

Supplementary Materials

for

“Prejudice in Disguise:

Which Features Determine the Subtlety of Ethnically Prejudicial Statements?”

Karolina Fetz^a

Martin Kroh^{a, b, c}

^aBerlin Institute for Integration and Migration Research (BIM), Humboldt-Universität zu Berlin, Berlin, Germany

^bFaculty of Sociology, Bielefeld University, Bielefeld, Germany

^cSocio-Economic-Panel (SOEP), German Institute for Economic Research (DIW) Berlin, Berlin, Germany

These Supplementary Materials contain an overview of all constructs included in the larger telephone survey (Table A1), an English translation of the experimental materials used in the study (Table A2), descriptive statistics (Table A3), results from further exploratory analyses (Tables A4–A8) as well results from an additional online survey (pp. 11-17; Tables A9–12).

Table A1

Overview of the Measures Included in the Full Telephone Survey Questionnaire

<i>Construct</i>	<i>Specification of measures</i>
<i>Prejudice endorsement</i>	2 items each of the Subtle and Blatant Prejudice Scale (Pettigrew & Meertens, 1995; German translation following Zick, 1997); Random selection of 4 items per participant (set 1) from a pool of 24 self-developed items (factorial survey experiment; see Table A2 of the Supplementary Materials); 6 self-developed prejudice items (set 2); Indication of agreement on a 1 (<i>completely disagree</i>) to 5 (<i>completely agree</i>) scale.
<i>Perceived prejudice endorsement of the general population and important referents</i>	3 self-developed prejudice items (set 3); Indication of estimated agreement of the general population in Germany and of important referents on a 1 (<i>completely disagree</i>) to 5 (<i>completely agree</i>) scale.
<i>Socio-demographic characteristics</i>	Age, country of birth, German citizenship, parents' countries of birth, religious affiliation, education, employment status, political interest, party preference; postal code and household income (the two last constructs were asked at the end of the interview)
<i>Xenophobia ratings</i>	2 items each from the Subtle and Blatant Prejudice Scale (Pettigrew & Meertens, 1995); 4 prejudice items (factorial survey experiment; set 1), 6 prejudice items (set 2); Two different instructions, one containing a definition of xenophobia, the other one not containing a definition of xenophobia; Rating on a 1 (<i>not xenophobic at all</i>) to 5 (<i>very xenophobic</i>) scale.
<i>Affect</i>	11 items; Indication of agreement with different emotional states (i.a. anger, joy, sadness, discomfort, negative self-directed emotions) on a 1 (<i>completely disagree</i>) to 5 (<i>completely agree</i>) scale.
<i>Affective and behavioral reaction to overhearing prejudicial statements</i>	3 self-developed prejudice items (set 3); Indication of whether overhearing such a statement would cause feelings of anger or encouragement or would lead to openly contradict the speaker on a 1 (<i>completely disagree</i>) to 5 (<i>completely agree</i>) scale.
<i>Egalitarian self-concept</i>	1 item, Indication of agreement on a 1 (<i>completely disagree</i>) to 5 (<i>completely agree</i>) scale.
<i>Humanitarianism-egalitarianism</i>	6 items of the humanitarianism-egalitarianism scale (Doll & Dick, 2000; Katz & Hass, 1988); Indication of agreement on a 1 (<i>completely disagree</i>) to 5 (<i>completely agree</i>) scale.
<i>Protestant work ethic</i>	4 items (Levin, Sidanius, Rabinowitz, & Federico, 1998); Indication of agreement on a 1 (<i>completely disagree</i>) to 5 (<i>completely agree</i>) scale.
<i>Social dominance orientation</i>	Full 8-item SDO _{7(s)} -Scale (Ho et al., 2015); Indication of agreement on a 1 (<i>completely disagree</i>) to 5 (<i>completely agree</i>) scale.
<i>Internal and external motivation to respond without prejudice scale</i>	2 items each of the Internal and External Motivation to Respond Without Prejudice Scale (Plant & Devine, 1998); Indication of agreement on a 1 (<i>completely disagree</i>) to 5 (<i>completely agree</i>) scale.
<i>Preference for consistency</i>	6 items of the Preference for Consistency Scale (Cialdini, Trost, & Newsom, 1995; German translation following Klocke, 2010); Indication of agreement on a 1 (<i>completely disagree</i>) to 5 (<i>completely agree</i>) scale.
<i>Social desirability</i>	Full 6-item KSE-G Scale (Kemper, Beierlein, Bensch, Kovaleva, & Rammstedt, 2012); Indication of agreement on a 1 (<i>completely disagree</i>) to 5 (<i>completely agree</i>) scale.

