Model 5). Hence, these varying effects underline the different cognitive processes for each political leader as a result of the exposure to video clips. These differences could have been caused by the varying levels of familiarity with each political leader.

Table 43: Quantile Regressions on the Response Time of Candidate Orientation

	Dependent Variable: Response Time Candidate Orientation					
	Male Politicians (1)	Female Politicians (2)	Merkel (E2) (3)	Gysi (E2) (4)	Gabriel (E3) (5)	Merkel (E3) (6)
Negative-Active	-0.758	0.105	-1.250*	1.582***	0.189	-1.531**
	(0.787)	(0.840)	(0.651)	(0.609)	(0.591)	(0.720)
Negative-Passive	0.500	-1.105				
	(0.765)	(0.762)				
Positive	-0.774	0.368	-1.983***	1.902***		
	(0.697)	(0.808)	(0.614)	(0.634)		
Neutral	-0.113	0.579	-1.150*	2.443***	0.057	-1.469**
	(0.713)	(0.744)	(0.626)	(0.663)	(0.548)	(0.595)
Age	0.210***	0.211***	0.150***	0.164***	0.160***	0.156***
	(0.019)	(0.020)	(0.014)	(0.018)	(0.016)	(0.019)
Gender: Female	-1.968***	-1.158**	0.133	0.131	0.009	-0.406
	(0.561)	(0.541)	(0.490)	(0.522)	(0.429)	(0.544)
PID	2.210***	1.158	-2.433***	-1.500	-1.104**	-1.125**
	(0.695)	(0.727)	(0.458)	(0.970)	(0.484)	(0.522)
Political Interest	-1.016***	-0.842**	-0.217	0.057	-0.217	-0.437
	(0.144)	(0.168)	(0.261)	(0.318)	(0.237)	(0.296)
Constant	23.620***	24.008***	19.501***	16.319***	17.473***	19.626***
	(0.807)	(0.922)	(0.578)	(0.468)	(0.549)	(0.615)
Observations	1,764	1,766	1,372	1,372	1,051	1,051

Note: *p < 0.1; **p < 0.05; ***p < 0.01. The reference category in each model is the experimental group without video treatment. The quantile regressions are based on the median (0.5 quantile). The abbreviation "PID" stands for "party identification".

To test whether these findings are consistent, the next table (Table 44) presents the results of a median quantile regression that focused on the evaluation according to the semantic differentials as dependent variable.

This second set of dependent variables largely confirms the previous findings. Participants formed significantly quicker judgments if they saw a video of Angela Merkel, while it took them longer to respond to the items regarding Gregor Gysi and in this instance also Sigmar Gabriel.

The baseline for Angela Merkel was the highest, as in the previous models, whereby participants who saw no video, were male, of average age and with an average interest in politics, without supporting the CDU had a median response time of 26.63 seconds, when answering the six semantic differentials regarding Angela Merkel. In comparison, the estimate for almost the same group of participants – except for those with no party identification or any other than the Left – is a median of only 19.93 seconds for Gregor Gysi. As in the previous case, if participants however saw video clips of him, it took them longer to respond to the questionnaire items with a median increase of three to four seconds for each of the experimental groups, with anger displays falling in the middle with 3.57 seconds between positive emotions (3.30 seconds) and neutral displays (3.96 seconds).

Regarding the semantic differentials, participants who saw videos of Sigmar Gabriel also took significantly longer to respond to the questionnaire items, once they were exposed to either video condition. This is in line with the treatment effects for Sigmar Gabriel, which could mainly be found for the semantic differentials, but not for the items that measured his candidate orientation.

After accounting for the exposure effects of the video treatments, the response time converges for all three politicians at a similar level. Although it took respondents longer to evaluate politicians as social groups, it has to be kept in mind that the semantic differentials for politicians in general included two additional items. These models however only tested whether the response time differed across groups; as yet, none of the models indicated that anger expressions had led to a systematic increase or decrease in response times.

