Routledge
Taylor & Francis Group

Journal of Ethnic and Migration Studies

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/cjms20

How to reduce discrimination? Evidence from a field experiment in amateur soccer

Robert Dur, Carlos Gomez-Gonzalez & Cornel Nesseler

To cite this article: Robert Dur, Carlos Gomez-Gonzalez & Cornel Nesseler (2022): How to reduce discrimination? Evidence from a field experiment in amateur soccer, Journal of Ethnic and Migration Studies, DOI: 10.1080/1369183X.2022.2071688

To link to this article: https://doi.org/10.1080/1369183X.2022.2071688

JOURNAL OF ETHNIC AND MIGRATION STUDIES
https://doi.org/10.1080/1369183X.2022.2071688

Routledge
Taylor & Francis Group

How to reduce discrimination? Evidence from a field experiment in amateur soccer

Robert Dur [a], Carlos Gomez-Gonzalez [b] and Cornel Nesseler [c]

[a]Department of Economics, Erasmus University Rotterdam, Rotterdam, The Netherlands; [b]Department of Business Administration, University of Zurich, Zurich, Switzerland; [c]NTNU Business School, Norwegian University of Science and Technology, Trondheim, Norway

ABSTRACT

A rich literature shows that ethnic discrimination is an omnipresent and highly persistent phenomenon. Little is known, however, about how to reduce discrimination. This study reports the results of a large-scale field experiment we ran together with the Norwegian Football Federation. The federation sent an email to a random selection of about 500 amateur soccer coaches, pointing towards the important role that soccer can play in promoting inclusivity and reducing racism in society and calling on the coaches to be open to all interested applicants. Two weeks later, we sent fictitious applications to join an amateur club, using either a native-sounding or a foreign-sounding name, to the same coaches and to a random selection of about 500 coaches who form the control group. In line with earlier research, we find that applications from people with a native-sounding name receive significantly more positive responses than applications from people with a foreign-sounding name. Surprisingly and unintentionally, the email from the federation substantially increased rather than decreased this gap. Our study underlines the importance of running field experiments to check whether well-intended initiatives are effective in reducing discrimination.

ARTICLE HISTORY
Received 27 January 2022
Accepted 25 April 2022

KEYWORDS
Ethnic discrimination; intervention; field experiment; correspondence test; amateur soccer

1. Introduction

Ethnic discrimination is a global and persistent phenomenon. A rich body of experimental research in a variety of contexts demonstrates that the ethnic background of a person matters a great deal for the opportunities one gets in society. In the labour market, for example, numerous correspondence tests have shown that ethnic minorities are less likely to receive callbacks for interviews when applying for jobs (Bertrand and Duflo 2017; Bertrand and Mullainathan 2004; Kaas and Manger 2012; Lancee 2021; Oreopoulos 2011; Pager, Bonikowski, and Western 2009; Quillian and Midtbøen 2021; Riach and Rich 2002; Thijssen et al. 2021; Weichselbaumer 2020; Zschirnt and Ruedin 2016). Research has also found that ethnic minorities face severe discrimination in housing

CONTACT Carlos Gomez-Gonzalez carlos.gomez@business.uzh.ch
Supplemental data for this article can be accessed https://doi.org/10.1080/1369183X.2022.2071688.

(Auspurg, Schneck, and Hinz 2019; Diehl et al. 2013; Sawert 2020), shopping (Bourabain and Verhaeghe 2019), transportation (Liebe and Beyer 2021; Mujcic and Frijters 2021), the sharing economy (Edelman, Luca, and Svirsky 2017; Ge et al. 2020), online dating (Jakobsson and Lindholm 2014), and sports (Gomez-Gonzalez, Nesseler, and Dietl 2021; Nesseler, Gomez-Gonzalez, and Dietl 2019). Meta-analyses by Quillian et al. (2017) and Heath and Di Stasio (2019) show that the extent to which ethnic minorities are discriminated against has hardly changed during the last three decades.

One reason why ethnic discrimination is so persistent over time is that we still know little about which policies reduce discrimination. In a recent review of more than 300 studies, Paluck et al. (2021) conclude that much research is 'ill-suited to provide actionable, evidence-based recommendations for reducing prejudice' (p.533). Similarly, Bertrand and Duflo (2017) write that 'While field experiments in the last decade have been instrumental in documenting the prevalence of discrimination, field experiments in the future decade should aim to play as large of a role in isolating effective methods to combat it' (p.383).

This paper contributes to this by analyzing the effectiveness of a low-cost intervention to reduce discrimination. The context we study is amateur soccer. Discrimination is prevalent in amateur soccer. In a correspondence test in 22 European countries, Gomez-Gonzalez, Nesseler, and Dietl (2021) show that when asking to join a training, people with foreign-sounding names are significantly less likely to receive a positive response from clubs than people with native-sounding names. In our study, we measure discrimination in the same way, i.e. through a correspondence test. The correspondence test is preceded by an anti-discrimination intervention, run in collaboration with the Norwegian Football Federation (NFF), and implemented for a random selection of amateur soccer clubs. The experimental design allows us to estimate the causal effect of the intervention on discrimination in the field.

The intervention is an email sent by the NFF encouraging soccer coaches to be open to people interested in membership independent of the ethnic background of the applicant. The email describes the important role that soccer can have in bringing people of diverse backgrounds together. It argues that, in this way, soccer can promote interaction and can be key to social inclusion (cf. Lowe 2021; Mousa 2020). It also mentions that the NFF finds it important that soccer is multicultural and diverse, reflecting diversity in society. In addition to encouragement, the email gives some information about current discrimination in amateur soccer, mentioning that studies have shown that 'players with foreign-sounding names are less likely to get a response when contacting a club for the first time'. Providing such information about the prevalence of discrimination can be essential – as shown by Boring and Philippe (2021) – because people are not always aware that they discriminate (Bertrand, Chugh, and Mullainathan 2005; Rooth 2010). The NFF sent the email two weeks preceding the correspondence test.

Our main predictions – which we pre-registered (https://doi.org/10.1257/rct.8049-1.0) – were that the email from the NFF would lead to an increase in positive responses to email requests to join a training session from people with a foreign-sounding name, whereas we expected no effects for email requests to join a training session from people with a native-sounding name. We defined a positive response as either an invitation to come to a training or a conditional acceptance (e.g. yes, you're welcome, but only if you are a defender). This was also pre-registered.

Our results show that coaches are less responsive to applications from people with a foreign-sounding name compared to identical applications from people with a native-sounding name. The gap in the positive response rate for the full sample is 11 percentage points, which is almost the same as the gap found two years ago in amateur soccer in Norway, reported in Gomez-Gonzalez, Nesseler, and Dietl (2021).

