



HAL
open science

Anti-Muslim behavior in everyday interaction: evidence from a field experiment in Paris

Martin Aranguren, Francesco Madrisotti, Eser Durmaz-Martins

► **To cite this version:**

Martin Aranguren, Francesco Madrisotti, Eser Durmaz-Martins. Anti-Muslim behavior in everyday interaction: evidence from a field experiment in Paris. *Journal of Ethnic and Migration Studies*, inPress, 10.1080/1369183X.2021.1953378 . hal-03285818

HAL Id: hal-03285818

<https://hal.science/hal-03285818>

Submitted on 13 Jul 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Title: Anti-Muslim behavior in everyday interaction: evidence from a field experiment in Paris

Authors

Martin Aranguren^{1*}, Francesco Madrisotti¹, Eser Durmaz-Martins¹

Affiliations

¹ Université de Paris, Centre National de la Recherche Scientifique, Unité de Recherches Migrations et Sociétés, Paris, France

*** Corresponding author**

E-mail: martin.aranguren@cnrs.fr

Forthcoming in Journal of Ethnic and Migration Studies (last update July 13th, 2021)

Authors contributions

MA: conceptualization, methodology, formal analysis, resources, data curation, writing-original draft presentation, writing-review and editing, visualization, supervision, project administration, funding acquisition

FM: investigation, writing-review and editing, project administration

ED-M: investigation, writing-review and editing

Acknowledgments

We thank Carlo Barone, Alessandro Bergamaschi, Jocelyne Streiff-Fénart and two anonymous reviewers for useful comments on earlier versions of the manuscript.

Funding statement

The present research was funded through a Momentum grant to Martin Aranguren awarded by the MITI departement of Centre National de la Recherche Scientifique.

Abstract

No other form of group antagonism affects the fate of so many people in France as anti-Muslim racism. While negative attitudes toward Muslims and Muslims' experience of discrimination are well documented, studies of anti-Muslim behavior are rare, especially in the context of everyday interpersonal encounters. To fill this void, we conducted a field experiment on platforms of the Paris metro (n=270) in which a bearded confederate asked for help to randomly selected passengers giving additional indirect cues of being Muslim in the experimental condition. The outcomes under investigation were the probability of helping the confederate and various behaviors indicative of interpersonal warmth or involvement. Interactions were videotaped, the outcomes objectively measured, and the data analyzed using Generalized Linear Models estimated with Bayesian inference. Passengers were found to offer help less often and to show lower interpersonal warmth in the experimental condition. Also, when considered in isolation the young turn out to discriminate but not the middle-aged. Given that these negative effects were observed despite the use of a minimal stimulus, the results probably underestimate the actual level of anti-Muslim discrimination that Muslim men face in their everyday dealings with non Muslims.

In France, the country with the largest Muslim population of Europe (Pew Research Center, 2017), no other form of racism affects the fate of so many people as anti-Muslimism (Halliday, 1999), broadly understood as the antagonistic attitudes and practices that non Muslims direct to Muslims because of their religious identity. Whereas a significant number of studies have dealt with anti-Muslimism, the various facets of the phenomenon have received unequal amounts of attention.

The attitudes of non Muslims towards Muslims are regularly measured by instruments such as the European Values Study (2017), the Eurobarometer (European Commission, 2019) or the Baromètre racisme (CNCDH, 2019). A number of surveys have documented Muslims' experiences of being discriminated in France (Brinbaum et al., 2018; Eberhard & Simon, 2016; FRA, 2017; Pew Research Center, 2006). In contrast, aside from a handful of correspondence studies on hiring discrimination in the job market (Adida et al., 2010; Koopmans et al., 2019; Pierné, 2013; Di Stasio et al., 2021; Valfort, 2020; for a review with a wider focus see also Adida et al., 2016), little is known about anti-Muslim discriminatory behavior, especially in the context of everyday interpersonal relations (but see Aranguren et al., 2021). The present article seeks to contribute to this emerging field.