Supplementary Materials for “Prejudice in Disguise: Which Features Determine the Subtlety of Ethnically Prejudicial Statements?”

<i>Contact with migrants/ non-migrants and contact valence</i>	4 items, Frequency of contact with (non-)migrants among family and relatives, at work, in the neighborhood, and among friends (ALLBUS); Indication of frequency on a 1 (<i>never</i>) to 5 (<i>very often</i>) scale; 1 item, Indication of contact valence on a 1 (<i>very negative</i>) to 5 (<i>very positive</i>) scale
<i>Worries</i>	Worries about the general and own economic situation, peace, criminality, xenophobia, immigration (Socio-Economic Panel, SOEP), and the conservation of language and culture, and values in Germany, Response on a 1 (<i>no worries</i>) to 3 (<i>big worries</i>) scale.

Note. Constructs are presented in their order of appearance in the telephone survey. All items of English scales were translated into German.

Table A2

English Translation of Systematically Varied Prejudice Items

		Topic		
		<i>Culture</i>	<i>Economic Utility</i>	<i>Danger/ Inner Security</i>
Target Group <i>Muslims</i>	Language			
	<i>Weakly essentialist</i>	Muslims need particular assistance to adapt to the idea that men and women have equal rights in Germany.	Muslims need particular assistance to adapt to the strong work ethic in Germany.	Muslims need particular assistance to adapt to the norm that conflicts are solved without violence in Germany.
	<i>Rather weakly essentialist</i>	Muslims cannot adapt to the idea that men and women have equal rights in Germany.	Muslims cannot adapt to the strong work ethic in Germany.	Muslims cannot adapt to the norm that conflicts are solved without violence in Germany.
	<i>Rather strongly essentialist</i>	Muslims are more sexist than Germans.	Muslims are more workshy than Germans.	Muslims are more inclined to violence than Germans.
	<i>Strongly essentialist</i>	Muslims are, by nature, more sexist than Germans.	Muslims are, by nature, more workshy than Germans.	Muslims are, by nature, more inclined to violence than Germans.
Target Group <i>Turks</i>	<i>Weakly essentialist</i>	Turks need particular assistance to adapt to the idea that men and women have equal rights in Germany.	Turks need particular assistance to adapt to the strong work ethic in Germany.	Turks need particular assistance to adapt to the norm that conflicts are solved without violence in Germany.
	<i>Rather weakly essentialist</i>	Turks cannot adapt to the idea that men and women have equal rights in Germany.	Turks cannot adapt to the strong work ethic in Germany.	Turks cannot adapt to the norm that conflicts are solved without violence in Germany.
	<i>Rather strongly essentialist</i>	Turks are more sexist than Germans.	Turks are more workshy than Germans.	Turks are more inclined to violence than Germans.
	<i>Strongly essentialist</i>	Turks are, by nature, more sexist than Germans.	Turks are, by nature, more workshy than Germans.	Turks are, by nature, more inclined to violence than Germans.

Note. Original German items can be obtained from the authors upon request.

Table A3

Descriptive Statistics: Xenophobia Ratings for the Systematically Varied Prejudice Items

	Topic								
	Culture			Economic utility			Danger		
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>
Target group: Muslims									
Language									
Weakly essentialist	148	3.05	1.24	157	3.17	1.36	155	3.11	1.32
Rather weakly essentialist	145	3.47	1.17	156	3.97	1.13	143	3.59	1.16
Rather strongly essentialist	151	3.63	1.29	142	4.04	1.24	170	3.94	1.29
Strongly essentialist	135	3.69	1.35	149	3.94	1.36	125	4.19	1.07
Target group: Turks									
Language									
Weakly essentialist	151	3.30	1.22	158	3.63	1.32	152	3.37	1.24
Rather weakly essentialist	162	3.47	1.21	136	3.87	1.17	154	3.77	1.14
Rather strongly essentialist	126	3.68	1.29	156	3.96	1.29	145	3.86	1.25
Strongly essentialist	144	3.50	1.30	150	4.25	1.06	145	3.97	1.24

Note. *N* = Observations per item; $N_{\text{Participants}} = 895$. Each respondent rated four randomly selected items. Items were rated on a 5-point scale from 1 (*not xenophobic at all*) to 5 (*very xenophobic*).

