



Social norm enforcement in ethnically diverse communities

Fabian Winter^a and Nan Zhang^{a,1}

^aMax-Planck Research Group “Mechanisms of Normative Change,” Max-Planck-Institute for Research on Collective Goods, 53113 Bonn, Germany

Edited by Karen S. Cook, Stanford University, Stanford, CA, and approved February 5, 2018 (received for review October 20, 2017)

Recent waves of immigration to Western nations have fueled a debate over the consequences of ethnic diversity for social cohesion. One prominent argument in this debate holds that diversity is detrimental to trust and cooperation because individuals in heterogeneous communities face difficulties in enforcing social norms across ethnic lines. We examine this proposition in a field experiment involving real-life interactions among residents of multiethnic German neighborhoods. We find significant ethnic asymmetries in the pattern of norm enforcement: Members of the majority “native” German population are more active in sanctioning norm violations, while ethnic minorities are more likely to find themselves the target of sanctions. We interpret these results in light of prevailing status inequalities between ethnic minorities and the native majority. We further calculate that, as a result of ethnic discrimination, social control is likely to rise in communities with moderate minority population shares.

ethnic diversity | social norms | field experiment | social control | enforcement

This paper investigates the influence of ethnicity on the enforcement of social norms in a modern European society. We engage with a wide-ranging debate across the social sciences about the negative effects of diversity on social cohesion (1–4). One prominent argument in this debate holds that individuals face difficulty in sanctioning antisocial behavior across ethnic lines, leading to lower levels of trust and cooperation in ethnically diverse communities (5–8). However, to date, evidence for this proposition is drawn mainly from the small-group context (6, 7) in which coethnicity overlaps with social networks that help individuals to identify defectors and communicate reputational information (5, 9, 10). By contrast, surprisingly little is known about how ethnicity influences social norm enforcement in anonymous encounters which characterize many interactions in complex Western societies (4, 11, 12).

This study presents evidence from a field experiment documenting ethnically driven patterns of social norm enforcement in natural encounters between strangers. Our study was conducted near metro entrances in ethnically diverse German neighborhoods. The intervention exposed passersby to a professional actor engaged in a widely understood deviant act (littering). We systematically vary the ethnicity and gender of our confederates and record the passerby’s reaction to the norm violation, focusing in particular on the use of verbal reprimands. This design allows us to examine two key determinants of the overall level of social control in diverse communities: (i) differences in enforcement propensity between ethnic minorities and members of the majority “native” population and (ii) differences between minorities and natives in the likelihood of being sanctioned for norm transgressions.

Results from 3,249 trials reveal significant ethnic differences in enforcement propensity: Holding the identity of confederates constant, native passersby react to transgressions more frequently than minority passersby. We also observe discrimination against minority confederates, who provoke reactions more frequently than native confederates, regardless of the identity of

the passerby. Finally, we find that natives sanction natives at the same rate that minorities sanction minorities. We discuss these results in light of research highlighting native–minority status inequalities in European societies (13–17) and their implications for social norm enforcement (18–21). In the last step, we combine these revealed enforcement patterns with information on local population shares to calculate the overall level and direction of social control in multiethnic communities.

This research advances our understanding of the ethnic dimensions of social norm enforcement in several respects. First, we consider the relationship between ethnicity and enforcement decisions in the absence of social network ties. While previous work has highlighted the role of ethnically bounded networks in facilitating peer punishment (5–7), many interactions in complex modern societies consist of anonymous, one-shot encounters (4, 11, 12). The present study therefore fills an important knowledge gap by examining how ethnicity conditions social norm enforcement in confrontations between strangers.

Second, we move beyond investigating the effects of diversity on social control at the aggregate community level (6, 8, 10, 22) and focus instead on direct observation of minority–majority interactions. In doing so, we are able to highlight ethnic asymmetries in norm enforcement that are obscured in more aggregated data. Simply put, enforcement patterns in a community composed of 70% natives and 30% minorities would look radically different from the dynamics we would expect in a community with 30% natives and 70% minorities, even though both areas would be rated as similarly diverse along commonly used indexes of ethnolinguistic fractionalization (ELF) (23–25). By studying individual majority–minority interactions, we are able

Significance

How does ethnic diversity influence the maintenance of social norms in complex, Western societies? We provide evidence from a natural field experiment examining reactions to norm violations in multiethnic German neighborhoods. We uncover asymmetric patterns of norm enforcement in interethnic encounters: “Native” Germans are more active in sanctioning norm violations, while ethnic minorities are more likely to be targeted for sanctions. We link these dynamics to prevailing status inequalities between minorities and natives in German society. We further show that, as a consequence of such asymmetries, social control tends to rise with ethnic diversity.

