Contextual effect of positive intergroup contact on outgroup prejudice

Oliver Christ[a,b,1], Katharina Schmid[c], Simon Lolliot[c], Hermann Swart[d], Dietlind Stolle[e], Nicole Tausch[f], Ananthi Al Ramiah[g], Ulrich Wagner[a], Steven Vertovec[h], and Miles Hewstone[c,1]

[a]Department of Psychology, Philipps-University Marburg, 35037 Marburg, Germany; [b]Department of Psychology, University of Hagen, 58084 Hagen, Germany; [c]Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, United Kingdom; [d]Department of Psychology, Stellenbosch University, Stellenbosch 7600, South Africa; [e]Centre for the Study of Democratic Citizenship, Department of Political Science, McGill University, Montreal, QC, Canada H3A 2T7; [f]School of Psychology and Neuroscience, University of St Andrews, St Andrews, Fife KY169JP, Scotland; [g]Department of Psychology, Yale-NUS College, Singapore 138614; and [h]Max Planck Institute for the Study of Religious and Ethnic Diversity, D-37073 Göttingen, Germany

Edited by Susan T. Fiske, Princeton University, Princeton, NJ, and approved February 3, 2014 (received for review November 8, 2013)

We assessed evidence for a contextual effect of positive intergroup contact, whereby the effect of intergroup contact between social contexts (the between-level effect) on outgroup prejudice is greater than the effect of individual-level contact within contexts (the within-level effect). Across seven large-scale surveys (five cross-sectional and two longitudinal), using multilevel analyses, we found a reliable contextual effect. This effect was found in multiple countries, operationalizing context at multiple levels (regions, districts, and neighborhoods), and with and without controlling for a range of demographic and context variables. In four studies (three cross-sectional and one longitudinal) we showed that the association between context-level contact and prejudice was largely mediated by more tolerant norms. In social contexts where positive contact with outgroups was more commonplace, norms supported such positive interactions between members of different groups. Thus, positive contact reduces prejudice on a macrolevel, whereby people are influenced by the behavior of others in their social context, not merely on a microscale, via individuals' direct experience of positive contact with outgroup members. These findings reinforce the view that contact has a significant role to play in prejudice reduction, and has great policy potential as a means to improve intergroup relations, because it can simultaneously impact large numbers of people.

diversity | trust | social norms | multilevel analysis

The world is becoming increasingly diverse, fueling debate about relations between, especially, ethnic and religious groups (1, 2). Earlier attempts to explain majority group members' prejudice toward minority outgroups proposed that social environments characterized by greater proportions of (minority) outgroup members inevitably invoke perceptions of competitive threat to the (majority) ingroup's position, provoking intergroup tension (3), and there is evidence that as minority group proportion increases, so does prejudice and threat (4–6). However, such analyses fail to include the role of intergroup contact (7). Positive intergroup contact provides a way to overcome intergroup tensions and conflict that are often associated with segregation (8), and extensive evidence shows that positive face-to-face contact, especially between cross-group friends (9), reduces outgroup prejudice among minority and especially majority group members (10). Many of these studies have used prejudice as the dependent variable, but outcome measures have also included threat, trust, and outgroup bias; we use the term "prejudice" loosely to include all these measures assessing the climate of intergroup relations. The impact of contact is, however, not limited to the effect of such direct contact. Extended contact, knowing that another ingroup member has positive outgroup contact, can also reduce outgroup prejudice (11). Almost all prior research on intergroup contact effects is, however, limited by its focus on the impact of individual, microlevel contact on prejudice. Could prejudice be a function of not only whom you know, but also where you live? If so, then this finding would elevate contact theory to a theoretical approach with macrolevel implications, and consequent policy implications for improving intergroup relations.

Though some studies have shown that for social contexts, like regions or districts, a higher context level of intergroup contact is associated with less prejudice (12), these studies did not test the contextual effect of contact (3), the difference between the effect of intergroup contact between social contexts (the between-level effect), and the effect of individual-level contact within contexts (the within-level effect) (13) on prejudice. Evidence for this contextual effect of positive contact would indicate that living in a place in which other ingroup members interact positively with members of the outgroup should reduce prejudice, beyond one's own contact experiences and irrespective of whether one knows the ingroup members experiencing intergroup contact. Thus, a person living in a context with a higher mean level of positive intergroup contact is likely to be less prejudiced than a person with the same level of direct positive contact, but living in a context with a lower mean level of intergroup contact (Fig. 1). Evidence, especially longitudinal data, for this contextual effect of contact would demonstrate that intergroup contact at the social context level has greatest consequences for individuals' attitudes (and behaviors), and that the processes involved cannot be reduced to characteristics of individuals or specific situations in which intergroup contact occurs (14), or selection bias. This

Significance

Although mixed social environments can provoke conflict, where this diversity promotes positive intergroup contact, prejudice is reduced. Seven multilevel studies demonstrate that the benefits of intergroup contact are broader than previously thought. Contact not only changes attitudes for individuals experiencing direct positive intergroup contact, their attitudes are also influenced by the behavior (and norms) of fellow ingroup members in their social context. Even individuals experiencing no direct, face-to-face intergroup contact can benefit from living in mixed settings where fellow ingroup members do engage in such contact. Two longitudinal studies rule out selection bias as an explanation for these findings on the contextual level. Prejudice is a function not only of whom you interact with, but also of where you live.

Author contributions: O.C., K.S., S.L., H.S., D.S., N.T., A.A.R., U.W., S.V., and M.H. designed research; O.C., K.S., H.S., D.S., A.A.R., and M.H. performed research; O.C. analyzed data; O.C., K.S., S.L., H.S., D.S., N.T., A.A.R., U.W., and M.H. commented on multiple prior analyses and article drafts; and O.C. and M.H. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]To whom correspondence may be addressed. E-mail: christ@staff.uni-marburg.de or miles.hewstone@psy.ox.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1320901111/-/DCSupplemental.

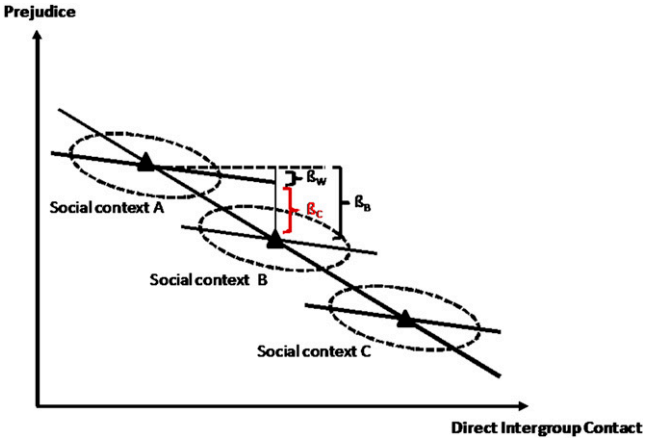

Fig. 1. Illustration of the proposed contextual effect of intergroup contact. ▲, mean level of intergroup contact within social context (group mean of intergroup contact); β_W, mean effect of intergroup contact within social contexts (within-level effect of intergroup contact); β_B, effect of intergroup contact between social contexts (between-level effect of intergroup contact); β_C, contextual effect of intergroup contact (difference between the between-level and within-level effect of intergroup contact).

evidence would make a theoretical contribution in better understanding the consequences of diversity, and a practical contribution in underlining the policy potential of contact as a social intervention to improve intergroup relations on a wider level (15, 16).