Table A4

Model of Fixed Effects of Item Dimensions on Xenophobia Ratings with Cross-Level Interactions Between Item Dimensions and Respondents' Social Dominance Orientation (SDO)

	<i>b (SE)</i>	<i>95% Confidence Interval</i>	
Main effects: Item dimensions			
Topic ^a			
Economic utility	0.427*** (0.037)	0.354	0.499
Danger	0.246*** (0.037)	0.173	0.319
Language ^b			
Rather weakly essentialist	0.332*** (0.042)	0.250	0.415
Rather strongly essentialist	0.561*** (0.042)	0.478	0.643
Strongly essentialist	0.613*** (0.043)	0.529	0.696
Target group ^c			
Turks	0.068* (0.030)	0.010	0.127
Cross-level interactions: Item dimensions x SDO			
Topic ^a x SDO			
Economic utility	-0.227*** (0.058)	-0.342	-0.113
Danger	-0.184** (0.058)	-0.298	-0.070
Language ^b x SDO			
Rather weakly essentialist	-0.085 (0.068)	-0.218	0.048
Rather strongly essentialist	-0.091 (0.066)	-0.220	0.038
Strongly essentialist	-0.072 (0.068)	-0.205	0.061
Target group ^c x SDO			
Turks	-0.008 (0.048)	-0.102	0.085
Intercept	3.048*** (0.039)	2.972	3.124
<i>R</i> ² _{within}	.140		
<i>R</i> ² _{between}	.116		
<i>R</i> ² _{overall}	.113		

Note. $N_{\text{Item ratings}} = 3,555$; $N_{\text{Participants}} = 895$. Unstandardized coefficients. Standard error in parentheses. The coefficient relates to a 5-point scale from 1 (*not xenophobic at all*) to 5 (*very xenophobic*). The variable SDO was mean centered. The main effect of SDO (level 2 variable) as a trait of individuals is invariant across manipulations and thus omitted by the FE model.

^a Reference category: Culture. ^b Reference category: Weakly essentialist. ^c Reference category: Muslims
 * $p < .05$; ** $p < .01$; *** $p < .001$.

Table A5

Model of Fixed Effects of Item Dimensions on Xenophobia Ratings with Cross-Level Interactions Between Item Dimensions and Respondents' Humanitarianism-Egalitarianism (HE)

	<i>b (SE)</i>	<i>95% Confidence Interval</i>	
Main effects: Item dimensions			
Topic ^a			
Economic utility	0.431*** (0.037)	0.358	0.503
Danger	0.245*** (0.037)	0.172	0.318
Language ^b			
Rather weakly essentialist	0.333*** (0.042)	0.250	0.416
Rather strongly essentialist	0.564*** (0.042)	0.481	0.646
Strongly essentialist	0.616*** (0.043)	0.533	0.700
Target group ^c			
Turks	0.070* (0.030)	0.011	0.129
Cross-level interactions: Item dimensions x HE			
Topic ^a x HE			
Economic utility	0.011 (0.063)	-0.113	0.135
Danger	-0.110 (0.064)	-0.235	0.016
Language ^b x HE			
Rather weakly essentialist	0.006 (0.073)	-0.138	0.150
Rather strongly essentialist	-0.005 (0.073)	-0.148	0.138
Strongly essentialist	0.002 (0.075)	-0.145	0.149
Target group ^c x HE			
Turks	0.002 (0.051)	-0.097	0.101
Intercept	3.043*** (0.039)	2.967	3.119
<i>R</i> ² _{within}	.136		
<i>R</i> ² _{between}	.007		
<i>R</i> ² _{overall}	.048		

Note. $N_{Item\ ratings} = 3,555$; $N_{Participants} = 895$. Unstandardized coefficients. Standard error in parentheses. The coefficient relates to a 5-point scale from 1 (*not xenophobic at all*) to 5 (*very xenophobic*). The variable HE was mean centered. The main effect of HE (level 2 variable) as a trait of individuals is invariant across manipulations and thus omitted by the FE model.