Concerning the effect of the intervention, we find a surprising result. Instead of reducing the gap in positive response rates, the gap actually increases as a result of the intervention, implying more discrimination. Underlying this effect is a strong increase in coaches' positive responses to applications from people with a native-sounding name. Positive responses to applications from people with a foreign-sounding name also go up, but only a little.

One possible explanation for our unexpected finding is that the intervention gave rise to feelings of resentment among some coaches, leading them to be more open towards people with a native-sounding name and less open to people with a foreign-sounding name. Such defiant behaviour in response to moral appeals by authorities has been found in earlier studies in other contexts, including tax compliance (Ariel 2012; Blumenthal, Christian, and Slemrod 2001), vaccination (Nyhan and Reifler 2015; Nyhan et al. 2014), and criminal offending (Bouffard and Piquero 2010). Other coaches, however, may have responded in line with our predictions: they were more open to applications with foreign-sounding names and equally open to applications with native-sounding names. On net, we may then observe an increase in the response to applications with native-sounding names and no, or a very small response, to applications with a foreign-sounding name.

We do some further, more exploratory analysis of our data. Inspired by earlier studies that find more severe discrimination in less populous regions (Huijsmans et al. 2021; Mayda 2006), we split our data by the population size of the region a club resides in. We show that, in the absence of the intervention, discrimination only occurs in less populous regions. Discrimination is severe there, amounting to a gap in positive response rates of more than 20 percentage points. Strikingly, the treatment effect of our intervention in these regions is close to zero. Clubs located in the most populous regions show no discriminatory responses in the absence of the intervention, but respond strongly to the intervention in the form of an increase in positive responses to applications with native-sounding names. Hence, if feelings of resentment drive the treatment effects, it should be the non-discriminating coaches who have been affected by this, which seems not very plausible.

A possible alternative interpretation could be that non-discriminating coaches were not aware that discrimination occurs in amateur soccer and that the information contained in the email about the presence of discrimination induced them to conform to the descriptive norm to discriminate. That people tend to conform their behaviour to others has been found in many earlier studies in a variety of contexts (Allcott and Rogers 2014; Bott et al. 2020; Bradler et al. 2016; Chen et al. 2010; Frey and Meier 2004; Gerber and Rogers 2009; Hallsworth et al. 2016; Schultz et al. 2007). Interventions that make harmful behaviour of others visible can then backfire, as it makes the harmful behaviour seem natural and socially acceptable (Bicchieri and Dimant 2019; Cialdini, Reno, and Kallgren 1990; Dur and Vollaard 2015; Kahan 1997; Keizer, Lindenberg, and Steg 2008). The descriptive norm theory can, however, not explain that the increased

discrimination mainly arose from a strong increase in coaches' positive responses to applications from people with native-sounding names.

Our third and last interpretation – which we owe to one of the reviewers – relates to the part of the email that reminds coaches that soccer can promote inclusivity and inter-ethnic interaction. Informing (or reminding) coaches about this may have encouraged them to invite more natives to expose them to this positive contact. This is also consistent with the finding that the increase in positive responses to applications with native-sounding names occurs amongst (counterfactually) non-discriminating coaches. Thus, our surprising result might be an unintentional consequence of a well-meaning gesture amongst non-discriminatory coaches.

Our paper is inspired by and contributes to a small but growing literature testing interventions to reduce discrimination using field experiments. The study that is closest to ours is Boring and Philippe (2021). They study the effects of a similar intervention – an email that includes information about existing discrimination and an appeal not to discriminate – but in a different context, namely student evaluations in higher education. Their results show that the email is effective in reducing gender bias. Likewise, Alesina et al. (2018) find that informing teachers about their implicit bias against immigrants reduces their bias in grading exams of immigrant students.

The rest of the paper is structured as follows. Section 2 describes the experimental design and the data. Section 3 presents the results and provides an interpretation. In Section 4, we give a brief summary and offer some concluding remarks.

2. Experimental setup and data

Our field experiment took place in Norway in the Fall of 2021. Norway has a substantial share of first- and second-generation immigrants. In 2021, 18.5% of inhabitants were immigrants or Norwegian-born to immigrant parents.[1] The largest groups come from Poland, Lithuania, and Somalia. They are spread out across the country, with some over-representation in big cities such as Oslo, Stavanger, and Bergen.

It is widely acknowledged that ethnic discrimination is present in Norway. For example, Midtbøen (2016) and Larsen and Di Stasio (2021) performed correspondence tests in the Norwegian labour market. They find that applicants with a Norwegian-sounding name were significantly more likely to receive a callback than applicants with a foreign-sounding name. Andersson, Jakobsson, and Kotsadam (2012) find similar results for the Norwegian housing market. In the context of amateur soccer, Gomez-Gonzalez, Nesseler, and Dietl (2021) find an 11 percentage points higher response rate to applications from native-sounding names.

The intervention we test was created together with and implemented by the NFF. The NFF's purposes and activities include organising and promoting soccer in Norway. It provides the framework for regional federations to organise and administer their leagues. Our experiment targets amateur soccer clubs in the lowest soccer divisions. While the clubs participating in these leagues are obligated to comply with the regulations and directions from the NFF, they have autonomy about the level of the membership fees and whom to admit as a member.

The NFF provided us with a dataset of the universe of the NFF's amateur adult soccer clubs in Norway. The dataset contains the email of the coach or the club,[2] postal code of

the club, league in which the club plays, and whether it is a male or female club. If a club had more than one adult team, we randomly selected one team to avoid contamination.

Together with the NFF we created an email message that aims to make coaches more open to admit people with a migration or non-native background to their club. The email points to the important role soccer can play in promoting inclusivity and reducing racism in society. The email also describes findings from studies showing that discrimination is present in amateur soccer. The email calls on the coaches to help keep soccer open to all people that are interested.

The exact text of the email was as follows (translated from Norwegian; the original Norwegian text is included in Appendix A):

Subject: *Soccer for everyone – for inclusion and against racism.*
Dear coach,
Soccer is the world's most popular sport and it gives us responsibility and the opportunity to unite people from different backgrounds. It is important for all of us that soccer is multicultural and diverse and reflects our entire society – not just for professionals, but for all players. At the amateur level, soccer facilitates integration and promotes interaction on and off the field. We aim to make soccer easily accessible to all members of our society.
Racism and exclusion are a societal problem and, thus, also a problem of soccer.
Scientific studies showed that it is more difficult for people with foreign names to join an amateur soccer club. Regardless of sport or language skills, players with foreign names are less likely to get a response when contacting a club for the first time.
This barrier to participating in sports is not only negative for foreigners, but also for Norwegians with a migration background and the Norwegian society. Creating diverse teams with members from different backgrounds is the key to improving social inclusion.
We have sent this email to all coaches in Norway to applaud you on the work you do and encourage everyone to continue to help keep the door to soccer open to all interested.
2021 Will be a strange year, but we hope that at least autumn will be almost normal for most people and that together we can look forward to a glorious 2022.
[Name of NFF's representative]
[Logo of NFF]

The email was sent from the email account of an NFF representative on 15 October 2021 at 1:41 pm.