We propose that perceived ingroup norms in social contexts where positive contact with outgroups is more commonplace are more tolerant, supporting positive interactions with outgroup members (17). These norms, based on perceptions of what ingroup members think, prescribe appropriate attitudes, values, and behaviors toward outgroup members, and prejudice is dependent on norms that support or oppose these prejudices (18). Moreover, norms influence people's attitudes toward, and willingness to interact with, outgroup members (8, 19). In four of our seven studies, we tested whether living in a social context in which individuals have, on average, more positive contact is associated with more tolerant social norms within these contexts. If so, this should lead to more tolerant outgroup attitudes, over and above the effect of individual contact experiences. To approximate norms at the neighborhood level, we measured diversity beliefs, which reflect the extent to which individuals value and endorse diversity (20–22). This construct was expressly developed to capture the belief that high diversity is instrumental to accomplishing the ingroup's goals (20); it is a construct that is theoretically and statistically related to, but distinct from, outgroup prejudice (when both were measured at the individual level; see *SI Text*, footnote) (23). When estimated at the social context level, for example, the neighborhood level (i.e., as a random effect, involving the average or aggregate level of diversity beliefs in the neighborhood), our measures of diversity beliefs reflect positive social norms in the neighborhood.

Studies 1a to 1e

We first sought evidence for the contextual effect of contact in five large cross-sectional survey data sets from a range of intergroup contexts, varying in the narrowness of the social context indicator (*SI Text*). The data for study 1a were taken from the European Social Survey (24), with context measured at the regional level. Study 1b used a 2002 probability survey of the German adult population, with context measured at the district level. Study 1c relied on data from a 2005 national survey of White Americans with context measured at the census tract level. Study 1d used survey data from White British respondents in a 2009–2010 national survey in England, with context measured at the neighborhood level.* Study 1e tested the contextual effect among minority respondents (in the sense of historically disadvantaged groups) in a 2011 city survey in Cape Town, South Africa, with context measured at the neighborhood level. All five data sets measured contact (cross-group friendships) and outgroup prejudice. Studies 1a, 1d, and 1e included indicators for social norms (for item wording, see *SI Text*). All studies included a range of pertinent individual (education) controls, and three studies (1b–1d) also included context-level (e.g., regional GDP) controls (for details, see *SI Text*). We used all available controls, but only had contextual-level measures of deprivation in specific studies (*SI Text*).

To assess the contextual effect of intergroup contact, we used multilevel modeling with the multilevel latent covariate approach (25) (for details, see *SI Text*). Respondents (within-level) were nested within contexts (between-level; e.g., neighborhoods). A contextual effect is indicated when the between-level effect of intergroup contact (β_B in Fig. 1) is significantly larger than the within-level effect (β_W in Fig. 1). We assessed the magnitude of the contextual effect by calculating an effect size measure (ES2; see *SI Text*) (26).

Results are summarized in Table 1. In all analyses, both at the individual level as well as at the social context level, intergroup contact was significantly negatively related to prejudice. As predicted, in all analyses the between-level effect of intergroup contact on prejudice was significantly larger than the within-level effect, yielding a relatively small effect size of the contextual effect of contact (ES2 ranged from 0.21 to 0.35). In study 1e, which used respondents from two groups (Black and so-called "Colored" in Cape Town, South Africa), we found a significant contextual effect, providing preliminary evidence for a contextual effect among disadvantaged (or minority) groups as well.

Next, we tested whether the contextual effect could be explained by differences in social norms between the different social contexts in studies 1a, 1d, and 1e. We reestimated the contextual effect after including norms as an additional predictor of prejudice on the between-level (including norms as a predictor on the within-level, too), thus controlling for between-context differences in ingroup norms. To test whether there was a significant reduction in the contextual effect when controlling for social norms, we tested the indirect effect of intergroup contact on prejudice via social norms on the between-level effect using multilevel mediational analysis (27). A significant indirect effect provides evidence for a reduction in the contextual effect of intergroup contact when controlling for social norms.

When norms were controlled, the difference between the within-level and the between-level effect of intergroup contact was substantially reduced in all cases (see significant indirect effects in Table 1). The contextual effect was rendered nonsignificant in all tests when between-context differences in norms were statistically controlled. Thus, living in a place where fellow ingroup members interact positively with outgroup members has a benign impact on prejudice, beyond one's own contact experiences, via social norms that value diversity (11).

Studies 2a and 2b

A possible interpretation of the study 1 results is that people who are low in prejudice are more likely to select places to live that are more diverse. Although this interpretation contradicts evidence that prejudice rises with minority group proportions (3–6), it remains possible that more tolerant people select places with higher context-level contact. We sought to rule out this self-selection account by

*Study 1d included a large sample of various minority groups in the United Kingdom (n = 798; e.g., Blacks, Asians). The ICC for the prejudice measure was small (0.02), resulting in a nonconvergence of our analysis, so we could not test the contextual effect for minority group members.

Table 1. Unstandardized estimates (SE in brackets) for the contextual effect (studies 1a–1e) and the contextual effect controlling for norms (studies 1a, 1d, and 1e)

	Study 1a[†]		Study 1b		Study 1c		Study 1d		Study 1e[‡]	
	β (SE)	*P*	β (SE)	*P*	β (SE)	*P*	β (SE)	*P*	β (SE)	*P*
Within-level effect	−0.189 (0.009)	<0.001	−0.351 (0.023)	<0.001	−0.082 (0.039)	0.035	−0.555 (0.101)	<0.001	−0.21 (0.05)	<0.001
Between-level effect	−0.738 (0.099)	<0.001	−0.663 (0.062)	<0.001	−0.416 (0.162)	0.010	−1.465 (0.342)	<0.001	−0.47 (0.12)	<0.001
Contextual effect	−0.549 (0.101)	<0.001	−0.311 (0.073)	<0.001	−0.334 (0.177)	0.059	−0.910 (0.377)	0.016	−0.25 (0.14)	<0.001
Effect size of contextual effect	0.33		0.22		0.35		0.34		0.21	
Contextual effect without controls	−0.495 (0.102)	<0.001	−0.282 (0.069)	<0.001	−0.270 (0.166)	0.104	−0.929 (0.379)	0.014	−0.226 (0.131)	0.085
Contextual effect controlling for norms	−0.135 (0.084)	0.106					0.33 (0.37)	0.37	0.42 (0.27)	0.542
Indirect effect of context[§]	−0.418 (0.097)	<0.001					−1.39 (0.44)	0.002	−0.71 (0.23)	0.010