^a Reference category: Culture. ^b Reference category: Weakly essentialist. ^c Reference category: Muslims

* $p < .05$; ** $p < .01$; *** $p < .001$.

Table A6

Model of Fixed Effects of Item Dimensions on Xenophobia Ratings with Cross-Level Interactions Between Item Dimensions and Respondents' Egalitarian Self-Concept (ES)

	<i>b (SE)</i>	<i>95% Confidence Interval</i>	
Main effects: Item dimensions			
Topic ^a			
Economic utility	0.430*** (0.037)	0.357	0.502
Danger	0.248*** (0.037)	0.175	0.321
Language ^b			
Rather weakly essentialist	0.330*** (0.042)	0.247	0.412
Rather strongly essentialist	0.566*** (0.042)	0.484	0.649
Strongly essentialist	0.616*** (0.043)	0.532	0.699
Target group ^c			
Turks	0.071* (0.030)	0.013	0.130
Cross-level interactions: Item dimensions x ES			
Topic ^a x ES			
Economic utility	0.066 (0.047)	-0.025	0.158
Danger	-0.001 (0.047)	-0.093	0.091
Language ^b x ES			
Rather weakly essentialist	0.086 (0.053)	-0.019	0.190
Rather strongly essentialist	0.117* (0.052)	0.016	0.219
Strongly essentialist	0.072 (0.055)	-0.035	0.179
Target group ^c x ES			
Turks	0.029 (0.037)	-0.043	0.102
Intercept	3.043*** (0.039)	2.967	3.120
<i>R</i> ² _{within}	.137		
<i>R</i> ² _{between}	.071		
<i>R</i> ² _{overall}	.086		

Note. $N_{\text{Item ratings}} = 3,551$; $N_{\text{Participants}} = 894$. Unstandardized coefficients. Standard error in parentheses. The coefficient relates to a 5-point scale from 1 (*not xenophobic at all*) to 5 (*very xenophobic*). The variable ES was mean centered. The main effect of ES (level 2 variable) as a trait of individuals is invariant across manipulations and thus omitted by the FE model.

^a Reference category: Culture. ^b Reference category: Weakly essentialist. ^c Reference category: Muslims

* $p < .05$; ** $p < .01$; *** $p < .001$.

Table A7

Model of Fixed Effects of Item Dimensions on Xenophobia Ratings with Cross-Level Interactions Between Item Dimensions and Respondents' Protestant Work Ethic (PWE)

	<i>b (SE)</i>	<i>95% Confidence Interval</i>	
Main effects: Item dimensions			
Topic ^a			
Economic utility	0.431*** (0.037)	0.358	0.503
Danger	0.245*** (0.037)	0.172	0.318
Language ^b			
Rather weakly essentialist	0.336*** (0.042)	0.253	0.419
Rather strongly essentialist	0.563*** (0.042)	0.481	0.645
Strongly essentialist	0.617*** (0.043)	0.534	0.701
Target group ^c			
Turks	0.072* (0.030)	0.013	0.131
Cross-level interactions: Item dimensions x PWE			
Topic ^a x PWE			
Economic utility	0.003 (0.049)	-0.094	0.099
Danger	0.084 (0.050)	-0.013	0.182
Language ^b x PWE			
Rather weakly essentialist	-0.006 (0.056)	-0.116	0.103
Rather strongly essentialist	-0.041 (0.056)	-0.151	0.069
Strongly essentialist	0.056 (0.057)	-0.056	0.169
Target group ^c x PWE			
Turks	-0.057 (0.039)	-0.134	0.021
Intercept	3.041*** (0.039)	2.965	3.117
<i>R</i> ² _{within}	.137		
<i>R</i> ² _{between}	.013		
<i>R</i> ² _{overall}	.054		

Note. $N_{Item\ ratings} = 3,555$; $N_{Participants} = 895$. Unstandardized coefficients. Standard error in parentheses. The coefficient relates to a 5-point scale from 1 (*not xenophobic at all*) to 5 (*very xenophobic*). The variable PWE was mean centered. The main effect of PWE (level 2 variable) as a trait of individuals is invariant across manipulations and thus omitted by the FE model.