To be able to estimate the effect of the email on discrimination, it was sent to a randomly selected half of all clubs in Norway.[3] The remaining clubs in Norway form the control group. We randomly assigned clubs to the treatment group and the control group using a block design. To achieve balance between treatment and control regarding region and type of club, we created two blocks for each of the federation's 18 regions, one with all male clubs in the region and one with all female clubs in the region. We randomised half of the clubs within each of the 36 blocks either to treatment or to control. Thus, 466 clubs were assigned to treatment and so were sent the email from the NFF, whereas 501 were assigned to control. The slight difference in the number of clubs assigned to treatment and control results from the block design. Figure 1 shows how clubs in treatment and control are distributed geographically.

Two weeks and three days after the NFF sent the email to the clubs in the treatment group, we sent fictitious applications to join a training session to all clubs. Sending an email is a common way in Europe to get in touch with a sports club. Soccer clubs in Norway typically provide an email address on their website. For practical reasons, we sent out the emails on two days (on Monday 1 November 2021 and Tuesday 2 November 2021) rather than on a single day. We randomly assigned all cubs (independent of treatment) to one of the two days. We constructed email accounts, using Gmail, with typical Norwegian- and foreign-sounding names. We tested the email accounts to confirm that

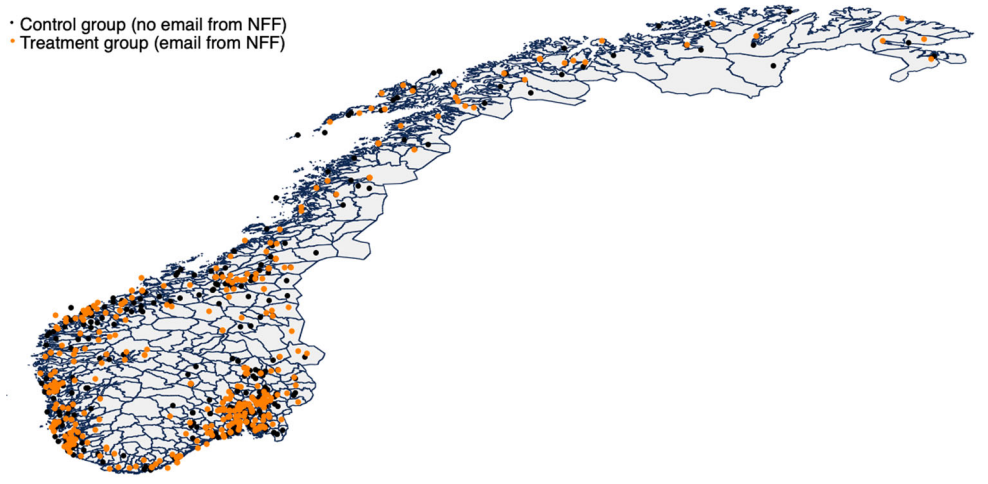

Figure 1. Geographic distribution of clubs by treatment assignment.

emails do not end up in a spam folder. Following the approach in Nesseler, Gomez-Gonzalez, and Dietl (2019), we created names for the three largest foreign groups in Norway (Polish, Lithuanian, and Somalian).[4,5] Each name was randomly assigned to each club. Block randomisation ensured that names were not overrepresented in specific blocks.[6]

Clubs received an email with the following text (translated from Norwegian; the original Norwegian text is included in Appendix A):

Subject*: Training possibility.*
Hello,
I am looking for a new soccer club. Do you think I could come and join a training session?
Thanks!
[Name]

We collected responses to fictitious applications for six weeks (from 1 November 2021 until 12 December 2021). Almost all responses arrived soon after we sent the application: 79% of responses arrived within two days and 94% of responses arrived in the first week after sending. We received two responses in the last two weeks; two responses in the fifth week, and no responses in the sixth week. In total, we received 554 responses (57% response rate). From the remaining 413 clubs, we received no response. We coded the responses as positive without additional inquiries (n = 291), positive with additional inquiries (n = 254), and negative (n = 9). Positive responses invited an individual to a training, oftentimes specifying day and time. Positive responses with additional inquiries typically asked applicants for information about their age, previous experience in other clubs or divisions, or their preferred playing position. Negative responses typically denied the opportunity to join a training because clubs were full or had had already too many other requests. The response rates are close to the ones found in 2019 in Norway, reported in Gomez-Gonzalez, Nesseler, and Dietl (2021).

Table 1 gives an overview of the data. The complete dataset is publicly available.[7]

We received ethical approval from the Human Subjects Committee of the Faculty of Economics, Business Administration, and Information Technology of the University of Zurich (OEC IRB # 2021-031) and registered the experiment before we sent out emails

(https://doi.org/10.1257/rct.8049-1.0). Our experimental setup had three potential ethical issues. First, the research involves subjects that are uninformed about their participation in the study. This approach has the disadvantage that respondents might participate who do not wish to participate in the research. However, not informing respondents has the advantage that we receive an undistorted response. Second, the subjects are misled. We tell the subjects that a person would like to join their club. However, this person does not exist. The subjects invest time to read the request and to answer it – if they answer. To minimise the effort of the subjects, we write a response email soon after a subject contacts us, informing that the applicant is no longer interested. This assures that subjects do not invest more time in the non-existing individual.[8] Third, the data is not anonymized at the beginning of the research project. However, the name of the club and the name of any individual have been deleted from the dataset right after processing the data. We published our completely anonymized dataset in HarvardDataverse (the data will be available upon publication). While we acknowledge these potential ethical issues, studying interventions to fight discrimination has the potential to create important benefits for society, which should be traded off against the ethical concerns (Asiedu et al. 2021; Glennerster 2017).

3. Results

Following the pre-registration, we pooled the responses to the fictitious applications into two categories: 'negative' and 'positive'. Negative responses include declines and non-responses and positive responses include positive responses with and without additional inquiries. Figure 2 shows our main results. It plots the share of positive responses to applications split by native- and foreign-sounding name of the applicant and treatment and control group of the coach.

Confirming previous research, we find evidence for discrimination: people with a foreign-sounding name receive fewer responses. This holds particularly for coaches in the treatment group – where the gap is 15 percentage points – but also in control, where the gap is 8 percentage points. In the Gomez-Gonzalez, Nesseler, and Dietl (2021) study, the gap found for Norway in 2019 in the same context was 11 percentage points.