Author contributions: F.W. and N.Z. designed research, performed research, analyzed data, and wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

This open access article is distributed under [Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 \(CC BY-NC-ND\)](#).

Data deposition: Replication data and code reported in this paper have been deposited in the Open Science Framework (doi.org/10.17605/OSF.IO/6JMN2).

¹To whom correspondence should be addressed. Email: zhang@coll.mpg.de.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1718309115/-DCSupplemental.

Published online February 26, 2018.

to take account of the social positioning of ethnic groups (14) in shaping sanctioning decisions.

Finally, this paper sheds light on social norms enforcement in a real-life setting. A handful of studies have examined how sanctioning patterns are influenced by group identity (20, 26–30) and ethnic considerations (7), using experimental games. However, Balafoutas and Nikiforakis (31) and Balafoutas et al. (32) show that sanctioning behavior differs substantially between the laboratory and the field, raising questions of external validity (33). Our design allows us to avoid these concerns by examining individual decisions in their natural context, free from selection and experimenter demand effects (34, 35).

Setting and Research Design

We investigate reactions to violations of a widely held social norm against littering (12, 31, 32, 36). The study consists of three distinct parts: (i) a preexperimental questionnaire to establish whether our research targets differ in their understanding of littering norms, (ii) a field-experimental intervention whereby passersby are exposed to public littering, and (iii) a postexperimental questionnaire to gather detailed information on nativity, citizenship, and ethnic identification. The three parts of the study involved different samples of research subjects, but were all conducted in the same ethnically diverse neighborhoods in the cities of Bonn and Cologne, Germany (see *Materials and Methods* for more details).

Experimental Protocol. The experimental intervention was staged at entrances to various metro stations in our research areas. As shown in Fig. 1, confederates were stationed about 2 m away from a public trashcan while holding a disposable coffee cup. Confederates were instructed to choose as their “observer” the next unaccompanied pedestrian walking toward the experimental area. As soon as this observer approached, the confederate was instructed to carelessly throw the cup on the ground in the general vicinity of the trashcan, thereby demonstrating an open disregard for the antilittering norm.

A research assistant stood ~5 m away from the confederate and discreetly recorded the reaction of the observer using a smartphone app. Research assistants could freely record a range of reactions including direct verbal sanctions, dirty looks, mumbled complaints, or attempts to pick up the cup. However, since subtle expressions of disapproval such as glances or headshakes can be ambiguous and fleeting, our main variable of interest is whether the observer directly confronted the confederate by issuing a ver-

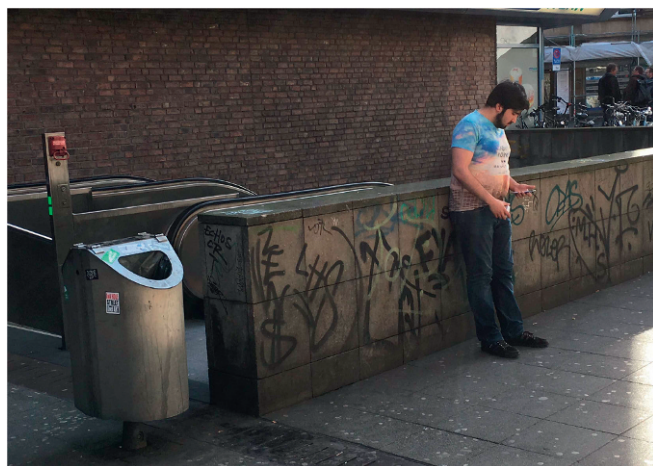


Fig. 1. The experimental setting. The actor holds a disposable coffee cup, which he carelessly throws in the vicinity of the trashcan upon the approach of a passerby.

bal reprimand. In *SI Appendix*, Fig. S2, we show that our main results are robust to coding a full range of negative reactions.

Research assistants were additionally instructed to collect information on the gender, approximate age, and height of the observer and to make their best guess as to whether the observer has a “migration background” (*migrationshintergrund*). This legal designation, which is distinct from citizenship or country of birth, effectively encompasses first- to third-generation immigrants to Germany. While it is admittedly difficult to visually distinguish Northern and Eastern Europeans from native Germans defined in this manner, “visible” ethnic minorities from Turkey, Southern Europe, Latin America, the Middle East, Africa, and Asia are largely covered by this category. *SI Appendix*, section 5 provides evidence from our postexperimental survey demonstrating that our coding of migration background closely matches information on the precise ancestry of individuals in our study locations.