^a Reference category: Culture. ^b Reference category: Weakly essentialist. ^c Reference category: Muslims
* $p < .05$; ** $p < .01$; *** $p < .001$.

Table A8

Model of Fixed Effects of Item Dimensions on Xenophobia Ratings with Cross-Level Interactions Between Item Dimensions and Respondents' Social Desirability (SD)

	<i>b (SE)</i>	<i>95% Confidence Interval</i>	
Main effects: Item dimensions			
Topic ^a			
Economic utility	0.426*** (0.037)	0.354	0.498
Danger	0.251*** (0.037)	0.178	0.324
Language ^b			
Rather weakly essentialist	0.327*** (0.042)	0.244	0.410
Rather strongly essentialist	0.560*** (0.042)	0.478	0.643
Strongly essentialist	0.611*** (0.043)	0.528	0.695
Target group ^c			
Turks	0.070* (0.030)	0.011	0.129
Cross-level interactions: Item dimensions x SD			
Topic ^a x SD			
Economic utility	-0.150** (0.054)	-0.257	-0.044
Danger	-0.088 (0.055)	-0.196	0.020
Language ^b x SD			
Rather weakly essentialist	-0.037 (0.063)	-0.160	0.087
Rather strongly essentialist	-0.023 (0.063)	-0.147	0.101
Strongly essentialist	-0.027 (0.065)	-0.154	0.101
Target group ^c x SD			
Turks	-0.067 (0.045)	-0.154	0.021
Intercept	3.046*** (0.039)	2.970	3.122
<i>R</i> ² _{within}	.136		
<i>R</i> ² _{between}	.032		
<i>R</i> ² _{overall}	.066		

Note. $N_{Item\ ratings} = 3,539$; $N_{Participants} = 891$. Unstandardized coefficients. Standard error in parentheses. The coefficient relates to a 5-point scale from 1 (*not xenophobic at all*) to 5 (*very xenophobic*). The variable SD was mean centered. The main effect of SD (level 2 variable) as a trait of individuals is invariant across manipulations and thus omitted by the FE model.

^a Reference category: Culture. ^b Reference category: Weakly essentialist. ^c Reference category: Muslims

* $p < .05$; ** $p < .01$; *** $p < .001$.

Results from an Additional Online Survey

We conducted an additional online survey with a convenience sample in order to further validate the systematic variations of the topic and language dimension of the constructed prejudice items ($N_{\text{Item ratings}} = 886$ from $N_{\text{Participants}} = 238$, 124 women, 77 men, 37 diverse/ no indication, $M_{\text{age}} = 35.69$ years, $SD = 13.87$, age range 19–82 years). Participants were asked to evaluate a randomized set of four (out of 24) prejudice items with respect to their perception of these dimensions. Regarding the topic dimension, results from three one-way ANOVAs (across observations) with planned Bonferroni-corrected contrasts (see Table A9) confirm that items within the culture topic ($F(2, 883) = 276.10, p < .001$) were indeed perceived as relating to an attribute of the respective target group that threatens the culture (i.e., gender equality) in Germany ($ps < .001$), items within the economic utility topic ($F(2, 883) = 207.56, p < .001$) were more strongly perceived as relating to threats to the productivity in Germany ($ps < .001$), and items of the danger topic ($F(2, 883) = 244.58, p < .001$) were more strongly judged as relating to threats to security ($ps < .001$), compared to items of the two other topics, respectively. Additionally, we examined whether our systematic variation of the linguistic phrasing actually manipulated the degree of essentialism. To this end, participants rated the prejudice items on several sub-dimensions of essentialism (adapting seven items from Haslam, Rothschild, & Ernst, 2000), which were collapsed into an average measure of essentialism for each prejudice item rated per participant. Correlational results (across observations; see Table A10) indicate that along the range of the manipulated linguistic phrasing, participants indeed perceived the items to be more essentialist ($r = .129, p < .001$), whereby a closer look at the sub-dimensions revealed an increase especially for the perceived immutability (i.e., whether membership in the respective target group is regarded as fixed; $r = .150, p < .001$) and stability (i.e., whether the group and its characteristics is perceived as stable over time; $r = .168, p < .001$). In addition, we also asked participants whether they perceived the target group’s described attribute to be unchangeable (in general, by themselves, or by others), as an additional, more straightforward assessment of perceived essentialism. Indeed, along the range of the linguistic phrasing the attribute of the respective target group was increasingly perceived as unchangeable (or uncontrollable) in general ($r = .252, p < .001$) and more specifically as increasingly unchangeable by others ($r = .287, p < .001$), rather than by the respective target group itself ($r = .057,$