The dearth of behavioral studies on anti-Muslim discriminatory behavior is regrettable for two important reasons. First, feeling and being discriminated are two distinct phenomena, the one manifesting itself in the realm of subjective experience and the other in the extra-subjective domain of objective behavior. It follows that mistakes are possible: one may feel discriminated without actually being so, or not feel discriminated even as one actually is. Second, in Western democracies the law and ordinary morals condemn discrimination on various grounds, including religious affiliation, and this consensus covers the entire political spectrum nowadays, at least officially. But the remedy for this unanimously acknowledged social evil remains ambiguous if a group's feeling of being discriminated cannot be shown to correspond to objective acts of discrimination perpetrated to the detriment of the group. It is not clear whether the fundamental solution lies in inviting the members of the afflicted group to change their perception of the way in which they are

treated by others or, alternatively, in urging others to stop discriminating against members of that group. If experiences and behavior diverge, the solution is in the psyche of the person or group who feels discriminated. If both converge instead, the remedy is in the discriminator's demeanor. In the presence of evidence of felt discrimination, behavioral studies are important for avoiding the equally undesirable poles of unduly "blaming the victim" when discrimination is not only felt but also observed, or unduly "manufacturing the culprit" when discrimination is only felt but not observed.

The study reported here is part of a broader project concerned with anti-Muslim discrimination in situations of day-to-day living. Elsewhere, we have examined whether a woman receives a different treatment when she wears an islamic headscarf or hijab in the metros of Brussels, Paris and Vienna (Aranguren et al., 2021). The aim of the present study is to investigate if a man is treated differently when he gives indirect signs of being of Islamic faith.

We conducted a field experiment on platforms of the Paris metro in which a bearded confederate in his early thirties approached randomly selected male or female passengers of all ages asking for directions. To help passengers plan a route for him, the confederate mentioned a landmark. In the control condition, the landmark was Université Sorbonne Nouvelle. In the treatment condition, it was the Grande mosquée de Paris (both the university and the mosque are located in the same area). Whereas the confederate is credible as a follower of Islam in the eyes of most passengers, he is a French national who speaks the language without a noticeable accent and whose physical appearance does not locally connote a foreign origin. Aside from the beard, which at the time of the experiment could locally index observance of orthodox Islam as much as adherence to hipster fashion, the confederate carried no outward religious signs. The choice of this experimental manipulation obeys an *a fortiori* rationale: if this minimal stimulus is enough to provoke predicted differences in behavior among passengers, *a fortiori* these differences will be present if passengers are exposed to a stronger stimulus, such as encountering a man with a voluminous beard who wears a *jellaba* (a loose robe traditionally used for ceremonies in the Arab

world), speaks with a foreign accent, whose physical appearance is locally associated with Maghreb, and who explicitly states that he is going to the mosque to attend a prayer.

To assess differences in behavior between these two conditions, we took note of passenger's readiness to help the confederate but we also measured various indicators of interpersonal warmth or involvement. Previous field studies on the topic similarly provoking interactions with a covert confederate have considered helping and other explicit behaviors (Aidenberger & Doehne, 2021; Choi et al., 2020; Diekmann et al., 2014), and in some cases the procedure has sought to capture more subtle responses by including subjective assessments of interpersonal warmth (Hebl et al., 2002; King & Ahmad, 2010). Going one step up in objectivity, ours are the first field experiments to incorporate measures (not just ratings) of involvement behaviors based on videotapes.

Literature review and hypotheses

Anti-Muslimism

Three recent nationally representative surveys locate the incidence of anti-Muslim attitudes in France at 8% of the population. In the 2017 European Values Survey, 8% of French respondents mentioned Muslims as undesirable neighbors (European Values Study, 2017); in the French sample of the 2019 special Eurobarometer on discrimination, 8% stated that they would not feel comfortable having a Muslim colleague (European Commission, 2019); in the 2019 edition of *Baromètre racisme*, 8% of respondents agreed that "French Muslims are not French like the others" (CNCDH, 2019). Anti-Muslim attitudes rise when other questions are asked. Thus, in the cited Eurobarometer 32% of French respondents would not be comfortable if one of their children were in a love relationship with a Muslim (another measure of social distance) and in the 2019 *Baromètre racisme* 36% agree that French Muslims represent a separate group in society (a reverse-coded index of perceived integration in national life).