[†]All estimates are controlled for between-country differences in variables.
[‡]The contextual effect for all groups is reported; due to sample sizes, further differentiation by ethnic group was not possible.
[§]The indirect effect reflects the effect of contact on prejudice via norms (assessed as diversity beliefs) on the social context level, and indicates whether the contextual effect is significantly reduced after controlling for norms on the social context level.

testing the contextual effect in two studies using longitudinal data (studies 2a and 2b), with controls, enabling us to compare the relation of intergroup contact with prejudice over time on the individual and social context levels (Fig. 1). Moreover, detecting a contextual effect of contact using longitudinal data would constitute a stronger case for the effect, compared with cross-sectional data, because we can demonstrate that contact at time 1 is associated with reduced prejudice at time 2, thus overcoming the so-called "causal sequence problem." In study 2a, only indicators for intergroup contact and prejudice were available, whereas study 2b also measured social norms (for item wording, see *SI Text*). Data for study 2a were drawn from respondents with no migration background, who participated in two waves (time 1, 2002; time 2, 2006) of a multiwave panel study representative of the German adult population. For study 2b, respondents, randomly sampled from 16 different cities in Germany, were surveyed in 2010 and 2011.

We estimated the contextual effect of intergroup contact over time using a multilevel cross-lagged panel model. We compared the cross-lagged effect of intergroup contact at time 1 on prejudice at time 2 at the social context level (between-level) with the same cross-lagged effect on the individual level (within-level), controlling for the autoregressive effects (associations between the same

Table 2. Unstandardized estimates (SE in brackets) for the autoregressive and cross-lagged effects in studies 2a and 2b

	Study 2a		Study 2b	
Model	β (SE)	*P*	β (SE)	*P*
Model 1				
Level 1				
$contact_{time1} \rightarrow contact_{time2}$	0.583 (0.031)	<0.001	0.624 (0.039)	<0.001
$prejudice_{time1} \rightarrow prejudice_{time2}$	0.611 (0.024)	<0.001	0.677 (0.027)	<0.001
$contact_{time1} \rightarrow prejudice_{time2}$	−0.055 (0.029)	0.059	−0.002 (0.004)	0.547
$prejudice_{time1} \rightarrow contact_{time2}$	−0.052 (0.026)	0.044	−0.560 (0.187)	0.003
Level 2				
$contact_{time1} \rightarrow contact_{time2}$	0.910 (0.101)	<0.001	0.984 (0.169)	<0.001
$prejudice_{time1} \rightarrow prejudice_{time2}$	0.439 (0.173)	0.011	0.755 (0.209)	<0.001
$contact_{time1} \rightarrow prejudice_{time2}$	−0.298 (0.078)	<0.001	−0.039 (0.015)	0.009
$prejudice_{time1} \rightarrow contact_{time2}$	−0.101 (0.206)	0.624	2.848 (2.107)	0.177
Contextual effect	−0.243 (0.085)	0.004	−0.037 (0.016)	0.024
Effect size of contextual effect	0.29		0.17	
Model 2				
Level 2				
$contact_{time1} \rightarrow contact_{time2}$			0.989 (0.160)	<0.001
$prejudice_{time1} \rightarrow prejudice_{time2}$			0.467 (0.221)	0.034
$contact_{time1} \rightarrow prejudice_{time2}$			−0.043 (0.015)	0.006
$prejudice_{time1} \rightarrow contact_{time2}$			4.028 (2.492)	0.106
$norms_{time1} \rightarrow norms_{time2}$			0.391 (0.146)	0.007
$contact_{time1} \rightarrow norms_{time\ 2}$			0.084 (0.030)	0.005
$norms_{time1} \rightarrow prejudice_{time2}$			−0.146 (0.057)	0.011
Contextual effect controlling for norms			0.036 (0.017)	0.033
Indirect effect of context			0.012 (0.007)	0.086

Note. Model 1 considers only the relationships between contact and prejudice; model 2 adds norms as a mediator. For model 2, level 1 results are shown in Table S1.

measures over time; for complete results of autoregressive and cross-lagged effects, see Table 2). For study 2a, the cross-lagged effect of intergroup contact was negative and significant at the between level ($\beta = -0.298$, SE = 0.078, $P < 0.001$), and negative and marginally significant at the within level ($\beta = -0.055$, SE = 0.029, $P = 0.059$). Thus, at both levels, more intergroup contact was associated with less prejudice over time. There was a small contextual effect of intergroup contact over time (ES2 = 0.29): the between-level cross-lagged effect of intergroup contact was significantly larger than the within-level cross-lagged effect ($\beta = -0.243$, SE = 0.085, $P = 0.004$; without controls: $\beta = -0.223$, SE = 0.085, $P = 0.009$). The between-level cross-lagged effect of prejudice on contact was nonsignificant ($\beta = -0.101$, SE = 0.206, $P = 0.624$), ruling out selection bias.

For study 2b, the cross-lagged effect of intergroup contact was negative and significant at the between level ($\beta = -0.039$, SE = 0.015, $P = 0.009$), and negative, but nonsignificant at the within level ($\beta = -0.002$, SE = 0.004, $P = 0.547$). Only at the between level was more intergroup contact associated with less prejudice over time. The between-level cross-lagged effect of intergroup contact was significantly larger than the within-level cross-lagged effect ($\beta = -0.037$, SE = 0.016, $P = 0.024$; without controls: $\beta = -0.037$, SE = 0.016, $P = 0.020$), demonstrating longitudinally a small contextual effect (ES2 = 0.17) of intergroup contact. Again, the between-level cross-lagged effect of prejudice on contact was nonsignificant ($\beta = 2.848$, SE = 2.107, $P = 0.177$). In study 2b we also tested the mediational effect of social norms longitudinally, which was negative and approached significance ($\beta = -0.012$, SE = 0.007, $P = 0.086$); the contextual effect remained significant ($\beta = -0.036$, SE = 0.017, $P = 0.033$). Although the power in study 2b was low at the social context level ($n = 50$), the results support the assumption that the contextual effect of intergroup contact can be partly explained by changes in social norms over time. In study 2b, on the within level, the significant longitudinal effect of time 1 prejudice on time 2 contact, together with the absence of a longitudinal effect from contact to prejudice, supports a pattern of self-selection (prejudice leads to contact). We have no ready explanation for this unpredicted result, but selection effects have been found in prior research (28). However, this result does not undermine our longitudinal demonstration of the contextual effect, an effect that cannot be explained with selection bias.

These findings show that the contextual effect cannot be totally explained with self-selection (because in the two longitudinal studies we control for selection). Selection would lead to a causal order from prejudice to norms to contact. We show, longitudinally, the opposite direction at the context level. Although prejudiced people might avoid individual contact, they still profit from the "contextual effect of contact," i.e., people in general have more intergroup contact in their environment.