Table 1. Descriptive statistics.*

Variable	*N*	Mean	Std. Dev.	Min	Max
Male or female soccer club (Male = 1)	967	.808	.394	0	1
Treatment group (email from federation) or control group (Treatment group = 1)	967	.482	.500	0	1
Native- or foreign-sounding name of applicant (Native = 1)	967	.481	.500	0	1
Application sent out on Monday or Tuesday (Monday = 1)	967	.515	.500	0	1
Responses to applications					
Any response**	554	0.57			
Positive response without further inquiries	291	0.30			
Positive response with further inquiries	254	0.26			
Negative response	9	0.01			
No response	413	0.43			

Notes: * In addition to the variables included in the table, we have data about the clubs' leagues (26 different leagues) and regions (18 regions). In 2017, Norway merged several regions into 11 administrative regions. However, the NFF still follows the previous 2017 version.
**If we only received an automatic response, this is counted as no response.

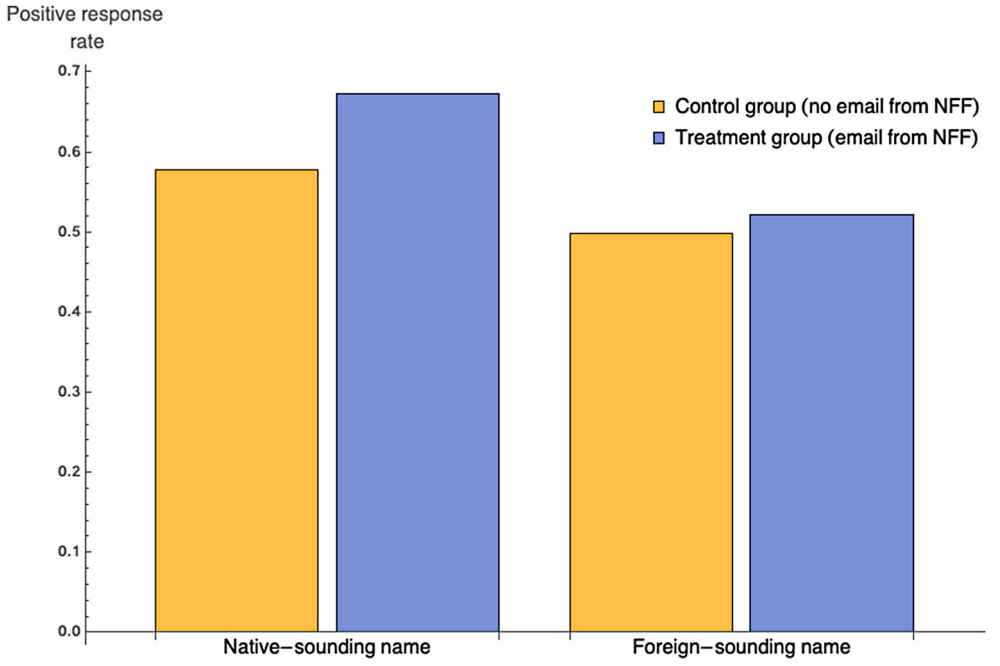

Figure 2. Share of positive responses to applications with native- and foreign-sounding names for treatment and control group coaches.

Next, let us consider the treatment effect of the intervention. Comparing the positive response rates to applicants with foreign-sounding names for the control and treatment group, we find a tiny increase of slightly more than two percentage points. Surprisingly, there is a large positive treatment effect of 10 percentage points on the positive response rate to people with a native-sounding name. The gap in response rates to applications from people with native- and foreign-sounding names thus widened as a result of the treatment. This runs counter to our prediction and was also not intended by the NFF.

Table 2 confirms these results using regression analyses, where the dependent variable $PositiveResponse_i$ is a dummy indicating whether we received a positive response to the fictitious application (dependent variable equals one) or not (dependent variable equals zero).

The regression equation reads:

$$\begin{aligned}\text{Positive Response}_i = &\alpha_0 + \beta_1 \text{Foreign} - \text{sounding name}_i + \beta_2 \text{Email from NFF}_i \\ &\times \text{Native} - \text{sounding name}_i \\ &+ \beta_3 \text{Email from NFF}_i \times \text{Foreign} - \text{sounding name}_i + \varepsilon_i\end{aligned}$$

The subscript i refers to a coach or club. β_1 is an estimate of how much a foreign-sounding name matters for getting a positive response, given that no email was sent to the coach or club by the NFF. β_2 is an estimate of the effect of an email from the NFF on the positive response rate to applications from people with native-sounding names. Likewise, β_3 is an estimate of the effect of an email from the NFF on the positive response rate to applications from people with foreign-sounding names. ε_i is the error term. In some

Table 2. Regression results.

	Dependent variable: positive response = 1, no or negative response = 0		
	Model 1	Model 2	Model 3
Native-sounding name	omitted	omitted	omitted
Foreign-sounding name	−0.08*	−0.08*	−0.10**
	(0.04)	(0.04)	(0.05)
Email from NFF × Native-sounding name	0.10**	0.09**	0.08*
	(0.04)	(0.04)	(0.05)
Email from NFF × Foreign-sounding name	0.02	0.02	0.04
	(0.04)	(0.04)	(0.05)
Male soccer team		−0.01	0.07
		(0.04)	(0.10)
Application sent out Monday or Tuesday (Monday = 1)		−0.07**	−0.06**
		(0.03)	(0.03)
League and region control			Yes
Constant	0.58***	0.62***	0.28
	(0.04)	(0.05)	(0.22)
Observations	967	967	967
R^2	0.017	0.022	0.053

Notes: Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

regression models, we also control for several background variables: male teams, the day the application was sent out (Monday or Tuesday), and in which league a team plays.

Table 2 shows that discrimination of people with a foreign-sounding name is statistically significant at the 10% level (and at the 5% level after adding controls, see column 3). The letter from the NFF has a statistically significant positive effect on the response rate to applicants with a native-sounding name and hardly any effect on the response rate to applications with a foreign-sounding name. Surprisingly, we find that emails sent on Monday 1 November 2021 receive fewer responses than those sent on Tuesday 2 November 2021.

How can we interpret these results? One possible interpretation is that the email led to feelings of resentment among some coaches, resulting in defiant responses. As discussed in the Introduction, defiant responses to moral appeals have been observed before in several other contexts. In our experiment, a defiant response would be to become less open to people with a foreign-sounding name and more open to people with a native-sounding name. The fact that, overall, we do not find a negative effect on the responses to applicants with foreign-sounding names might result from other coaches responding in the predicted and intended way: more open to people with a foreign-sounding name and equally open to people with a native-sounding name. On net, the pattern that we see in the data can result: more positive responses to people with a native-sounding name and no change in responses to people with a foreign-sounding name. However, as we shall see in the remainder of this section, some further exploratory analyses point toward two alternative interpretations – a descriptive norm effect and an encouragement effect.