Experimental Conditions. To play the role of confederates, we recruited actors who could also be clearly identified as either native German or of migration background (see *Materials and Methods* for further details. Actor profiles are provided in *SI Appendix*, Fig. S1). We thus study four types of encounters: (i) native observers and native actors, (ii) native observers and minority actors, (iii) minority observers and native actors, and (iv) minority observers and minority actors. Relative frequencies are displayed in Table 1. A total of 66.4% of observers for whom ethnicity could be confidently assessed were classified as native, and 62.3% of all interventions were conducted by native confederates. However, native observers were no more likely than minority observers to interact with a native confederate, suggesting that our confederates did not select passersby based on visible criteria. Additional balance checks are provided in *SI Appendix*, Table S2.

Results

A Widely Shared Norm Against Littering. We first check whether popular understanding of the antilittering norm is consistent across our four experimental conditions. Here, we draw upon data from our preexperimental survey in which respondents were presented with a visual depiction of a confederate throwing trash in a public space and asked about the extent to which they would be bothered by this situation (0 = “not at all,” 0.5 = “somewhat,” 1 = “very much”).

The overwhelming majority of survey respondents (84.6%) stated that they would be bothered very much by the act of littering. We find no significant differences in antilittering attitudes between native and minority respondents [$n = 227$, $\chi^2(2) = 3.32$, $P = 0.19$, χ^2 test]. Moreover, as shown in Fig. 24, the identity of the confederate does not affect judgments about the acceptability of littering. Finally, a majority (55.5%) of respondents stated that they would issue a verbal sanction if confronted with a similar situation. Again, we find no statistically significant differences with respect to the ethnicity of the respondent [$n = 227$, $\chi^2(1) = 0.08$, $P = 0.78$, χ^2 test].

While our survey results could be construed as “cheap talk,” they also indicate that most respondents (irrespective of ethnic considerations) both are aware of the antilittering norm and consider verbal sanctions to be a legitimate response to violations. These results thus provide assurance that behavioral differences in the field experiment capture variation in the willingness to sanction norm breakers, as opposed to different understandings of the norm itself.

Basic Description of Sanctioning Behavior. Although the majority of survey respondents indicated a willingness to verbally sanction norm violations, only 10.2% of experimental interventions actually resulted in a verbal reprimand. This number is comparable to results from previous field experiments using similar

Table 1. Type and frequency of encounters

Encounter	No. of trials	%
Native observer, native confederate	1,348	41.49
Native observer, minority confederate	810	24.93
Minority observer, native confederate	676	20.81
Minority observer, minority confederate	415	12.77
Total	3,249	100

interventions (12, 31, 36). We also observe a gender difference in sanctioning behavior: Whereas 12.3% of violations observed by women were sanctioned, the enforcement rate falls to 8.1% in interventions observed by men [$n = 3,249$, $\chi^2(1) = 16.43$, $P < 0.001$, χ^2 test]. Finally, we note a higher tendency to react to norm violations in trials involving older passersby [$n = 3,249$, $\chi^2(5) = 167.91$, $P < 0.001$, χ^2 test].

Native Observers React to Violations More Frequently Than Minority Observers. Averaging across confederates, native observers reacted to 12.6% of interventions, compared with 5.3% for minority observers [$n = 3,249$, $\chi^2(1) = 42.18$, $P < 0.001$, χ^2 test]. Fig. 2B shows that this difference in baseline enforcement propensities holds across both native and minority confederates. When confronted with native confederates, native observers sanction 8.7% of the time, compared with 3.9% for minority observers [$n = 2,024$, $\chi^2(1) = 16.02$, $P < 0.001$, χ^2 test]. An even larger gap obtains when we consider reactions to minority confederates. Here, native observers sanction 19.1% of the time, compared with 7.7% for minority observers [$n = 1,225$, $\chi^2(1) = 27.69$, $P < 0.001$, χ^2 test]. *SI Appendix, Table S3* provides corroborating results from random-effects logistic regression models, which take account of the crossed matching in our design between confederates and research assistants. *SI Appendix, section 5* also presents evidence from Monte Carlo simulations showing the robustness of our findings to potential ethnic misclassification of observers.

Minority Confederates Elicit More Frequent Reactions Than Native Confederates. Fig. 2B also shows that observers apply conditional enforcement strategies depending upon the ethnicity of confederates. Averaging across all observers, 7.1% of trials involving native confederates resulted in a verbal reprimand, compared with 15.3% of trials involving minority confederates [$n = 3,249$, $\chi^2(1) = 56.23$, $P < 0.001$, χ^2 test]. Again, this difference remains when we focus separately on observer subgroups. Native observers sanction minority confederates significantly more often than native confederates [19.1% vs. 8.7%, $n = 2,158$, $\chi^2(1) = 50.22$, $P < 0.001$, χ^2 test], and the same holds true for minority observers [7.7% vs. 3.9%, $n = 1,091$, $\chi^2(1) = 7.63$, $P = 0.006$, χ^2 test]. Further statistical and robustness tests are provided in *SI Appendix, section 5* and *Table S3*.