One possible alternative explanation of the contextual effect is that individuals high in prejudice avoid social contexts in which people have frequent intergroup contact. Thus, in social contexts with a high mean level of prejudice, the contextual effect should be smaller or even absent. In studies 2a and 2b, we were able to test whether the contextual effect was dependent on the mean prejudice level within the social contexts. We included the interaction between contact and prejudice at time 1, both measured on the between level. In study 2a, the interaction was negative and significant ($\beta = -0.201$, SE = 0.084, $P = 0.017$), whereas in study 2b, the interaction was negative but not significant ($\beta = -0.008$, SE = 0.020, $P = 0.677$). Results for study 2b showed that the contextual effect was not influenced by the mean level of prejudice within a social context. In study 2a, the contextual effect was even stronger in social contexts with a high mean level of prejudice. Together, these results do not support self-selection as a possible alternative explanation of our data on the contextual effect.

Discussion

To conclude, our data show consistently across seven studies that individuals' outgroup attitudes are more positive when living in social contexts in which people have, on average, more positive intergroup contact. Moreover, we found a consistent contextual effect of contact on prejudice in each study: indeed, the effect of intergroup contact between social contexts is greater than the effect of individual-level contact within contexts. In four studies we provided evidence that this contextual effect is accompanied by more tolerant social norms that possibly explain the larger effect of intergroup contact on the social-context level of analysis. Thus, positive intergroup contact is associated with reduced prejudice on a macro- and not merely microlevel, whereby people are influenced by the behavior of others in their wider social context.

Three key considerations speak to the robustness of our findings. First, results were replicated over seven studies, using a range of measures (of contact, norms, and prejudice), contexts (regions, districts, and neighborhoods), respondents (from both majority and minority groups), and countries. Second, we obtained evidence for the contextual effect both with no statistical controls and with several controls at the individual and context level. Third, we demonstrated the contextual effect in two longitudinal studies. This finding confirms that, over time, the context-level effect of contact is greater than its individual-level effect, and that contact impacts prejudice. In sum, we are confident that our findings are high in generalizability, reliable, and demonstrate an effect from positive contact to reduced prejudice.

We do, however, acknowledge some limitations of our program of research and hence areas for future research. First, the measures of some key constructs in some studies were suboptimal (measured with a single indicator, or as difference scores) because we sought to demonstrate the contextual effect using both archival and our own data from a wide range of contexts; notwithstanding, the effect size of the contextual effect is robust across studies (Tables 1 and 2), so we may well have underestimated the size of the contextual effect. Second, thus far the longitudinal evidence is based exclusively on German data, and these results should be replicated in other contexts. Third, in two of the studies (studies 1a and 1e), context-level controls were not available, although given that we detected contextual effects in four other studies that included such controls, it is unlikely that results reported are solely attributable to some other variable. Finally, in study 1e, we reported a significant contextual effect for lower status groups in Cape Town, South Africa, but due to sample size we were unable to disaggregate results for Coloreds (the numerical majority in this city) and Blacks (both a numerical and social minority), in a society where Whites continue to benefit from the historical advantages of their majority group status. Future research should test the contextual effect on minority groups who have lower status and are in a numerical minority.

These findings have two notable implications. First, macrolevel diversity should not be equated with actual intergroup contact. It is not sufficient to report the proportion of outgroup members in an area; one must report the extent to which members of different groups engage in positive contact. Second, contact has even more beneficial effects than was previously thought (10, 15). Contact does not merely change attitudes on a microscale, in the case of those people who experience direct positive contact with members of the outgroup, nor do interventions on that microlevel offer the only means of reducing prejudice. Rather, contact also affects prejudice on a macrolevel, whereby people are influenced by the behavior of others in their social context. Prior research that has prioritized the interpersonal nature of contact has ignored its potential widespread impact. Even individuals who have no direct intergroup contact experience can benefit from living in mixed

settings, provided that fellow ingroup members do engage in positive intergroup contact: Prejudice is a function not only of whom you know, but also of where you live.

These findings demonstrate the policy potential of contact at the context level, because it can be implemented via macrolevel contexts such as mixed schools, neighborhoods, and workplaces. Our research demonstrates the value of living in mixed settings where positive intergroup interaction occurs, over and above positive effects of each individual's own positive contact experiences. This potential positive impact of diversity, via intergroup contact, is, however, constrained by segregation (29), which precludes contact. The full potential of positive intergroup contact can only be realized with a reduction in segregation that results in increased opportunities for contact, and of course when members of different groups take up those opportunities and engage in more frequent, positive, face-to-face contact.

ACKNOWLEDGMENTS. This work was supported by Deutsche Forschungsgemeinschaft Research Fellowship CH 743, 2-1 (to O.C.) and a Leverhulme Trust Programme Grant (to M.H.). Data collection was supported by Volkswagen Stiftung and Freudenberg Stiftung (studies 1b and 2a; awarded to Wilhelm Heitmeyer, University of Bielefeld); the Weidenbaum Center on the Economy, Government, and Public Policy (the U.S. Citizenship, Involvement, Democracy Survey used in study 1c, which was commissioned by Georgetown University's Center for Democracy and Civil Society); the Leverhulme Trust (study 1d); and the Max Planck Institute for the Study of Religious and Ethnic Diversity (studies 1d, 1e, 2b).