Further exploratory analyses

We performed a few more regression analyses, which were not pre-registered and so are of a more exploratory nature.

First, we studied how much our results depend on the exact definition of a positive response. Table S2 in Appendix B uses positive responses without any additional inquiries as the dependent variable instead of the broader definition used above, which includes positive responses with additional inquiries. Using this stricter definition, we find slightly more discrimination in the control group and a much smaller treatment effect for applicants with native-sounding names, which is statistically indistinguishable from zero. Hence, the sizeable and statistically significant treatment effect reported in Table 2 stems mainly from coaches moving from a negative or no response to a conditional positive response, and not so much to an unconditional positive response. Otherwise, the results in Table S2 are by and large similar to those in Table 2.

Next, we study whether our results differ between more and less populous regions. Earlier studies have shown that ethnic discrimination is particularly prevalent in less populous regions (Huijsmans et al. 2021; Mayda 2006).[9] Figure 3 and the accompanying regression results in Table 3 confirm this. Focusing on the control group, we find that discrimination is absent (or even slightly positive) in regions with more than 100,000 citizens. In contrast, in regions with less than 100,000 citizens, there is a sizeable difference of more than 20 percentage points between the response rates to applications from people with native-sounding names and those from people with foreign-sounding names.

Moreover, Figure 3 and Table 3 show a striking difference in the estimated effect of the treatment. Whereas the treatment did not have any effect in regions with less than 100,000 citizens, it had a sizeable effect of 18 percentage points in regions with more than 100,000 citizens on positive responses to applications from people with a native-sounding name. Positive responses to applications from people with a foreign-sounding name also went up in these regions, but to a smaller extent (about 8 percentage points) and statistically insignificant.

Table 3. Regression results for subsamples by population size of the region.[#]

	Dependent variable: positive responses = 1, no or negative response = 0	
	Regions with less than 100,000 citizens	Regions with more than 100,000 citizens
Native-sounding name	omitted	omitted
Foreign-sounding name	−0.21***	0.01
	(0.07)	(0.06)
Email from NFF × Native-sounding name	−0.04	0.18***
	(0.07)	(0.06)
Email from NFF × Foreign-sounding name	−0.01	0.08
	(0.07)	(0.06)
Male soccer team	0.09	0.04
	(0.14)	(0.15)
Application sent out Monday or Tuesday (Monday = 1)	−0.04	−0.08*
	(0.05)	(0.04)
League and region control	Yes	Yes
Constant	0.21***	0.34
	(0.07)	(0.29)
Observations	439	528
R^2	0.090	0.074

Notes: [#]Excluding/Including only Oslo, Hordaland, Rogaland, Trøndelag, Buskerud, and Østfold.
Standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

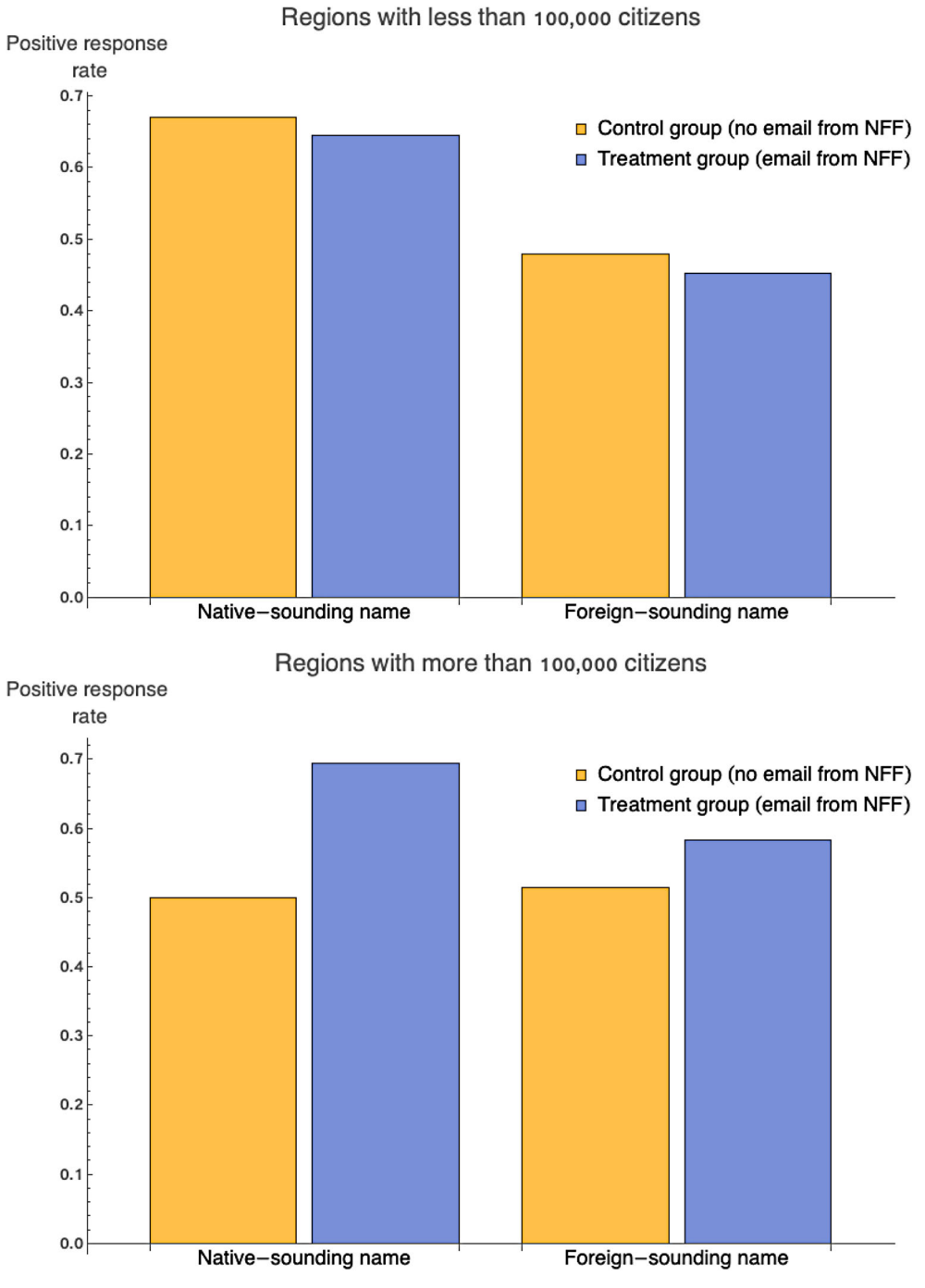

Figure 3. Share of positive responses to applications with native- and foreign-sounding names for treatment and control group coaches by population size of the region.