Similar Enforcement Rates in Native–Native and Minority–Minority Encounters in Combination with Asymmetric Sanctioning Across Ethnic Lines. We observe that the rate at which natives sanction natives is statistically indistinguishable from the propensity of minorities to sanction minorities [$n = 1,763$, $\chi^2(1) = 0.38$, $P = 0.54$, χ^2 test]. On the other hand, natives sanction minorities at dramatically higher rates than minorities sanction natives [19.1% vs. 3.9%, $n = 1,486$, $\chi^2(1) = 80.53$, $P < 0.001$, χ^2 test]. These results have important implications for our explanation of enforcement patterns discussed below.

Discussion

Explaining Ethnic Variation in Enforcement Patterns. Prevailing research on norm enforcement in diverse communities has argued that sanctioning is more likely to occur within, rather than

across, ethnic boundaries (5–7). This result is linked to the sanctioning advantages conveyed by ethnic social networks that help individuals to identify defectors and communicate reputational information. We note, however, that these benefits accruing to social networks are absent in the stranger-to-stranger encounters we study. Thus, our experimental design ensures that informational and reputational considerations cannot account for the enforcement patterns we document.

Instead, our results are consistent with an alternative framework focusing on the role of status inequalities in sanctioning decisions (20). Previous studies show that high status is positively associated with more assertive behavior (18, 37–39) and that lower-status targets are more likely to elicit social sanctions (19, 21). In many European societies including Germany, immigrants and visible ethnic minorities tend to occupy the lower rungs of the status hierarchy (16, 17, 40–42). Additionally, status differences based on ethnicity are often overlaid upon minority disadvantages in language proficiency, educational attainment,

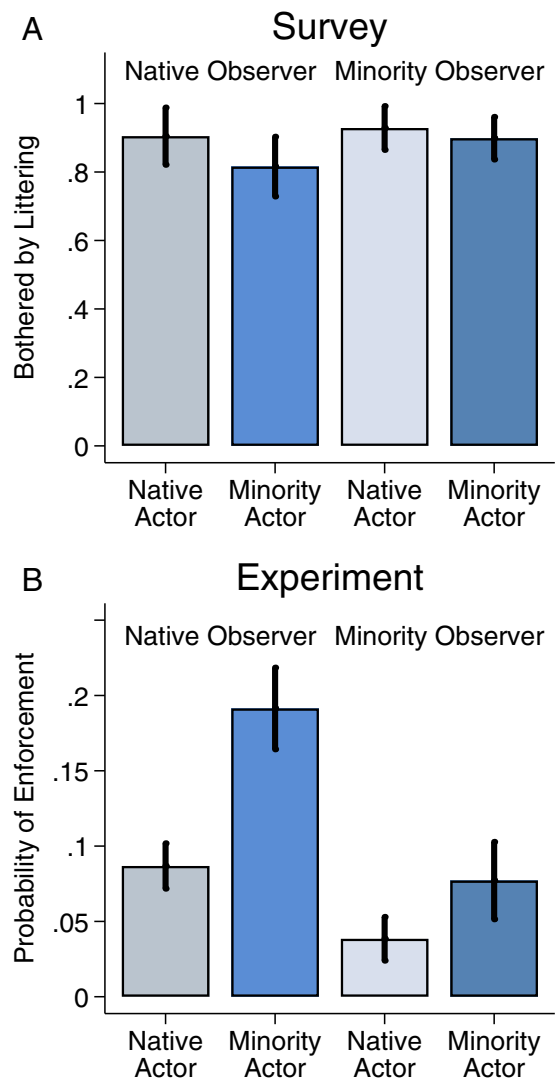


Fig. 2. Popular understanding of the antilittering norm and observed reactions to norm violations. (A) The average expressed disapproval toward an act of littering on a scale from 0 to 1 ($N = 227$), separated by the ethnicity (native, minority) of the observer and the violator. (B) The observed sanctioning rate in native–native, native–minority, minority–native, and minority–minority encounters ($N = 3,249$). Error bars represent 95% confidence intervals.

and labor market participation (13, 15, 43, 44), which we take as further evidence of marginalization. These dynamics imply that natives should be more vigilant when confronting minority violators, while minorities should behave more leniently toward native transgressors. By contrast, in native–native and minority–minority interactions where status differentials are equalized, we should expect similar enforcement levels to prevail. The enforcement patterns we observe in the field are fully consistent with these propositions.