1. Hooghe M, Reeskens T, Stolle D, Trappers A (2009) Ethnic diversity and generalized trust in Europe: A cross-national multilevel study. *Comp Polit Stud* 42(2):198–223.
2. Putnam RD (2007) E Pluribus Unum: Diversity and community in the twenty-first century. The 2006 Johan Skytte Prize Lecture. *Scand Polit Stud* 30(2):137–174.
3. Blalock HM (1984) *Annual Review of Sociology*, eds Turner RH, Short, Jr JF (Annual Reviews, Inc., Palo Alto, CA), Vol 10, pp 353–372.
4. Quillian L (1995) Prejudice as a response to perceived group threat: Population composition and anti-immigrant and racial prejudice in Europe. *Am Sociol Rev* 60(4): 586–611.
5. Quillian L (1996) Group threat and regional change in attitudes toward African Americans. *Am J Sociol* 102(3):816–860.
6. Taylor MC (1998) How white attitudes vary with the racial composition of local populations: Numbers count. *Am Sociol Rev* 63(4):512–535.
7. Schmid K, Al Ramiah A, Hewstone M (2014) Neighborhood ethnic diversity and trust: The role of intergroup contact and perceived threat. *Psychol Sci*, in press.
8. Allport GW (1954) *The Nature of Prejudice* (Addison-Wesley, Cambridge, MA).
9. Davies K, Tropp LR, Aron A, Pettigrew TF, Wright SC (2011) Cross-group friendships and intergroup attitudes: A meta-analytic review. *Pers Soc Psychol Rev* 15(4):332–351.
10. Pettigrew TF, Tropp LR (2006) A meta-analytic test of intergroup contact theory. *J Pers Soc Psychol* 90(5):751–783.
11. Wright SC, Aron A, McLaughlin-Volpe T, Ropp SA (1997) The extended contact effect: Knowledge of cross-group friendships and prejudice. *J Pers Soc Psychol* 73(1):73–90.
12. Wagner U, Christ O, Pettigrew TF, Stellmacher J, Wolf C (2006) Prejudice and minority proportion: Contact instead of threat effects. *Soc Psychol Q* 69(4):380–390.
13. Raudenbush SW, Bryk AS (2002) *Hierarchical Linear Models* (Sage, Newbury Park, CA), 2nd Ed.
14. Oishi S, Graham J (2010) Social ecology: Lost and found in psychological science. *Perspect Psychol Sci* 5(4):356–377.
15. Hewstone M (2009) Living apart, living together? The role of intergroup contact in social integration. *Proc Br Acad* 162:243–300.
16. Pettigrew TF, Tropp LR (2011) *When Groups Meet: The Dynamics of Intergroup Contact* (Psychology Press, Philadelphia).
17. De Tezanos-Pinto P, Bratt C, Brown R (2010) What will the others think? In-group norms as a mediator of the effects of intergroup contact. *Br J Soc Psychol* 49(3): 507–523.
18. Crandall CS, Eshleman A, O'Brien L (2002) Social norms and the expression and suppression of prejudice: The struggle for internalization. *J Pers Soc Psychol* 82(3): 359–378.
19. Pettigrew TF (1998) Intergroup contact theory. *Annu Rev Psychol* 49:65–85.
20. Tropp LR, Bianchi RA (2006) Valuing diversity and interest in intergroup contact. *J Soc Issues* 62(3):533–551.
21. van Knippenberg D, Haslam SA, Platow MJ (2007) Unity through diversity: Value-in-diversity beliefs, work group diversity, and group identification. *Group Dyn* 11(3): 207–222.
22. van Dick R, van Knippenberg D, Hägele S, Guillaume YRF, Brodbeck FC (2008) Group diversity and group identification: The moderating role of diversity beliefs. *Hum Relat* 61(10):1463–1492.
23. Adesokan AA, Ullrich J, van Dick R, Tropp LR (2011) Diversity beliefs as moderator of the contact-prejudice relationship. *Soc Psychol* 42(4):271–278.
24. Norwegian Social Science Data Services (2002) *European Social Survey Round 1 Data.* Data file edition 6.2. Available at www.europeansocialsurvey.org/data/download.html?r=1.
25. Lüdtke O, Marsh HW, Robitzsch A, Trautwein U (2011) A 2×2 taxonomy of multilevel latent contextual models: Accuracy-bias trade-offs in full and partial error correction models. *Psychol Methods* 16(4):444–467.
26. Marsh HM, et al. (2009) Doubly-latent models of school contextual effects: Integrating multilevel and structural equation approaches to control measurement and sampling error. *Multivariate Behav Res* 44(6):764–802.
27. Preacher KJ, Zyphur MJ, Zhang Z (2010) A general multilevel SEM framework for assessing multilevel mediation. *Psychol Methods* 15(3):209–233.
28. Binder J, et al. (2009) Does contact reduce prejudice or does prejudice reduce contact? A longitudinal test of the contact hypothesis among majority and minority groups in three European countries. *J Pers Soc Psychol* 96(4):843–856.
29. Uslaner EM (2012) *Segregation and Mistrust: Diversity, Isolation, and Social Cohesion* (Cambridge Univ Press, New York).

Supporting Information

Christ et al. 10.1073/pnas.1320901111

SI Text

Computation of Contextual Effect. We followed Raudenbush and Bryk (1) when estimating the contextual effect of intergroup contact by specifying the following relation:

$$\text{Level 1}: \text{prejudice}_{ij} = \beta_{0j} + \beta_1 \text{direct intergroup contact}_{CWC} + r_{ij}, \quad \textbf{[S1]}$$

where prejudice_{ij} is the outcome for respondent i in context j modeled as a function of the intercept β_{0j} of context j, the slope β_{1j} for direct intergroup contact in context j, and an error r_{ij}. The predictor, direct intergroup contact, is centered at the group mean (centering within cluster), removing all between-context variation in direct intergroup contact and yielding a pooled-within (i.e., level 1) estimate (β_W) for the relation of direct intergroup contact with prejudice. The level 1 coefficients β_{0j} and β_{1j} are then modeled at level 2. Level 2 coefficients are typically notated as γ.

$$\text{Level 2}: \begin{matrix} \beta_{0j} = \gamma_{00} + \gamma_{01} \text{direct intergroup contact}_{\,Group\ Mean} + u_{0j} \\ \beta_{1j} = \gamma_{10,} \end{matrix} \quad \textbf{[S2]}$$

where γ_{00} and γ_{10} are the level 1 intercepts, and γ_{01} is the slope relating the group mean of direct intergroup contact to the intercepts from the level 1 equation. The slope γ_{01} captures the between-context relation between direct intergroup contact (mean of intergroup contact within contexts) and prejudice. It is important to note that we model a random-intercept model (1) here; therefore, only the level 1 intercepts have a level 2 residual u_{0j}. The effect of direct intergroup contact on prejudice at level 1 is fixed, allowing for no variation in this effect between contexts. A contextual effect is present if γ_{01} is significantly larger or smaller than γ_{10}, meaning that the relationship at the aggregated level is stronger or weaker than the relationship at the individual level.

As Eq. **S2** shows, at level 2 the group mean is used as an estimate for the level 2 effect of intergroup contact. However, sampling error in the group mean can cause biased and less efficient estimates of the true population effect (here γ_{01}) with a consequential biased estimate for the contextual effect, therefore a multilevel latent covariate (MLC) approach that corrects for the unreliability in the level 2 construct has been recommended, resulting in unbiased estimates of the level 2 effect (2).* We therefore applied the MLC approach to estimate the contextual effect of intergroup contact in all studies. We implemented MLC using the maximum-likelihood procedure in Mplus 6.1 (3). Due to "missingness," we used full-information maximum-likelihood estimates with robust SEs. Missing values in no case exceeded 2.2%.

To assess the effect size of the contextual effect of contact we used an effect size measure (ES2) proposed by Marsh et al. (4). The effect size is calculated with the following formula:

$$ES2 = \frac{2 * B * SD_{intergroup\ contact}}{\sigma^2_{intergroup\ contact}}, \quad \textbf{[S3]}$$

where B is the unstandardized regression coefficient of the contextual effect in the multilevel model, $SD_{intergroup\ contact}$ is the SD of intergroup contact at the social context level, and σ^2 is the total variance of intergroup contact at the individual level. The resulting effect size describes the difference in the dependent variable between two level 2 groups that differ by two SDs on the predictor variable. This effect size is comparable with Cohen's d (5). An effect can be defined as small with d (or ES2) = 0.2, medium with d (or ES2) = 0.5, and large with d (or ES2) = 0.8.