$p = .092$). More detailed results from a one-way ANOVA (across observations; see Table A11) with Bonferroni-corrected pairwise comparisons regarding the perceived unchangeability of the attribute in general ($F(3, 882) = 39.03, p < .001$) and by others ($F(3, 882) = 43.88, p < .001$) show that mainly items of the weakly essentialist phrasing differed from items of the three stronger levels of the essentialist linguistic phrasing ($ps < .001$). The locus for that perceived uncontrollability, varying between the first and other three levels of the language dimension, thus primarily lay in others, rather than the target group itself. Overall, these additional survey results validate that the systematic item variations were indeed effective in manipulating the intended levels of the topic and language, i.e., essentialist phrasing, dimension.

Table A9

Planned Contrasts after One-Way ANOVAs for the Perception of the Item Dimension Topic

	<i>Contrast (SE)</i>	<i>95% Confidence Interval</i>	
Items refer to threats to culture			
Culture (ref.)			
Economic utility	-2.341*** (0.102)	-2.569	-2.113
Danger	-1.644*** (0.102)	-1.874	-1.415
Items refer to threat to productivity			
Economic utility (ref.)			
Culture	-1.845*** (0.104)	-2.079	-1.611
Danger	-1.755*** (0.101)	-1.982	-1.528
Items refer to threat to security			
Danger (ref.)			
Culture	-1.344*** (0.103)	-1.575	-1.113
Economic utility	-2.177*** (0.099)	-2.400	-1.954

Note. One-way ANOVAs were conducted across observations, based on $N_{\text{Item ratings}} = 886$ from $N_{\text{Participants}} = 238$. Standard error in parentheses. Reference categories are indicated by (ref.). Significance levels are Bonferroni-corrected for two comparisons, respectively.

* $p < .05$; ** $p < .01$; *** $p < .001$.

Table A10

Pairwise Correlations between the Item Dimension Language with Perceived Essentialism and Its Subdimensions

	<i>Language (Essentialist Phrasing)</i>
Subdimensions of Essentialism	
Discreteness	.048
Naturalness	.076*
Immutability	.150***
Stability	.168***
Uniformity	.092**
Informativeness	.066*
Inherence	.074*
Essentialism (Sum-Score)	.129***

Note. Correlations were calculated across observations, based on $N_{\text{Item ratings}} = 886$ from $N_{\text{Participants}} = 238$. The item dimension language (essentialist phrasing) was entered as a continuous variable.

* $p < .05$; ** $p < .01$; *** $p < .001$.

Table A11

Pairwise Comparisons after One-Way ANOVAs for the Perceived Uncontrollability (General and by Others) for Different Levels of the Item Dimension Language