Summarising, for the regions where we find sizeable discrimination in the absence of treatment, we find no effect of the treatment whatsoever. The resulting policy implication is that, if one wishes to reduce discrimination where discrimination is prevalent, the intervention we tested in this paper does not help. For the regions where there is no discrimination in the absence of treatment, we find that the treatment leads to a large increase in positive responses to applicants with native-sounding names, and thus leads to discrimination. As we conjectured above, feelings of resentment may drive the average treatment effect of the intervention. However, this interpretation does not seem very likely given our finding that the average treatment effect originates entirely from the non-discriminating coaches.

The results, however, also suggest an alternative interpretation. As discussed in the Introduction, a rich body of empirical evidence shows that people tend to conform to other people's behaviour, i.e. they follow descriptive norms. The email from the NFF mentioned that studies have shown that discrimination is present in amateur soccer. In the less populous regions, this message may contain no news: as our data clearly show, discrimination is prevalent there, and coaches may be well aware of this. In more populous regions, however, discrimination is absent and so the message in the email from the NFF that discrimination occurs may change coaches' beliefs about what is the apparent social norm and they may adjust their behaviour accordingly. As a result, in those regions, the intervention may have unintentionally backfired. The descriptive norm theory can, however, not explain that the increased discrimination only arose from a substantial increase in coaches' positive responses to applications from people with a native-sounding name.

The third interpretation, which was offered by one of the reviewers, is perhaps the most compelling. The email from the NFF informs or reminds coaches that amateur soccer can unite people from different ethnic backgrounds and facilitate integration on and off the field. Thus, driven by this message, some coaches may have decided to invite more native players to expose them to their ethnically diverse soccer teams. This interpretation is also consistent with our finding that the increase in responses to individuals with native-sounding names happens among counterfactually non-discriminating coaches only.

4. Concluding remarks

We performed a field experiment together with the NFF to examine if a low-cost intervention decreases discrimination in the context of amateur soccer. The intervention was an email from the NFF which described the important role of soccer for inclusivity and fighting racism and asked coaches to remain open to all interested applicants. Specifically, the goal was to keep a high response rate for all applicants but to decrease the gap between people with a native- and a foreign-sounding name. The email also mentioned that studies have shown that discrimination occurs in amateur soccer. We measured the effect of the intervention by sending fictitious applications to join a training using a native-sounding or a foreign-sounding name.

Our data show that discrimination is prevalent in Norwegian amateur soccer. The gap in positive response rates to applications with native- and foreign-sounding names

amounts to 11 percentage points in the full sample. These results contribute to the debate regarding discrimination of ethnic minorities.

Surprisingly, the low-cost intervention did not reduce discrimination. On the contrary, we find that the intervention increased discrimination. The increased inequality is driven by coaches being more open to applicants with native-sounding names. We discussed three possible interpretations of our findings. The email might have led some coaches to feel resentful and respond in a defiant way. Alternatively, the information in the email that discrimination occurs in amateur soccer may have changed coaches' perception of the descriptive norm, leading them to discriminate more. A third and, we believe, most convincing interpretation of the findings focuses on the part of the email that stresses the role that soccer can play in reducing racism. This could have encouraged coaches of diverse teams to invite more natives. This interpretation is in line with the fact that the increase in positive responses to applications with native-sounding names occurred among coaches from areas where discrimination in amateur soccer is rare.

To further explore the empirical relevance of possible interpretations of our findings, it would have been great if we could have collected more data – e.g. questionnaire data or interview data – about how the intervention is perceived by different types of coaches. This was outside of the scope of the present study, but we intend to include it in possible follow-up studies. Other interesting avenues for future research – that we were not able to do because of data limitations or lack of statistical power – include: i) studying whether the ethnicity and other background characteristics of the coach and the team matter for the severity and direction of discrimination and for the response to the treatment; ii) studying whether the ethnicity of the applicant matters for these outcomes; iii) studying the effects of adding or taking away particular elements of the intervention email (e.g. regarding the information about current discrimination in amateur soccer); iv) studying the influence on the response rate of different days or weeks on which an application is sent.

Our findings are important for practice and future research as they show that, even though an intervention is well-intended and has been proven to be effective in another context (Boring and Philippe 2021), it can have no or even a negative effect. Many organisations throughout the world have taken initiatives to fight discrimination. However, interventions are rarely tested through large-scale field experiments. Our results show that refraining from doing so implies running the risk of making things worse rather than better (cf. Behaghel, Crépon, and Le Barbanchon 2015).

Notes

1. For detailed statistics see https://www.ssb.no/en/befolkning/innvandrere/statistikk/innvandrere-og-norskfodte-med-innvandrerforeldre.
2. 620 of the email addresses we used are the personal email addresses of coaches. The remaining 347 email addresses are the email addresses of administrative personnel of the clubs. The rate and nature of responses to applications sent to the two types of email addresses are very similar and do not differ significantly.
3. This contrasts with the message in the email from the NFF that the email was sent to all coaches. In agreement with the NFF, sending out the rest of the emails was postponed at least until the end of the experiment.

4. The Norwegian statistics bureau (SSB) categorises the largest foreign groups. For more information, see https://www.ssb.no/en/befolkning/innvandrere/statistikk/innvandrere-og-norskfodte-med-innvandrerforeldre.
5. We asked students before the start of a lecture (in classes in Ålesund and Trondheim in Spring 2021) to fill out a questionnaire and state if a name sounded Norwegian or foreign. The results of the survey are available in Table S1 in Appendix B and show that the names we use are identified as native or foreign by close to 100% of respondents.
6. As a result of this randomisation, some clubs received an application with a name associated with a group that is not greatly represented in their area, e.g., an application with a Somalian-sounding name to a club in a rural area. However, since we used names for the three largest foreign groups in Norway and immigrants are relatively widely spread geographically in Norway, we do not think that this caused any irregularities. In Norway, young people with an immigrant background historically participate less in sports than those without an immigrant background (Walseth 2008). The differences are particularly salient among young women. While self-limiting behaviours related to culture or religion could at least partially explain this underrepresentation, our experiment studies whether exclusion by clubs also plays a role.
7. The data (upon publication) is available at https://doi.org/10.7910/DVN/ZETZDZ (the page will be activated upon publication of the paper). Individually identifiable data is excluded.
8. We considered the possibility of debriefing subjects after the experiment, but decided against it following the advice by the Human Subjects Committee, who argued it may create more harm than good.
9. 439 clubs are in regions with less than 100,000 citizens and 528 are in regions with more than 100,000 citizens. A map showing where these clubs are located is available in Figure S1 in Appendix B. In regions with less than 100,000 citizens, 21.0% of the teams are female and 79.0% are male teams. In regions with more than 100,000 citizens, 17.8% of the teams are female and 82.2% are male teams.