Available measures vary less as one moves from the attitudes of the non Muslim majority to Muslims' self-reported experience of being discriminated. The nationally representative Trajectoires et Origines survey indicated that the proportion of Muslims who had experienced some form of religion-based discrimination is 35% or 47%, depending on whether discrimination was directly self-reported or inferred by the researchers on the basis of indirect questions (Brinbaum et al., 2018). In a survey restricted to the Parisian region, 40% of Muslim respondents reported having experienced discrimination because of their religion (Eberhard & Simon, 2016). In the Pew Global Attitudes Survey the share of Muslim respondents who reported "a bad experience" because of their religion amounted to 39% in France (Pew Research Center, 2006), whereas in the French sample of the last EU-MIDIS 20% of Muslim respondents declared having experienced discrimination on the grounds of their religion or their religious beliefs in the five years preceding the survey (FRA, 2017).

Behavioral studies confirm to some extent that Muslims' experience of being discriminated is the subjective correlate of objective discrimination on the part of non Muslims, but the evidence is not without nuance. A number of correspondence studies found that when a CV is sent in response to a job opening in various sectors of the French labor market, the chance of being called for a hiring interview diminishes if the fake application indicates that the candidate is of Islamic faith (Adida et al., 2010; Pierné, 2013; Valfort, 2020). However, our own field experiment inquiring into the effects of the islamic headscarf or *hijab* on helping and interpersonal warmth does not support the expectation of monolithic hostility against Muslims (Aranguren et al., 2021). The only overall trend, in Paris as well as Brussels and Vienna, was to show not less but *more* nonverbal warmth in interaction with a hijab-wearing confederate. The hijab also gave rise to unfriendly behaviors, but these were mediated by the gender of the passenger, as elaborated below.

These uncertainties, nuances and exceptions notwithstanding, the existing evidence still favors the hypothesis of anti-Muslimism. We predict that, in interaction with the bearded

confederate in the mosque condition, passengers will offer less help and show lower interpersonal warmth.

Moderators: gender and age

In our previous work we have considered the possibility that gender moderates the expression of anti-Muslimism through behaviors that convey interpersonal warmth, such as sustained eye contact, close interpersonal distance and active participation in the conversation. These “involvement behaviors” have been hypothesized to serve two main alternative functions: spontaneously expressing one’s evaluation of the interaction partner or deliberately managing the impression that the partner forms of oneself (Patterson, 1982). The default hypothesis is that the degree of involvement shown in interaction with another person expresses the degree to which that person is positively evaluated. If a Muslim person is evaluated less positively than a non-Muslim one, we predict lower involvement with the Muslim person.

Now, a surprising yet robust finding is the fact that an actor may show *more* involvement (not less) when the interaction partner is negatively evaluated, perhaps in an effort to make the interaction less unpleasant (Bond, 1972; Coutts et al., 1980; Ickes et al., 1982). In the domain of intergroup relations, such deliberate management of impressions has been observed to occur among women but not men (Aranguren et al., 2021; Lemasson et al., 2021; Littleford et al., 2005), suggesting the hypothesis of an interaction effect between attitude and gender: confronted with a negatively viewed partner, men express their negative feelings by decreasing involvement, whereas women manage impressions by increasing it. Women’s tendency to exaggerate positivity in interaction with stigmatized others can be imputed to a stronger internalization, compared to men, of the social norm to control prejudice, a gender difference confirmed to be at work at least in Norway, Sweden and the UK (Harteveld & Ivarsflaten, 2018). From this, we predict that in the mosque condition the level of involvement will fall among men but rise among women.