Implications for Social Control at the Community Level. We can use these observed enforcement patterns to calculate the aggregate level of social control—defined as the probability that an “average” resident is sanctioned for norm violations—for communities with varying minority population shares. To do so, we weight the likelihood of all possible interactions (native observes native, native observes minority, and so on) by the associated enforcement rate observed in the field. More formally, we define $p_{ij} \in [0, 1]$ as the “interaction probability” that a member of group $i \in \{Native, Minority\}$ observes a norm violation committed by a member of group $j \in \{Native, Minority\}$. If we assume that individuals commit norm violations at some constant rate which is independent of ethnicity, then the interaction probabilities can be calculated from the respective proportions of natives and minorities within any community. We define $e_{ij} \in [0, 1]$ as the conditional probability that a member of i enforces a norm against a member of j . The aggregate level of social control P can therefore be calculated as

$$P = \sum_i^n \sum_j^m p_{ij} e_{ij}. \quad [1]$$

Fig. 3 plots the predicted level of social control for the population shares and sanctioning rates observed in our experiment. Across all 3,249 trials in which ethnicity was confidently assessed, we coded 33.6% of passersby as having a migration background. This share is indicated by the vertical dashed line in Fig. 3. Given our observed enforcement rates, we predict that an average transgressor will be sanctioned 9.76% of the time (denoted by the red diamond).

Fig. 3 is also useful for illustrating the comparative statics of Eq. 1: Holding enforcement rates constant, we can study the effect of shifts in the relative size of the native and minority populations on aggregate social control. Compared with a homogeneous native community, we predict that social control increases in the presence of minorities (due to their higher likelihood of being sanctioned). Importantly, however, this tightening of social control comes at the cost of a greater ethnic divide between enforcers and their targets. This dynamic is illustrated by the solid surfaces in Fig. 3 which decompose aggregate social control into its constituent components. At the highest levels of social control, the majority of sanctions are imposed by natives against ethnic minorities. The model thus highlights an uneasy trade-off between discrimination and sanctioning capacity in diverse communities.

Of course, the model we present here paints a simplified picture of the relationship between ethnic diversity and norm enforcement at the community level. A more dynamic version might model p_{ij} as a function of population shares or relax the assumption of random interaction probabilities. For example, natives living in diverse areas may mix more readily across ethnic lines and exhibit less discrimination compared with natives in areas with few minorities (45). For present purposes, the static model suffices to illustrate the importance of asymmetric sanctioning patterns in shaping the overall enforcement rate in multiethnic communities.

Composition of Enforcement by Proportion of Minority

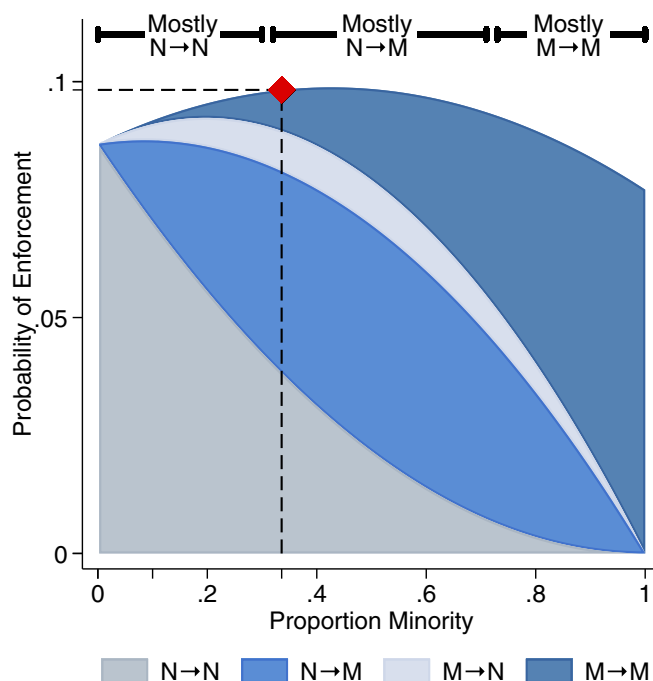


Fig. 3. Aggregate social control as a function of the share of ethnic minorities in the population. The shape of the curve is calculated from sanctioning patterns we observe in the field. The red diamond indicates the predicted enforcement rate for the observed minority population share in our research locations. The solid colors decompose the enforcement rate by the ethnicity of the observer and the norm violator. For instance, N→M denotes the component of aggregate social control that is attributable to native observers sanctioning minority transgressors. In the trivial case of a population that is composed of only natives, 100% of observed enforcement is due to natives sanctioning natives. In the case of a population composed of 50% natives and 50% minorities, the greatest share of enforcement is due to natives sanctioning minorities.