228

Table 44: Quantile Regressions on the Response Time of Semantic Differentials

	Male Politicians (1)	Female Politicians (2)	Merkel (E2) (3)	Gysi (E2) (4)	Gabriel (E3) (5)	Merkel (E3) (6)
	Dependent Variable: Response Time Semantic Differentials					
Negative-Active	-0.950	-0.264	-1.960**	3.565***	2.027***	-2.226**
	(1.571)	(1.473)	(0.844)	(0.821)	(0.707)	(0.943)
Negative-Passive	2.550*	-1.971				
	(1.515)	(1.497)				
Positive	-1.750	-0.829	-2.960***	3.304***		
	(1.291)	(1.343)	(0.884)	(0.659)		
Neutral	0.650	-0.486	-2.040**	3.957***	2.730***	-2.358***
	(1.627)	(1.440)	(0.941)	(0.882)	(0.741)	(0.889)
Age	0.450***	0.464***	0.240***	0.239***	0.216***	0.264***
	(0.037)	(0.032)	(0.020)	(0.021)	(0.022)	(0.029)
Gender: Female	0.400	-0.293	0.080	1.022*	-0.486	1.113
	(1.084)	(0.922)	(0.669)	(0.617)	(0.654)	(0.836)
PID	0.350	-1.350	0.320	2.261**	-0.000	1.340
	(1.328)	(1.273)	(0.647)	(1.100)	(0.765)	(0.876)
Political Interest	-0.700	-0.271**	-0.440	0.565	0.081	-0.377
	(0.617)	(0.571)	(0.391)	(0.391)	(0.342)	(0.379)
Constant	44.100***	46.580***	26.630***	19.925***	21.282***	26.070***
	(1.495)	(1.530)	(0.771)	(0.619)	(0.582)	(0.861)
Observations	1,764	1,766	1,372	1,372	1,051	1,051

Note: *p < 0.1; **p < 0.05; ***p < 0.01. The reference category in each model is the experimental group without video treatment. The quantile regressions are based on the median (0.5 quantile). The abbreviation "PID" stands for "party identification".

However, as stated in hypothesis H8, it is possible that participants who feel conflicted about the video treatment could take longer to derive a judgment about the politician in question. To test this assumption, the next table (Table 45) presents two additional models for Gregor Gysi. The case of Gregor Gysi seems especially suited to test this hypothesis H8, due to the fact that most participants did not have a high opinion of him before the experimental treatment was administered – at least given that they did not identify with the Left, according to scalometer ratings in the first panel wave (M = 4.42, SD = 2.85). In contrast, participants who identified with the Left had on average a more favorable evaluation of him (M = 8.20, SD = 1.80). Across political parties, participants had the strongest reactions towards Gysi's emotional expressions compared to the other politicians, i.e., the changes in his evaluation were the most noticeable according to the average treatment effects (see Subchapter 5.1).

Since participants who support the Left had already a high opinion of Gysi in wave 1, the following analysis solely focuses on the participants who did not support the Left and therefore, 130 participants were exclud-

ed from the following analysis. This is in line with the theoretical assumptions, since supporters of the Left should not feel irritated by any positive reactions towards Gregor Gysi's emotional expressions.

Hence, the two models in the following table (Table 45) included an additional dichotomous variable in the models that indicated whether a positive change in the evaluation of Gysi had occurred. A positive change was determined as being larger than the average change for those participants that did not support the Left.

Since the response time is only measured for two sets of indicators regarding the evaluation of political leaders, the dichotomous variable indicating change is also measured across all four items of candidate orientation (warmth and competence) and for three out of six items that measured candidate evaluations according to semantic differentials and indicated positive evaluation, such as politeness, peacefulness and modesty.[82] An interaction between such a positive reaction (change) and the experimental treatment was then added into the models.