A prerequisite for estimating the contextual effect of intergroup contact is sufficient between-level variance in all relevant measures. In studies 1a to 1e, the intraclass correlations (ICC) for all indicators and the composite measures of direct intergroup contact, ingroup norms, and prejudice were small to large in size (M = 0.17, SD = 0.08, minimum = 0.07, maximum = 0.30), showing that there was substantial between-level variance. Likewise, ICCs in studies 2a and 2b were small to medium for intergroup contact, social norms, and prejudice at time 1 and 2 (M = 0.08, SD = 0.06, minimum = 0.02, maximum = 0.18), indicating sufficient between-level variability in these measures.

SI Materials and Methods: Sampling Information and Measures in Study 1 and Study 2

Study 1a. ***Sampling information.*** Data came from the first round of the European Social Survey (ESS) (6). This representative cross-national survey was conducted from September 2002 to October 2003 and covered 22 countries (21 European countries and Israel). In total, 42,359 face-to-face interviews were achieved. We dropped all respondents without national citizenship, with place of birth outside the country of data collection, or who classified themselves as belonging to a minority ethnic group in their country, resulting in a reduced sample size for all analyses of n = 36,334 respondents (for four respondents, regional codes were not available).

As the regional level measure, we used the country-specific indicator variables available in the ESS to group respondents into regional units corresponding to the nomenclature of statistical units classification scheme (NUTS) (7). The NUTS classifies European regions according to socioeconomic, cultural, and historical characteristics (7). Conceptually, the NUTS comprises three different regional subdivisions that divide each country into large-scale (NUTS 1), medium-scale (NUTS 2) and small-scale (NUTS 3) regions. However, the NUTS levels provided by the ESS differed somewhat between countries. Whereas for the majority of countries, respondents are grouped into NUTS 2 regions, for four countries, respondents were grouped according to NUTS 1. To base our analyses on regions of comparable size, we reclassified the NUTS 2 codes into NUTS 1 codes and recalculated all analyses using only NUTS 1 regions (n = 91 NUTS 1 regions). The pattern of results was almost identical compared with the analyses using the mix of NUTS 1 and NUTS 2 regions. We therefore report only the results using the original regional codes provided with the ESS data (n = 248 NUTS 1/NUTS 2 regions). The mean number of

*As Lüdtke et al. (2) have recently shown, the estimation of the contextual effect cannot only be biased due to sampling error, but also due to measurement error. The authors (2) distinguished between different approaches to correct for sources of error in estimating contextual effects, proposing a 2 × 2 taxonomy of multilevel contextual models correcting for no error source, for only one source of error, or for all error sources. Lüdtke et al. (2) showed in a simulation study that depending on specific data circumstances, the uncorrected and the partial correction approaches can result in biased estimates of the contextual effect. However, when the data provides only limited information on the level 2 constructs (i.e., small number of groups, low intraclass correlations), partial correction approaches outperform the doubly latent approach. The authors therefore suggest that researchers juxtapose the different approaches (where possible) and use the estimates from the different approaches as bounds for the true parameter. We were able to implement these different approaches in study 1b because multiple items for intergroup contact and prejudice were available and therefore latent variables on both levels could be specified. In all approaches, a significant estimate for the contextual effect emerged, ranging from −0.142 to −0.331.

respondents per NUTS 1/ NUTS 2 region was $M = 146.51$. In all analyses, we controlled for between-country differences, using country as a level 3 unit in the analyses.

Measures. Direct intergroup contact was measured with a single indicator: "Do you have any friends who have come to live in [country] from another country?" (1 = no, none at all, 2 = yes, a few, 3 = yes, several).[†]

Prejudice toward foreigners was assessed with four items (Cronbach's $\alpha = 0.72$): "Average wages and salaries are generally brought down by people coming to live and work here"; "People who come to live and work here generally harm the economic prospects of the poor more than the rich"; "If people who have come to live and work here are unemployed for a long period, they should be made to leave"; "If people who have come to live here commit any crime, they should be made to leave" (1 = disagree strongly to 5 = agree strongly).

Social norms were measured with two items ($r = 0.61$, $P < 0.001$): "Would you say that [country]'s cultural life is generally undermined or enriched by people coming to live here from other countries?" (1 = cultural life undermined to 10 = cultural life enriched); "Is [country] made a worse or a better place to live by people coming to live here from other countries?" (1 = worse place to live to 10 = better place to live).[‡]

Control variables were age, sex, and education on the individual level. There were no controls available on the social context level.

Study 1b. ***Sampling information.*** Data were obtained from a probability survey of the German adult population (16 y of age and older) conducted in May/June 2002, excluding those with a migration background. Respondents were randomly selected from a two-stage probability sample, resulting in a representative sample of the German adult population. A total of $n = 2,722$ respondents were interviewed by a survey company using computer-assisted telephone interviews (CATI). These data contain district codes that indicate the place of residence of each respondent interviewed. A district is a state organizational unit usually composed of a big city or a number of smaller cities, towns, or rural areas. Sizes of districts vary between 35,700 and 3,382,200 inhabitants. Altogether, Germany is divided into 440 districts, of which 418 districts were sampled. The mean number of observations per district was $M = 6.50$ respondents.

Measures. Direct intergroup contact was measured with three items ("How many of your friends are foreigners living in Germany?"; "How often have you had an interesting conversation with a foreigner?"; "How often have you been helped by a foreigner?"; Cronbach's $\alpha = 0.75$). All items were answered using four-point rating scales ranging from 1 = none/never, 2 = few/sometimes, 3 = fairly many/often, to 4 = very many/very often.

Prejudice toward foreigners was assessed with three items ("If jobs become scarce, foreigners should be sent back to their home countries?"; "There are too many foreigners in Germany"; "Foreigners are a burden for our social security system"; Cronbach's $\alpha = 0.82$). Each item was answered on a four-point rating-scale (1 = fully disagree to 4 = fully agree).

Control variables were age, sex, and education on the individual level and an index of regional deprivation (gross domestic product, unemployment rate, rate of people receiving social welfare) on the social context level.

Study 1c. ***Sampling information.*** The data were collected from mid-May to mid-July 2005 as part of the U.S. Citizenship, Involvement, Democracy Survey conducted by the Center for Democracy and Civil Society at Georgetown University (8). This national survey is comprised of 1,001 face-to-face interviews of adults throughout the United States. We restricted our analyses to White respondents ($n = 725$) because sample sizes for Blacks, Hispanics, and Asians were too small for analysis. As the regional level measure, we used information at the level of census tracts ($n = 174$). Census tracts generally have a population size between 1,200 and 8,000 people. The mean number of observations per district was $M = 4.08$ respondents.