Conclusion

This study used a natural field experiment to examine the association between ethnicity and social norm enforcement in multiethnic German neighborhoods. We find evidence of asymmetric enforcement patterns: Natives are more active in sanctioning norm violations, while ethnic minorities are more likely to find themselves the target of sanctions. We link these dynamics to status inequalities between minorities and the majority native population. We also show that, as the result of such asymmetries, social control tends to increase in both frequency and bias in communities with a moderate-size ethnic minority population.

One important implication of this study is to highlight the distinction between two conceptions of ethnic diversity (4, 24). One approach (embodied in ELF measures) is to interpret diversity as the probability that two randomly drawn individuals from a population will belong to the same ethnic group (23). An alternative approach (often implicit in discussions of diversity) is to define diversity in terms of the size of the nonnative population. The choice of definitions is important, as it pushes scholars to focus on different mechanisms linking diversity to outcomes of interest. The first interpretation draws attention to forms of social solidarity rooted in cultural and ethnic similarity (1), while encouraging us to examine relations between groups in a “color-” and “status-blind” way. In contrast, the second approach

highlights the historical and institutionalized inequalities separating natives and ethnic minorities and their associated consequences for social cohesion (4, 46).

While many prominent studies of ethnic diversity have primarily taken a color-blind approach (1–3, 5, 47, 48), our work shows that much can be gained by focusing on the hierarchies and inequalities inherent in ethnic relations. The conjecture that social control increases with diversity is not obvious and even runs counter to evidence suggested by previous studies (8, 22). The reasoning behind this result becomes apparent only once we take account of asymmetric sanctioning patterns in native–minority interactions. Of course, it remains an open question whether discriminatory reactions to norm violations actually bring about greater norm compliance. Nonetheless, the present study has shed more light on the dynamics of norm enforcement in multiethnic Western societies.

Materials and Methods

IRB approval for the study was granted by the Ethics Commission of the University of Bonn. As is standard in natural field experiments of this type, we obtained a waiver of informed consent.

Setting. The field experiment and surveys were conducted in the vicinity of the Bad Godesberg train station in Bonn, between the Kalk-Post and Kalk-Kapelle U-Bahn stations in Cologne, and also in the vicinity of the Venloer Strasse U-Bahn station in the Ehrenfeld section of Cologne. The selection of locations ensured that the presence of visible minorities (i.e., our confederates) would not arouse curiosity or suspicion and that we would be able to record the reaction of an adequate number of minority passersby.

Information on the demographic composition of our local research areas comes from our postexperimental survey. A total of 39.5% of respondents for whom we have complete data are coded as having migration background, defined as either (i) being born outside of Germany or (ii) having at least one parent born outside of Germany. In addition to being diverse in terms of the size of the minority population, our research areas are also diverse in terms of the number of ethnic groups represented. First- and second-generation Turks (the largest minority ethnic group in Germany) constitute only 24.6% of “nonnative” respondents in our sample. A further 27.0% of respondents have roots in countries in the Middle East and North Africa. The remaining 48.4% of minority respondents claim ancestry from over 25 different countries.

Field Experiment. Interventions were staged Monday through Saturday from 3:00 PM to 6:00 PM over the period March 21 to April 25, 2017. Teams of two (a confederate and a research assistant) simulated the act of littering and recorded the reactions of passersby according to the procedure defined

in the main text. In all, seven confederates and six research assistants took part in the study. In forming teams each day, we cross-matched confederates with research assistants. All research assistants and confederates were blind to the purposes of the study.

We recruited four native (two male and two female) and three minority (one male and two female) actors for the study. All confederates were young people between the ages of 18 y and 25 y. Actors were instructed to dress “naturally” such that passersby would feel comfortable approaching them (i.e., clothing could vary depending on the actor and day of the experiment, but actors were always dressed in a way to not arouse suspicion or scare people off).

The intervention was performed in front of single pedestrians only. Confederates were instructed to not conduct the intervention if individuals were on bikes, wearing earphones, talking on the phone, accompanied by children, or walking in groups. In addition, the research assistant recorded only those interventions in which the passerby clearly took note of the act of littering by glancing at the thrown cup. Once the observer and all other passersby who had potentially witnessed the norm violation had cleared the area, the cup was retrieved and the next trial was begun.