The interaction effect is also displayed in Figure 25. From this analysis it can be seen that participants who were exposed to Gysi's negative active emotional expressions and changed their opinion in favor of Gregor Gysi did not take longer to evaluate his warmth and competence than those participants who did not respond positively. The same finding held true for the evaluation of his character traits according to semantic differentials (see Figure 26). Hence, the hypothesis H8 is not supported by the experimental findings. Exposure to video clips of Gregor Gysi resulted in longer response times regardless of the emotional expressions that were presented. Anger expressions also led to slightly quicker judgements compared to other emotional expressions; however, these differences are not statistically significant (see Figure 25 and Figure 26).
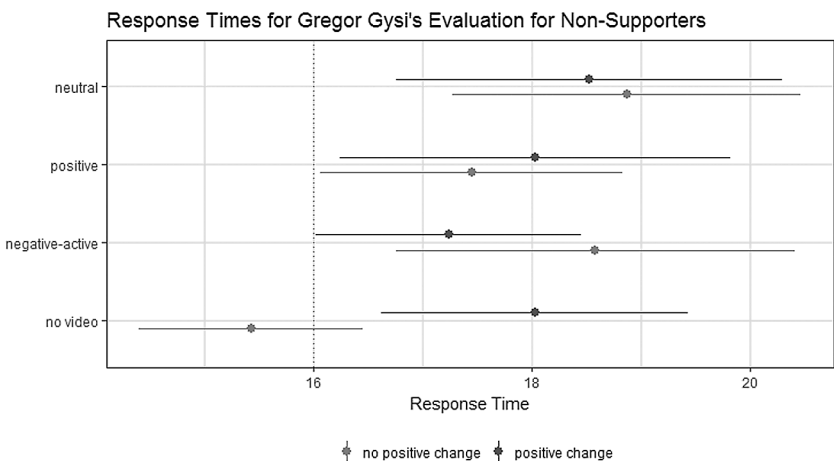
---

82 Positive changes in the evaluation of arousal (agitation), emotionality/rationality, and attractiveness were not considered as relevant positive changes. The level of arousal could be interpreted as a manipulation check for the experimental treatment.

Table 45: Interaction Between the Experimental Treatment and Positive Evaluation of Gregor Gysi on the Response Time for Non-Supporters

	Dependent Variable:	Response Time
	Candidate Orientation (1)	Semantic Differentials (2)
Negative-Active	3.158***	4.114***
	(0.993)	(1.014)
Positive	2.026**	3.250***
	(0.805)	(1.031)
Neutral	3.447***	5.977***
	(0.901)	(1.215)
Positive Reaction	2.605***	0.682
	(0.871)	(1.093)
Age	0.158***	0.227***
	(0.019)	(0.022)
Gender: Female	0.553	1.045
	(0.591)	(0.664)
Political Interest	0.184	0.432
	(0.330)	(0.388)
Negative-Active*Positive Reaction	-3.947***	-1.341
	(1.393)	(1.707)
Positive*Positive Reaction	-2.026	0.205
	(1.427)	(1.525)
Neutral*Positive Reaction	-2.947**	-2.773
	(1.440)	(1.743)
Constant	15.424***	19.488***
	(0.556)	(0.702)
Observations	1,260	1,257

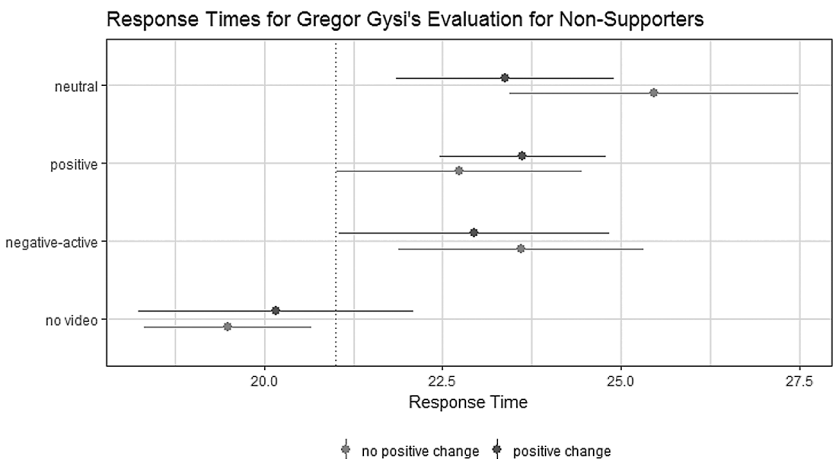*Note*: *p < 0.1; **p < 0.05; ***p < 0.01. The reference category is the experimental group without video treatment.