Age is another possible moderator of anti-Muslimism on helping and involvement behaviors. The literature is remarkably split into research leading to the prediction of an intensification of prejudice with age, on the one hand, and work supporting the opposite expectation of a weakening of intergroup antagonism with the passing of biographical time, on the other.

The intensification hypothesis points to both cohort (or generational) effects and the consequences of aging. The expected cohort effect can be put as follows: in recent history, as generations follow one another the acceptability of prejudice decreases; having been socialized into a less prejudiced moral environment, the members of younger generations will show lower levels of prejudice than their seniors (e.g. Ford, 2008). When it comes to aging, it has been observed that with the years individuals become more conservative (Cornelis et al., 2009; Truett, 1993) and less open to experience (Allemand et al., 2008; Donnellan & Lucas, 2008; McCrae et al., 1999; Roberts et al., 2006; Soto et al., 2011; Srivastava et al., 2003). Both increased conservatism (Federico & Sidanius, 2002) and decreased openness (Ekehammar & Akrami, 2003), especially to values (Ekehammar & Akrami, 2007), are correlated with higher prejudice. Further, even when underlying prejudice is held constant, a consequence of aging on cognitive functioning is a reduced ability to inhibit the behavioral expression of prejudiced views (Gonsalkorale et al., 2009; Stewart et al., 2009; von Hippel et al., 2000), leading to the prediction of more discrimination among the old compared to the young.

The opposite weakening hypothesis relies on the observation that, as they age, individuals progressively become more agreeable, a personality trait exemplified by facets such as altruism and tender-mindedness (Allemand et al., 2008; Donnellan & Lucas, 2008; McCrae et al., 1999; Roberts et al., 2006; Soto et al., 2011; Srivastava et al., 2003). These findings have been questioned on the grounds that respondents may be conforming to social desirability pressures rather than giving an unbiased description of their basic tendencies (Köötus-Ausmees et al., 2020; Soubelet & Salthouse, 2011). The bottom line, then, is that as they age individuals desire more and more to portray themselves in a favorable light. Independently from research on personality, other work has

similarly found that age is positively correlated with trust in others (Li & Fung, 2013), including strangers (Bailey & Leon, 2019). To the extent that agreeableness (Ekehammar & Akrami, 2003), trust (Dhont & van Hiel, 2011) and social desirability (Batson et al., 1978) are all three negatively associated with prejudice, individuals are expected to become decreasingly antagonistic towards Muslims as aging makes them increasingly agreeable, trusting and/or avid for social approval.

Put in schematic form, these are the predictions that we undertake to test:

P1: passengers will offer less help and show less involvement in the mosque condition.

P2: the effect of the mosque condition on involvement will be positive among women but negative among men.

P3a: the effect predicted by P1 will be weaker among the young (or absent among the young but present among the middle-aged)

P3b: the effect predicted by P1 will be stronger among the young (or present among the young but absent among the middle-aged).

Method

Design

The experiment follows a 2 (experimental condition, see below) x 2 (passenger's gender) x 6 (metro stations: see below) between-subjects design with random assignment of participants to conditions and with equal sampling time devoted to each of the unique factor combinations. Age was measured as a covariate. For more details on the design and other aspects of the method, see (Aranguren et al., 2021).

Stations selection

The six observation sites were randomly drawn from the list of stations with intermediate ridership (quartiles 2 and 3) and with similarly arranged platforms (only one entrance, serving the trains of a unique line all going in the same direction). The selected stations are Boucicaut, Michel Bizot, Notre Dame de Lorette, Pyrénées, Ranelagh and Riquet.

Sampling

After a pilot study in September 2018, the main study proceeded between October 2 and November 2 of the same year. Acting as an Institutional Review Board, the CIL division of Centre National de la Recherche Scientifique approved the study (application number: 2-18019) and the transportation authority RATP gave us formal clearance to conduct the experiment on its premises.