Surveys. Two surveys were conducted in the same experimental locations as the field experiment. A preexperimental survey was conducted between March 13 and March 22, 2017, and a postexperimental survey was conducted between July 12 and August 18, 2017. Survey enumerators worked in teams of two. To increase the likelihood of sampling nonnative German speakers, each survey team included at least one bilingual German/Turkish or German/Arabic speaker.

Potential survey respondents were solicited using a randomization procedure. Specifically, our survey app instructed the interviewers to let a random number of people pass before approaching the next potential respondent. Complete questionnaire data from 227 individuals were collected in the preexperimental survey and from 319 individuals in the postexperimental survey.

Data Availability. Replication data and code can be accessed at doi.org/10.17605/OSF.IO/6JMN2.

ACKNOWLEDGMENTS. We thank Kamal Abbadi, Aya Alahrachi, Mustapha Berdahoum, Ayse Celik, Carina Derikartz, Betül Eroglu, Dietrich Napiontek, Viktoria Oellers, Ogulcan Ozturk, Lea Stentenbach, Christian Vollmer, Marie von Oepen, and particularly Tobias Werner for excellent research assistance. Confederate roles were played by Loxy Diercks, Florentin Eikelmann, Annalena Kayser, Gina Makurat, Zühre Nur, Stefanos Triantafyllou, and Reza Vaghef. Valuable feedback was provided by Maria Abascal, Amelie Aidenberger, Delia Baldassarri, Anna Baumert, Diego Gambetta, Johanna Gereke, David Hugh-Jones, Merlin Schaeffer, Max Schaub, and conference participants at Bocconi University, Universitat de Barcelona, and the Max-Planck-Institute for Research on Collective Goods. Financial support of the Max-Planck-Society for the Max-Planck Research Group “Mechanisms of Normative Change” is gratefully acknowledged.

- Putnam R (2007) E pluribus unum: Diversity and community in the twenty-first century. *Scand Polit Stud* 30:137–174.
- Alesina A, La Ferrara E (2000) Participation in heterogeneous communities. *Q J Econ* 115:847–904.
- Alesina A, La Ferrara E (2002) Who trusts others? *J Public Econ* 85:207–234.
- Abascal M, Baldassarri D (2015) Love thy neighbor? Ethnoracial diversity and trust reexamined. *Am J Sociol* 121:722–782.
- Fearon J, Laitin D (1996) Explaining interethnic cooperation. *Am Polit Sci Rev* 90:715–735.
- Miguel E, Gugerty M (2005) Ethnic diversity, social sanctions, and public goods in Kenya. *J Public Econ* 89:2325–2368.
- Habyarimana J, Humphreys M, Posner D, Weinstein J (2007) Why does ethnic diversity undermine public goods provision? *Am Polit Sci Rev* 101:709–725.
- Algan Y, Hémet C, Laitin D (2016) The social effects of ethnic diversity at the local level: A natural experiment with exogenous residential allocation. *J Polit Econ* 124:696–733.
- Koopmans R, Schaeffer M (2015) Relational diversity and neighbourhood cohesion. Unpacking variety, balance and in-group size. *Soc Sci Res* 53:162–176.
- Legewie J, Schaeffer M (2016) Contested boundaries: Explaining where ethnoracial diversity provokes neighborhood conflict. *Am J Sociol* 122:125–161.
- Portes A, Vickstrom E (2011) Diversity, social capital, and cohesion. *Annu Rev Sociol* 37:461–479.
- Balafoutas L, Nikiforakis N, Rockenbach B (2014) Direct and indirect punishment among strangers in the field. *Proc Natl Acad Sci USA* 111:15924–15927.
- Heath A, Rothson C, Kilpi E (2008) The second generation in Western Europe: Education, unemployment, and occupational attainment. *Annu Rev Sociol* 34:211–235.
- Abascal M (2015) Us and them: Black-White relations in the wake of Hispanic population growth. *Am Sociol Rev* 80:789–813.
- Adida C, Laitin D, Valfort MA (2016) *Why Muslim Integration Fails in Christian Heritage Societies* (Harvard Univ Press, Cambridge, MA).
- Potărcă G, Mills M (2014) Racial preferences in online dating across European countries. *Eur Sociol Rev* 31:326–341.
- Hagendoorn L (1995) Intergroup biases in multiple group systems: The perception of ethnic hierarchies. *Eur Rev Soc Psychol* 6:199–224.
- Diekmann A, Jungbauer-Gans M, Krassnig H, Lorenz S (1996) Social status and aggression: A field study analyzed by survival analysis. *J Soc Psychol* 136:761–768.
- Doob A, Gross A (1968) Status of frustrator as an inhibitor of horn-honking responses. *J Soc Psychol* 76:213–218.
- Hoff K, Kshetramade M, Fehr E (2011) Caste and punishment: The legacy of caste culture in norm enforcement. *Econ J* 121:F449–F475.
- Wolbring T, Bozoyan C, Langner D (2013) “Links gehen, rechts stehen!” Ein Feldexperiment zur durchsetzung informeller normen auf rolltreppen [“Walk left, stand right!” A field experiment on the enforcement of informal norms on escalators]. *Z Soziol* 42:239–258. German.
- Sampson R, Raudenbush S, Earls F (1997) Neighborhoods and violent crime: A multi-level study of collective efficacy. *Science* 277:918–924.
- Fearon J (2003) Ethnic and cultural diversity by country. *J Econ Growth* 8:195–222.
- van der Meer T, Tolsma J (2014) Ethnic diversity and its effects on social cohesion. *Annu Rev Sociol* 40:459–478.
- Posner D (2004) Measuring ethnic fractionalization in Africa. *Am J Pol Sci* 48:849–863.
- Chen Y, Li S (2009) Group identity and social preferences. *Am Econ Rev* 99:431–457.