Acknowledgements

This study was approved by the Human Subjects Committee of the Faculty of Economics, Business Administration, and Information Technology of the University of Zurich (OEC IRB # 2021-031) and pre-registered at the AEA RCT Registry (AEARCTR-0008049; https://doi.org/10.1257/rct.8049-1.0). We are grateful to two reviewers of this journal, Anne Boring, Francesco Capozza, Josse Delfgaauw, Donald Green, Sacha Kapoor, and Arnfinn Midtbøen for helpful comments on an earlier version of this paper. We are also thankful for the support of the Norwegian Football Federation and, especially, the help of Henrik Lunde.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Robert Dur http://orcid.org/0000-0001-9081-4179
Carlos Gomez-Gonzalez http://orcid.org/0000-0001-8610-4828
Cornel Nesseler http://orcid.org/0000-0001-7900-4175

References

Alesina, A., M. Carlana, E. La Ferrara, and P. Pinotti. 2018. Revealing Stereotypes: Evidence from Immigrants in Schools. National Bureau of Economic Research Working Paper No. 25333.

Allcott, H., and T. Rogers. 2014. "The Short-run and Long-run Effects of Behavioral Interventions: Experimental Evidence from Energy Conservation." *American Economic Review* 104 (10): 3003–3037.

Andersson, L., N. Jakobsson, and A. Kotsadam. 2012. "A Field Experiment of Discrimination in the Norwegian Housing Market: Gender, Class, and Ethnicity." *Land Economics* 88 (2): 233–240.

Ariel, B. 2012. "Deterrence and Moral Persuasion Effects on Corporate Tax Compliance: Findings From a Randomized Controlled Trial*." *Criminology; An interdisciplinary Journal* 50 (1): 27–69.

Asiedu, E., D. Karlan, M. P. Lambon-Quayefio, and C. R. Udry. 2021. "A Call for Structured Ethics Appendices in Social Science Papers." *Proceedings of the National Academy of Sciences* 118 (29): e2024570118.

Auspurg, K., A. Schneck, and T. Hinz. 2019. "Closed Doors Everywhere? A Meta-Analysis of Field Experiments on Ethnic Discrimination in Rental Housing Markets." *Journal of Ethnic and Migration Studies* 45 (1): 95–114.

Behaghel, L., B. Crépon, and T. Le Barbanchon. 2015. "Unintended Effects of Anonymous Résumés." *American Economic Journal: Applied Economics* 7 (3): 1–27.

Bertrand, M., D. Chugh, and S. Mullainathan. 2005. "Implicit Discrimination." *American Economic Review* 95 (2): 94–98.

Bertrand, M., and E. Duflo. 2017. "Field Experiments on Discrimination." In *Handbook of Field Experiments, Volume 1*, edited by A. V. Banerjee, and E. Duflo, 309–393. Amsterdam: North-Holland Publishing.

Bertrand, M., and S. Mullainathan. 2004. "Are Emily and Greg More Employable Than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination." *American Economic Review* 94 (4): 991–1013.

Bicchieri, C., and E. Dimant. 2019. Nudging with Care: The Risks and Benefits of Social Information. Public Choice, forthcoming.

Blumenthal, M., C. Christian, and J. Slemrod. 2001. "Do Normative Appeals Affect tax Compliance? Evidence from a Controlled Experiment in Minnesota." *National Tax Journal* 54 (1): 125–138.

Boring, A., and A. Philippe. 2021. "Reducing Discrimination in the Field: Evidence from an Awareness Raising Intervention Targeting Gender Biases in Student Evaluations of Teaching." *Journal of Public Economics* 193: 104323.

Bott, K. M., A. W. Cappelen, E. Sorensen, and B. Tungodden. 2020. "You've Got Mail: A Randomized Field Experiment on Tax Evasion." *Management Science* 66 (7): 2801–2819.

Bouffard, L. A., and N. L. Piquero. 2010. "Defiance Theory and Life Course Explanations of Persistent Offending." *Crime & Delinquency* 56 (2): 227–252.

Bourabain, D., and P. P. Verhaeghe. 2019. "Could you Help me, Please? Intersectional Field Experiments on Everyday Discrimination in Clothing Stores." *Journal of Ethnic and Migration Studies* 45 (11): 2026–2044.

Bradler, C., R. Dur, S. Neckermann, and A. Non. 2016. "Employee Recognition and Performance: A Field Experiment." *Management Science* 62 (11): 3085–3099.

Chen, Y., F. M. Harper, J. Konstan, and S. X. Li. 2010. "Social Comparisons and Contributions to Online Communities: A Field Experiment on Movielens." *American Economic Review* 100 (4): 1358–1398.

Cialdini, R. B., R. R. Reno, and C. A. Kallgren. 1990. "A Focus Theory of Normative Conduct: Recycling the Concept of Norms to Reduce Littering in Public Places." *Journal of Personality and Social Psychology* 58 (6): 1015–1026.

Diehl, C., V. A. Andorfer, Y. Khoudja, and K. Krause. 2013. "Not in my Kitchen? Ethnic Discrimination and Discrimination Intentions in Shared Housing among University Students in Germany." *Journal of Ethnic and Migration Studies* 39 (10): 1679–1697.

Dur, R., and B. Vollaard. 2015. "The Power of a Bad Example." *Environment and Behavior* 47 (9): 970–1000.

Edelman, B., M. Luca, and D. Svirsky. 2017. "Racial Discrimination in the Sharing Economy: Evidence from a Field Experiment." *American Economic Journal: Applied Economics* 9 (2): 1–22.