	<i>Contrast (SE)</i>	<i>95% Confidence Interval</i>	
Perceived general uncontrollability			
Rather weakly essentialist vs. weakly essentialist	1.103*** (0.117)	0.795	1.412
Rather strongly essentialist vs. weakly essentialist	0.915*** (0.117)	0.606	1.224
Strongly essentialist vs. weakly essentialist	1.050*** (0.117)	0.740	1.360
Rather strongly essentialist vs. rather weakly essentialist	-0.188 (0.116)	-0.495	0.119
Strongly essentialist vs. rather weakly essentialist	-0.053 (0.116)	-0.361	0.254
Strongly essentialist vs. rather strongly essentialist	0.135 (0.116)	-0.173	0.443
Perceived uncontrollability by others			
Rather weakly essentialist vs. weakly essentialist	1.134*** (0.122)	0.813	1.456
Rather strongly essentialist vs. weakly essentialist	1.102*** (0.122)	0.780	1.424
Strongly essentialist vs. weakly essentialist	1.190*** (0.122)	0.867	1.513
Rather strongly essentialist vs. rather weakly essentialist	-0.032 (0.121)	-0.352	0.288
Strongly essentialist vs. rather weakly essentialist	0.056 (0.121)	-0.265	0.377
Strongly essentialist vs. rather strongly essentialist	0.088 (0.121)	-0.233	0.409

Note. One-way ANOVAs were conducted across observations, based on $N_{\text{Item ratings}} = 886$ from $N_{\text{Participants}} = 238$. Standard error in parentheses. Significance levels are Bonferroni-corrected for six comparisons, respectively.

* $p < .05$; ** $p < .01$; *** $p < .001$.

In the additional online survey, we also tested the prejudice items for their perceived negativity and assessed whether this perceived negative valence of the prejudice items varied for the manipulated dimensions. Results from a one-way ANOVA (across observations) with Bonferroni-corrected pairwise comparisons and an independent t-test (across observations), respectively, show that the perceived negative valence of the prejudice items neither differed by the topic ($F(2, 883) = 0.56, p = .571$), nor the target group ($t(884) = -0.631, p = .528$) they referred to. However, correlational results (across observations) indicate that the perceived negative valence increased along the range of the manipulated essentialist linguistic phrasing ($r = .119, p < .001$), which is in line with previous research suggesting that variations of abstract-essentialist language are often accompanied by variations in perceived valence (e.g., Douglas & Sutton, 2006, 2010). In order to make sure that the effect of our manipulation of the linguistic-essentialist phrasing on the degree to which the prejudicial statements are (not) perceived as xenophobic is independent of their perceived negativity, we entered the mean negativity ratings of the 24 prejudice from the additional survey into the dataset of the main study. We then assessed the effect of the linguistic-essentialist phrasing on the xenophobia ratings, while controlling for these mean negativity ratings. Results from this FE model showed that the effect of the language dimension remained robust (see Table A12), suggesting that the manipulation of the essentialist linguistic phrasing affected the subtlety of the prejudicial statements, i.e., the degree to which they are perceived as (not) xenophobic, beyond their perceived negativity.

Table A12

Model of Fixed Effects of the Item Dimension Language on Xenophobia Ratings Controlling for the Items' Average Perceived Negativity Ratings

	<i>b (SE)</i>	<i>95% Confidence Interval</i>	
Model 1			
Language			
Weakly essentialist (ref.)			
Rather weakly essentialist	0.332*** (0.043)	0.247	0.416
Rather strongly essentialist	0.566*** (0.043)	0.482	0.651
Strongly essentialist	0.619*** (0.044)	0.533	0.705
Intercept	3.306*** (0.029)	3.248	3.364
R^2_{within}	.088		
$R^2_{between}$.013		
$R^2_{overall}$.039		
Model 2			
Language			
Weakly essentialist (ref.)			
Rather weakly essentialist	0.277*** (0.055)	0.169	0.384
Rather strongly essentialist	0.522*** (0.051)	0.422	0.622
Strongly essentialist	0.561*** (0.056)	0.451	0.672
Perceived negativity	0.175 (0.108)	-0.037	0.386
Intercept	2.563*** (0.459)	1.662	3.463
R^2_{within}	.089		
$R^2_{between}$.013		
$R^2_{overall}$.039		

Note. $N_{\text{Item ratings}} = 3,555$; $N_{\text{Participants}} = 895$. Unstandardized coefficients. Standard error in parentheses. The coefficient relates to a 5-point scale from 1 (*not xenophobic at all*) to 5 (*very xenophobic*). Reference category is indicated by (ref.). Model 2 controls for average perceived negativity ratings from the additional online survey for each prejudice item.

* $p < .05$; ** $p < .01$; *** $p < .001$.