27. Goette L, Huffman D, Meier S (2006) The impact of group membership on cooperation and norm enforcement: Evidence using random assignment to real social groups. *Am Econ Rev* 96:212–216.
28. Goette L, Huffman D, Meier S, Sutter M (2012) Competition between organizational groups: Its impact on altruistic and antisocial motivations. *Manage Sci* 58:948–960.
29. Bernhard H, Fischbacher U, Fehr E (2006) Parochial altruism in humans. *Nature* 442:912–915.
30. Bernhard H, Fehr E, Fischbacher U (2006) Group affiliation and altruistic norm enforcement. *Am Econ Rev* 96:217–221.
31. Balafoutas L, Nikiforakis N (2012) Norm enforcement in the city: A natural field experiment. *Eur Econ Rev* 56:1773–1785.
32. Balafoutas L, Nikiforakis N, Rockenbach B (2016) Altruistic punishment does not increase with the severity of norm violations in the field. *Nat Commun* 7:13327.
33. Guala F (2012) Reciprocity: Weak or strong? What punishment experiments do (and do not) demonstrate. *Behav Brain Sci* 35:1–15.
34. Harrison G, List J (2004) Field experiments. *J Econ Lit* 42:1009–1055.
35. Levitt S, List J (2007) What do laboratory experiments measuring social preferences reveal about the real world? *J Econ Perspect* 21:153–174.
36. Berger J, Hevenstone D (2016) Norm enforcement in the city revisited: An international field experiment of altruistic punishment, norm maintenance, and broken windows. *Rationality Soc* 28:299–319.
37. Piff P, Stancato D, Côté S, Mendoza-Denton R, Keltner D (2012) Higher social class predicts increased unethical behavior. *Proc Natl Acad Sci USA* 109:4086–4091.
38. Keltner D (2017) *The Power Paradox: How We Gain and Lose Influence* (Penguin, New York).
39. Ridgeway C (2014) Why status matters for inequality. *Am Sociol Rev* 79:1–16.
40. Lee T, Fiske S (2006) Not an outgroup, not yet an ingroup: Immigrants in the stereotype content model. *Int J Intercult Relat* 30:751–768.
41. Snellman A, Ekehammar B (2005) Ethnic hierarchies, ethnic prejudice, and social dominance orientation. *J Community Appl Soc Psychol* 15:83–94.
42. Steinbach A (2004) *Soziale Distanz: Ethnische Grenzziehung und die Eingliederung von Zuwanderern in Deutschland* (VS Verlag für Sozialwissenschaften, Wiesbaden, Germany).
43. Worbs S (2003) The second generation in Germany: Between school and labor market. *Int Migr Rev* 37:1011–1038.
44. Kristen C, Granato N (2007) The educational attainment of the second generation in Germany: Social origins and ethnic inequality. *Ethnicities* 7:343–366.
45. Schaeffer M (2012) Which groups are mostly responsible for problems in your neighbourhood? The use of ethnic categories in Germany. *Ethn Racial Stud* 36:156–178.
46. Baldwin K, Huber J (2010) Economic versus cultural differences: Forms of ethnic diversity and public goods provision. *Am Polit Sci Rev* 104:644–662.
47. Delhey J, Newton K (2005) Predicting cross-national levels of social trust: Global pattern or Nordic exceptionalism? *Eur Sociol Rev* 21:311–327.
48. Alesina A, Baqir R, Easterly W (1999) Public goods and ethnic divisions. *Q J Econ* 114:1243–1284.