Frey, B. S., and S. Meier. 2004. "Social Comparisons and Pro-Social Behavior: Testing "Conditional Cooperation" in a Field Experiment." *American Economic Review* 94 (5): 1717–1722.
Ge, Y., C. R. Knittel, D. MacKenzie, and S. Zoepf. 2020. "Racial Discrimination in Transportation Network Companies." *Journal of Public Economics* 190: 104205.
Gerber, A. S., and T. Rogers. 2009. "Descriptive Social Norms and Motivation to Vote: Everybody's Voting and so Should you." *The Journal of Politics* 71 (1): 178–191.
Glennerster, R. 2017. "The Practicalities of Running Randomized Evaluations: Partnerships, Measurement, Ethics, and Transparency." In *Handbook of Field Experiments, Volume 1*, edited by A. V. Banerjee, and E. Duflo, 175–243. Amsterdam: North-Holland Publishing.
Gomez-Gonzalez, C., C. Nesseler, and H. M. Dietl. 2021. "Mapping Discrimination in Europe Through a Field Experiment in Amateur Sport." *Humanities and Social Sciences Communications* 8 (1): 1–8.
Hallsworth, M., T. Chadborn, A. Sallis, M. Sanders, D. Berry, F. Greaves, L. Clements, and S. C. Davies. 2016. "Provision of Social Norm Feedback to High Prescribers of Antibiotics in General Practice: A Pragmatic National Randomised Controlled Trial." *The Lancet* 387 (10029): 1743–1752.
Heath, A., and V. Di Stasio. 2019. "Racial Discrimination in Britain, 1969–2017: A Meta-Analysis of Field Experiments on Racial Discrimination in the British Labour Market." *The British Journal of Sociology* 70: 1774–1798.
Huijsmans, T., E. Harteveld, W. Van der Brug, and B. Lancee. 2021. "Are Cities Ever More Cosmopolitan? Studying Trends in Urban-Rural Divergence of Cultural Attitudes." *Political Geography* 86: 102353.
Jakobsson, N., and H. Lindholm. 2014. "Ethnic Preferences in Internet Dating: A Field Experiment." *Marriage & Family Review* 50 (4): 307–317.
Kaas, L., and C. Manger. 2012. "Ethnic Discrimination in Germany's Labour Market: A Field Experiment." *German Economic Review* 13 (1): 1–20.
Kahan, D. M. 1997. "Social Influence, Social Meaning, and Deterrence." *Virginia Law Review* 83: 349–395.
Keizer, K., S. Lindenberg, and L. Steg. 2008. "The Spreading of Disorder." *Science* 322: 1681–1685.
Lancee, B. 2021. "Ethnic Discrimination in Hiring: Comparing Groups Across Contexts. Results from a Cross-National Field Experiment." *Journal of Ethnic and Migration Studies* 47 (6): 1181–1200.
Larsen, E. N., and V. Di Stasio. 2021. "Pakistani in the UK and Norway: Different Contexts, Similar Disadvantage. Results from a Comparative Field Experiment on Hiring Discrimination." *Journal of Ethnic and Migration Studies* 47 (6): 1201–1221.
Liebe, U., and H. Beyer. 2021. "Examining Discrimination in Everyday Life: A Stated Choice Experiment on Racism in the Sharing Economy." *Journal of Ethnic and Migration Studies* 47 (9): 2065–2088.
Lowe, M. 2021. "Types of Contact: A Field Experiment on Collaborative and Adversarial Caste Integration." *American Economic Review* 111 (6): 1807–1844.
Mayda, A. M. 2006. "Who is Against Immigration? A Cross-Country Investigation of Individual Attitudes Toward Immigrants." *Review of Economics and Statistics* 88 (3): 510–530.
Midtbøen, A. H. 2016. "Discrimination of the Second Generation: Evidence from a Field Experiment in Norway." *Journal of International Migration and Integration* 17 (1): 253–272.
Mousa, S. 2020. "Building Social Cohesion Between Christians and Muslims Through Soccer Inpost-ISIS Iraq." *Science* 369 (6505): 866–870.
Mujcic, R., and P. Frijters. 2021. "The Colour of a Free Ride." *The Economic Journal* 131 (634): 970–999.
Nesseler, C., C. Gomez-Gonzalez, and H. Dietl. 2019. "What's in a Name? Measuring Access to Social Activities with a Field Experiment." *Palgrave Communications* 5: 1–12.
Nyhan, B., and J. Reifler. 2015. "Does Correcting Myths About the flu Vaccine Work? An Experimental Evaluation of the Effects of Corrective Information." *Vaccine* 33 (3): 459–464.
Nyhan, B., J. Reifler, S. Richey, and G. L. Freed. 2014. "Effective Messages in Vaccine Promotion: A Randomized Trial." *Pediatrics* 133 (4): e835–e842.

Oreopoulos, P. 2011. "Why do Skilled Immigrants Struggle in the Labor Market? A Field Experiment with Thirteen Thousand Resumes." *American Economic Journal: Economic Policy* 3 (4): 148–171.

Pager, D., B. Bonikowski, and B. Western. 2009. "Discrimination in a Low-Wage Labor Market." *American Sociological Review* 74 (5): 777–799.

Paluck, E. L., R. Porat, C. S. Clark, and D. P. Green. 2021. "Prejudice Reduction: Progress and Challenges." *Annual Review of Psychology* 72: 533–560.

Quillian, L., and A. H. Midtbøen. 2021. "Comparative Perspectives on Racial Discrimination in Hiring: The Rise of Field Experiments." *Annual Review of Sociology* 47: 391–415.

Quillian, L., D. Pager, O. Hexel, and A. H. Midtbøen. 2017. "Meta-analysis of Field Experiments Shows no Change in Racial Discrimination in Hiring Over Time." *Proceedings of the National Academy of Sciences* 114 (41): 10870–10875.

Riach, P. A., and J. Rich. 2002. "Field Experiments of Discrimination in the Market Place." *The Economic Journal* 112 (483): F480–F518.

Rooth, D. O. 2010. "Automatic Associations and Discrimination in Hiring: Real World Evidence." *Labour Economics* 17 (3): 523–534.

Sawert, T. 2020. "Understanding the Mechanisms of Ethnic Discrimination: A Field Experiment on Discrimination Against Turks, Syrians and Americans in the Berlin Shared Housing Market." *Journal of Ethnic and Migration Studies* 46 (19): 3937–3954.

Schultz, P. W., J. M. Nolan, R. B. Cialdini, N. J. Goldstein, and V. Griskevicius. 2007. "The Constructive, Destructive, and Reconstructive Power of Social Norms." *Psychological Science* 18 (5): 429–434.

Thijssen, L., B. Lancee, S. Veit, and R. Yemane. 2021. "Discrimination Against Turkish Minorities in Germany and the Netherlands: Field Experimental Evidence on the Effect of Diagnostic Information on Labour Market Outcomes." *Journal of Ethnic and Migration Studies* 47 (6): 1222–1239.

Walseth, K. 2008. "Bridging and Bonding Social Capital in Sport—Experiences of Young Women with an Immigrant Background." *Sport, Education and Society* 13 (1): 1–17.

Weichselbaumer, D. 2020. "Multiple Discrimination Against Female Immigrants Wearing Headscarves." *ILR Review* 73 (3): 600–627.

Zschirnt, E., and D. Ruedin. 2016. "Ethnic Discrimination in Hiring Decisions: A Meta-Analysis of Correspondence Tests 1990–2015." *Journal of Ethnic and Migration Studies* 42 (7): 1115–1134.