Figure 25: *Interaction Effect Between the Experimental Treatment and Positive Changes Towards Gysi for Non-Supporters (Candidate Orientation)*

Note: The figure displays median quantile regression coefficients and confidence intervals based on Table 45 (Model 1). The dotted line indicates median response time for the control group without video treatment.

Figure 26: *Interaction Effect Between the Experimental Treatment and Positive Changes Towards Gysi for Non-Supporters (Semantic Differentials)*

Note: The figure displays median quantile regression coefficients and confidence intervals based on Table 45 (Model 1). The dotted line indicates median response time for the control group without video treatment.

5.3.4 Summary of the Broader Implications

This section provides an overview of the broader implications of this study. The previous analysis focused on the longevity of the experimental effects, possible spillover effects as well as response time measurements as potential indicators of underlying mechanisms. The longevity of the experimental effects has been assessed by analyzing data from a third panel wave. At least seven days have passed between the third and second panel waves, in which the experimental treatment had been initially administered. By and large, most effects were not long-lasting, since no difference between the first wave and the third wave could be detected. This is not a surprise, as most effects were very small – even immediately after the experiment was administered. However, one exception occurred for Gregor Gysi: the treatment effects on his evaluations were the largest compared to the other politicians (Subchapter 5.1). They were also the ones that lasted the longest, although some decay occurred in comparison to the initial effects. For his overall evaluation and the evaluation of his likeability and trustworthiness, small treatment effects could be found when comparing the differences in evaluations between the first and third panel wave. Thus, the hypothesis H9 about short-term effects is partially rejected, as the analysis of Gysi's anger expressions indicates that effects can still be present one week after the initial administration of the experimental treatment.

As a second criterion to test broader implications, spillover effects were analyzed that might occur based on the evaluation of key political leaders on their respective parties. However, such spillover effects were not clearly linked to the expression of anger. A spillover effect could be found for Angela Merkel and the CDU only due to her positive emotional expressions. For the Left, a positive spillover occurred for non-supporters due to the exposure to the video clips – regardless of Gysi's emotional expressions. Only for Sigmar Gabriel was a negative spillover more closely linked to his expressions of incivility and moderated by the individual party identification. Respondents who supported the SPD remained unaffected by Gabriel's incivility in their evaluation of the Social Democratic Party. In contrast to those who supported the party, other participants (with a different or no party identification) reacted negatively to Gabriel's anger expressions and rated the SPD less favorably based on the scalometer evaluation. Overall, the analysis did not clearly support hypothesis H10, that exposure to emotional expressions of party leaders could also affect the evaluation of their respective parties. The initial spillover effects were not long-lasting, as they had disappeared one week after the initial exposure. However, contin-

uous uncivil media appearances of party leaders as well as strong emotional expressions could affect the evaluation of political parties more gradually.

Lastly, the response time was analyzed in order to test the reaction time as an indicator of the underlying causal mechanisms that could shape the candidate evaluation. The analysis provided some insights into differences between the politicians, whereby participants took longer answering items after being exposed to Gregor Gysi – a candidate that might be lesser known. After seeing Angela Merkel – the most well-known German politician – participants were faster in their responses.

This indicates that response times and cognitive processing might depend on how much previous knowledge is available to the participants. For this study, the response times were not affected by specific emotional expressions, but rather by exposure itself. Hence the hypothesis H7 about longer response times for the experimental group of anger expression is not supported by the analysis. In addition, the hypothesis H8 – whereby it was assumed that participants would take longer after reacting positively to emotional displays of a politician with low initial ratings – could also not be supported by the experimental data. This hypothesis was only tested by using data in relation to Gregor Gysi and those who did not support the Left.