Measures. Direct intergroup contact was measured with one item: "Now I want to ask you some questions about people you are really close to, that is, people you feel at ease with and can talk to about whatever's on your mind, or call on for help. Though this may include family members, in the questions that follow I will refer to these people as your close friends. How many of your close friends are of a different race from yours? By race I mean such groups as Asians, Blacks, Hispanics, and Whites" (1 = none to 9 = all).

Prejudice toward Blacks, Hispanics, and Asians was assessed with the difference between an indicator for liking Whites (the ingroup) and a composite based on three indicators for liking Blacks, Hispanics, and Asians, respectively ("How do you feel about the following groups, in general?"; 1 = dislike a great deal to 11 = like a great deal).

Control variables were age, sex, education, and income on the individual level, and educational level and income on the social context level. Moreover, on the individual level, we also controlled for the quantity of close friends (whether ingroup or outgroup, to control for being more outgoing or sociable, which may be related to contact and/or prejudice).

Study 1d. ***Sampling information.*** The data came from a survey in England (2009–2010) with $n = 868$ White British respondents (level 1) from $n = 217$ neighborhoods (level 2; mean number of observations per neighborhood was $M = 4.00$). Neighborhoods constituted so-called middle-layer super output areas in England, which are small geographical units with an average size of 7,200 residents. Data collection was subcontracted to a professional survey organization that used computer-assisted personal interviewing by trained social survey interviewers, involving face-to-face interviews in respondents' own homes. Random location quota

[†]An anonymous reviewer of an earlier version of this paper queried whether respondents all understand the distinction between "yes, a few" (coded 2) and "yes, several" (coded 3) in the same manner. As an authority on survey methodology has noted, though there is random and systematic variation in the meanings of some verbal labels to respondents, "many labels do appear to have sufficiently universal meanings to be used in attitude measurement in this manner" (9, p. 151). If there were such random and/or systematic variation in understanding this distinction, this would likely cause an unsystematic relation between contact and prejudice scores at these two levels of the contact measure. A trend test using regression analysis showed, however, a strong and significant linear trend between the contact and the prejudice measure ($\beta = -0.66$, $t = 16.75$, $P < 0.001$) over the whole scale of contact and prejudice, and a significant but smaller quadratic trend ($\beta = 0.09$, $t = 8.61$, $P < 0.001$). Although the linear trend is less steep on higher ratings in the friendship measure (i.e., between scale point 2 and scale point 3), there is still a strong relationship between contact and prejudice on the two higher scale points, as corroborated by a significant difference in prejudice scores between these two scale points ($M = 3.05$, SD = 0.85 for respondents with score 2 on the contact measure; $M = 2.84$, SD = 0.85 for respondents with score 3 on the contact measure; $t = 14.02$, df = 16, 022, $P < 0.001$). These results do not support the assumption that respondents vary systematically in their understanding of the distinction between scale points 2 and 3. Finally, given that the size of the contextual effect in this study is comparable to those found in the other studies, we are confident in the robustness and validity of our findings in study 1a.

[‡]Exploratory factor analysis of the data from study 1a confirmed that our measures for norms and prejudice are related, but separable, constructs. In study 1a, we were able to use multilevel exploratory factor analysis (ML-EFA) as implemented in Mplus (3) to examine the factorial structure for both the norms and prejudice measure on the individual and social context levels simultaneously. In Mplus, ML-EFA is based on maximum-likelihood estimates, allowing us to compare different factorial solutions by means of fit statistics known from a structural equation modeling framework. The two within- and two between-factor solutions (involving separate factors of norms and prejudice on both levels, respectively) showed the best fit to the data compared with all other possible combinations ($\chi^2 = 2,771.42$, df = 8, $P < 0.001$; Comparative Fit Index = 0.894; RMSEA = 0.098; Standardized Root Mean Square Residual $(SRMR)_{within} = 0.041$; $SRMR_{between} = 0.043$). These findings are consistent both with prior theoretical work on the construct of diversity beliefs (10) and prior empirical work distinguishing diversity beliefs and prejudice (11). We could not test the factorial structure of norms and prejudice in studies 1d, 1e, and 2b, due to single-item measures of prejudice in studies 1d and 1e, and small sample size on the social context level in study 2b. Full results of the EFA are available from the first author.

sampling ensured that the profile of respondents interviewed in each neighborhood reflected the profile of the neighborhood with regard to key demographics (age, sex, working status, and ethnicity). Data collection took place from October 2009 to February 2010.

Measures. Direct intergroup contact was measured with a single item: "What proportion of your close friends are from ethnic minorities?" (1 = none or very few, 2 = a few, 3 = about half, 4 = a lot, 5 = almost all or all).

Prejudice toward ethnic minorities was assessed with an ingroup bias measure computed by subtracting the outgroup rating from the ingroup rating: "How warm or cold do you feel about White British people?" and "How warm or cold do you feel about ethnic minorities?" (feeling thermometer ranged from 0 = cold to 100 = warm).

Social norms were measured with two items ($r = 0.40$, $P < 0.001$): "The mix of different ethnic groups in my neighborhood enriches local life" and "The mix of different ethnic groups in my neighborhood creates social disorder" (reverse coded; 1 = definitely disagree, 2 = tend to disagree, 3 = neither agree nor disagree, 4 = tend to agree, 5 = definitely agree).

Control variables were age, sex, and education on the individual level and an index of multiple deprivation score on the social context level.

Study 1e. ***Sampling information.*** Data came from a city survey in Cape Town, South Africa, conducted in September and November 2011 with $n = 897$ respondents (level 1) from $n = 97$ neighborhoods (level 2; mean number of observations per neighborhood was $M = 9.25$ respondents). The survey included information from Black ($n = 438$) and Colored respondents ($n = 459$). Data collection was subcontracted to a professional survey organization that used trained interviewers to undertake face-to-face interviews in respondents' own homes.

Measures. Direct intergroup contact with White South Africans was measured with one item: "What proportion of your close friends are White people?" [1 = none to 5 = (almost) all].

Prejudice toward White South Africans was measured with one item asking how much the respondents felt that White South Africans can be trusted (1 = cannot trust them at all to 5 = all can be trusted). This item was reverse coded for the analyses, so that high scores denoted prejudice.

Social norms were measured with four items (Cronbach's $\alpha = 0.86$): "How important do you think it is to have people from different racial backgrounds in your workplace, or place of study?"; "How important do you think it is to have people from different racial backgrounds among your friends?"; "How important do you think it is to have people from different racial backgrounds in your neighborhood?"; and "How important do you think it is to have people from different racial backgrounds in South Africa?" (1 = not at all important to 5 = very important).

Control variables were age, sex, and education on the individual level. There were no controls available on the social context level.