We made five data collection visits within each of the six selected stations. All visits, scheduled at different weekdays within the same station, had a duration of two hours. Of these, we assigned the first hour to one experimental condition and the second hour to the other condition, balancing for the entire experiment the number of times that each condition was placed chronologically first or second. During the hour devoted to each condition, we sought to recruit an equal number of male and female passengers, which involved the combination of a method for approximating random selection and another one for stratifying the sampling of men and women.

Strict randomization was approximated with a method of systematic selection: during the time period comprised between the departure of the last train and the arrival of the following one, the confederate approached the first passenger who arrived at the platform. The stratification technique consisted in starting with the method of systematic selection regardless of the sex of the passenger, recruiting one passenger (for example, a man), and then reapplying the method of systematic selection but only to passengers of the opposite sex (in this case, women). The third passenger was again selected regardless of sex, the fourth by stratifying by sex, and so on. This means that, in stratifying our sample, we relied on our own commonsensical understandings of

sexual dimorphism to identify passengers as men or women, and not on passengers' self-reported gender identity.

Data collection visits took place on regular weekdays between 12pm and 2pm. In Paris, this is the only period of the working day in which waiting times are in the range of 3-5 minutes (instead of 1-2), maximizing the chances that the confederate will get to complete the script before the following train arrives.

Procedure

On one of the six platforms, a non-immigrant bearded confederate in his early thirties (see Figure 1) approaches the selected passenger and asks for help following a script. Before initiating the interaction, he waits until the selected passenger stops walking and stays standing somewhere on the platform. The passenger stands typically in a position that is perpendicular, on the frontal or coronal plane, to the rails. The confederate, carrying a portable metro map, approaches walking parallel to the rails and stops when the tip of his shoe is at a rough 10 cm distance from the passenger's. The result is a side-by-side arrangement in which confederate and participant form an approximate right angle on the frontal plane.

[Insert Figure 1 here]

Figure 1: confederate's appearance

The script divides the interaction in two stages involving different verbal contents and body postures. The first stage consists in locating items on a portable map with confederate and passenger side-by-side (see Supplemental Material, Method, Figure i). In the second stage, the confederate shifts to a close face-to-face position, asking the passenger to estimate the duration of the trip ahead of him (SM, Method, Figure ii). After the passenger's reply, the confederate laments being late for an important appointment, emphasizes that he needs to contact the person

he has to meet, but regrets that his cell phone has run out of battery. After the passenger's reply to this indirect request, a researcher intervenes to unmask the plot and inform the passenger that the interaction has been recorded. The researcher hands the passenger a leaflet with a statement of the research objectives and contact information, provides a short oral explanation and requests consent to process the collected images and audio files. The passenger is then invited to answer to a short questionnaire.

Experimental treatment

Description. The experimental manipulation intervenes in the opening lines of the script, as follows:

Bonjour, excusez-moi de vous déranger. Je suis un peu perdu. J'aurais voulu savoir où se trouvait la station X sur le plan.

Hello. Excuse me, I'm a bit lost. I'd like to know where station X is located on the map [station X is the station where the interaction takes place].

(Passenger replies)

Merci. Et moi, je vais à Censier-Daubenton...

Thank you. And I'm going to [station] Censier-Daubenton...

At that point, in the control condition the confederate adds

...à côté de l'Université Sorbonne Nouvelle.

...next to Université Sorbonne Nouvelle.

In the experimental condition, he completes the sentence by saying instead

...à côté de la Grande mosquée de Paris.

...next to Grande mosquée de Paris.

Rationale. In contrast to areas such as education, where the treatment usually involves an intervention and the control its absence, experiments investigating discrimination never use a genuine control condition. The treatment is not truly compared to a nonintervention, but to another treatment. This baseline treatment will never be “the average population”, that statistical abstraction, but always an embodied, concrete type of person characterized by a gender, an age, and so on. This is obviously the case when discrimination is studied, as in the present research, in the context of live face-to-face interactions with confederates. But even when the treatment is such a “thin” stimulus as a fake CV in a correspondence study, fake applicants always have a gender, an age, an address, a nationality, and possibly many other socially relevant characteristics.