Study 2a. ***Sampling information.*** Data were collected as waves 1 and 4 of a multiwave panel study representative of the German adult population (16 y and above) with no migration background. A total of 1,024 respondents were interviewed via a survey company using CATI in both waves in 2002 (time 1) and 2006 (time 2). These respondents form a part of the sample used in study 1b. As such, the level 2 units were the same districts as in study 1b. We sampled a random subsample of respondents from study 1b who agreed to be recontacted for the panel survey ($n = 2,363$; response rate = 43%). Therefore, the number of districts was reduced to $n = 345$; the average number of respondents within districts was $n = 2.97$. Missing data were negligible (<1%). Systematic panel mortality was negligible.

Measures. The panel included the same indicators for direct intergroup contact (Cronbach's $\text{alpha}_{\text{time 1}} = 0.73$; Cronbach's $\text{alpha}_{\text{time 2}} = 0.75$) and prejudice (Cronbach's $\text{alpha}_{\text{time 1}} = 0.81$; Cronbach's $\text{alpha}_{\text{time 2}} = 0.82$) as used in study 1b.

Control variables were age, sex, and education on the individual level, and an index of regional deprivation (gross domestic product, unemployment rate, rate of people receiving social welfare) on the social context level.

Study 2b. ***Sampling information.*** Data were collected in Germany by a professional survey organization, using trained social survey interviewers and CATI (www.mmg.mpg.de/en/publications/working-papers/2012/wp-12-21/). Respondents were purposefully sampled from neighborhoods varying in their proportional share of foreign residents, resulting in a two-level hierarchical data structure with respondents nested in neighborhoods. Fifty neighborhoods (so-called "Wohnviertel"; minimum $n = 2,800$ residents, average $n = 7,500$ residents) from 16 different cities in Germany were randomly sampled. Data collection took place from May to July 2010 for wave 1 ($n = 1,976$) and May to July 2011 for wave 2 ($n = 1,056$; response rate: 53.44%). The final sample size for this study was $n = 1,056$ respondents (level 1) from $n = 50$ neighborhoods (level 2) who took part at both time points. The average number of respondents within districts was $n = 21.12$. Missing data were negligible (<1%), as was systematic panel mortality.

Measures. Direct intergroup contact was measured with an index based on the product of two items assessing the frequency and quality of contact with foreigners within the neighborhood of the respondents: "In your neighborhood, how often do you talk to people who are themselves not native Germans or whose parents are not from Germany?" and "How do you perceive the conversations with immigrants in your neighborhood?". Both items were answered on five-point rating scales (1 = never to 5 = daily; 1 = very unpleasant to 5 = very pleasant).

Prejudice toward foreigners was assessed with four items (Cronbach's $\text{alpha}_{\text{time 1}} = 0.72$; Cronbach's $\text{alpha}_{\text{time2}} = 0.73$): "Foreigners in Germany threaten the German way of life"; "The values of foreigners living in Germany are incompatible with the values of Germans"; "Foreigners living in Germany make it more difficult for Germans to find jobs"; and "Foreigners living in Germany are a burden on the social welfare system". The items were answered on a five-point rating scale (1 = fully disagree to 5 = fully agree).

Social norms were measured with two items ($r_{\text{time 1}} = 0.48$, $P < 0.001$; $r_{\text{time2}} = 0.48$, $P < 0.001$): "It is enriching for a city when people come from different backgrounds and cultures" and "Muslims living in Germany should have the right to build mosques, including in your own neighborhood". The items were answered on a five-point rating scale (1 = fully disagree to 5 = fully agree).

Control variables were age, sex, and education on the individual level, and unemployment rate on the social context level.

1. Raudenbush SW, Bryk AS (2002) *Hierarchical Linear Models* (Sage, Newbury Park, CA), 2nd Ed.
2. Lüdtke O, Marsh HW, Robitzsch A, Trautwein U (2011) A 2 × 2 taxonomy of multilevel latent contextual models: Accuracy-bias trade-offs in full and partial error correction models. *Psychol Methods* 16(4):444–467.
3. Muthén LK, Muthén BO (1998–2010) *Mplus User's Guide* (Muthén & Muthén, Los Angeles), 6th Ed.
4. Marsh HM, et al. (2009) Doubly-latent models of school contextual effects: Integrating multilevel and structural equation approaches to control measurement and sampling error. *Multivariate Behav Res* 44(6):764–802.
5. Cohen J (1988) *Statistical Power Analysis for the Behavioral Sciences* (Erlbaum, Mahwah, NJ).
6. Norwegian Social Science Data Services (2002) *European Social Survey Round 1 Data*. Data file edition 6.2. Available at www.europeansocialsurvey.org/data/download.html?r=1.
7. Eurostat (2003) *European Regional Statistics Reference Guide* (Office for Official Publication of the European Communities, Luxemburg).
8. Howard MH, Gibson JL, Stolle D (2005) *The U.S. Citizenship, Involvement, Democracy Survey* (Center for Democracy and Civil Society, Georgetown University, Washington, DC).
9. Krosnick JA, Fabrigar LR (1997) *Survey Measurement and Process Quality*, eds Lyberg L, et al. (Wiley-Interscience, New York), pp 141–164.

10. Tropp LR, Bianchi RA (2006) Valuing diversity and interest in intergroup contact. *J Soc Issues* 62(3):533–551.

11. Adesokan AA, Ullrich J, van Dick R, Tropp LR (2011) Diversity beliefs as moderator of the contact-prejudice relationship. *Soc Psychol* 42(4):271–278.

Table S1. Unstandardized estimates (SE in brackets) for model 2 in study 2b

	Study 2b	
Model 2	β (SE)	*P*
Level 1		
$contact_{time1} \rightarrow contact_{time2}$	0.636 (0.039)	<0.001
$prejudice_{time1} \rightarrow prejudice_{time2}$	0.569 (0.031)	<0.001
$contact_{time1} \rightarrow prejudice_{time2}$	−0.007 (0.004)	0.089
$prejudice_{time1} \rightarrow contact_{time2}$	−0.239 (0.244)	0.327
$norms_{time1} \rightarrow norms_{time2}$	0.056 (0.035)	0.108
$contact_{time1} \rightarrow norms_{time\ 2}$	0.051 (0.006)	<0.001
$norms_{time1} \rightarrow prejudice_{time2}$	0.004 (007)	0.615
Level 2		
$contact_{time1} \rightarrow contact_{time2}$	0.989 (0.160)	<0.001
$prejudice_{time1} \rightarrow prejudice_{time2}$	0.467 (0.221)	0.034
$contact_{time1} \rightarrow prejudice_{time2}$	−0.043 (0.015)	0.006
$prejudice_{time1} \rightarrow contact_{time2}$	4.028 (2.492)	0.106
$norms_{time1} \rightarrow norms_{time2}$	0.391 (0.146)	0.007
$contact_{time1} \rightarrow norms_{time\ 2}$	0.084 (0.030)	0.005
$norms_{time1} \rightarrow prejudice_{time2}$	−0.146 (0.057)	0.011