In creating the contrast between the experimental treatments, we started with two constraints. First, the relevant type of person had to be a French man in his thirties. Second, the Muslim condition had to involve an outer religious sign, in analogy with the previously conducted experiment in which a female confederate had signaled Islamic faith with a hijab (Aranguren et al., 2021). There is no perfect male equivalent of Muslim women’s headscarf, because there is no publicly recognizable item of clothing that orthodox Muslim men *must* use in all public situations. But the outer sign that comes closest to the hijab in terms of intensity of use in public is a voluminous beard. What distinguishes the beard from the hijab, however, is its public ambiguity. Whereas the hijab is almost universally recognized in France as a sign of Islamic affiliation, the beard might mean many other things depending on the context.

Having set a French man in his thirties as the relevant type of person and the beard as an outer (yet ambiguous) sign of Islamic affiliation, the next task consisted in constructing a credible contrast between a baseline non-Muslim bearded French man in his thirties, on the one hand, and a Muslim bearded French man in his thirties, on the other. The common practice in discrimination research conducted in Western countries is to set some fraction of the White

middle-class as the baseline. After consulting the catalogue of available public types of men, we concluded that the “hipster”, as a representative of a young and trendy fraction of the White middle-class (Arsel & Thompson, 2011; Cronin et al., 2014; le Grand, 2020), supplied a valid type of baseline bearded man with which the Muslim bearded man could be compared.

To help passengers associate the beard to either type, the confederate mentioned the university in one condition but the mosque in the other, and he wore a (fashionable) fedora hat only in the baseline non-Muslim condition. At the time of the experiment, the fedora hat happened to be common among hipster young men of the Parisian White middle-class, and so we decided to include it in the non-Muslim condition to further help passengers disambiguate the confederate’s social type. Mentioning Université Sorbonne Nouvelle as a landmark can be considered a valid baseline destination in the particular context of the Paris metro and the French educational system for the following reasons. First, as can be concluded from Table 1 (see below), across the surveyed stations the median level of educational achievement is located between 3 and 5 years of higher education, indicating familiarity with a university setting. Second, high school education in France is stratified into two broad classes of institutions: elite schools with very demanding admission conditions that usually require special preparation (e.g. École Normale Supérieure, Sciences Po) and what the French call “la fac”, that is nonelite regular universities such as Sorbonne Nouvelle where admission is less selective. In short, for the average passenger Sorbonne Nouvelle can be reasonably expected to represent a nonelite, familiar university setting.

It goes without saying that, logically, there is no mutual exclusion between hipster and orthodox Muslim. Empirically, however, at the time of conducting the experiment to all appearances in Paris most hipsters were not orthodox Muslims and vice-versa.

Measurements and outcome variables

The demographic variables measured with the questionnaire are age, educational achievement, income, and religion. The outcome variables analyzed pertain only to the face-to-face part of the interaction between confederate and passenger. The outcomes measuring interpersonal warmth or involvement are distance in cm, proportion of time spent speaking (speech rate), proportion of time spent establishing eye contact (gaze rate) and incidence of shoulder shrugs in reply to the confederate's requests or complaints (for more details, see Supplemental Material, Method, Measurements).

Data preparation

The continuous covariate age was dichotomized by setting the threshold between the young and the middle-aged at 30 years of age. Only 16 passengers were older than 60, i.e. above the upper limit of what is usually considered the middle-age range, precluding a separate analysis for the old for lack of a sufficiently large sample. All reported models were performed either by pooling these 16 older participants with the 139 middle-aged ones or by excluding them from the dataset. As the qualitative results do not differ, in what follows we report the pooled, more complete models.

Statistical analyses (for details, see Supplemental Material, Method)

The five models that we estimated are analogous in logic to traditional ANOVAs but computed with Bayesian inference in the context of the Generalized Linear Model (Gelman et al., 2013; Kruschke, 2015).

Note on reporting style. It is inherent to Bayesian inference to describe the output from each model, namely parameter values, as intervals (more precisely, as posterior probability distributions) instead of point estimates. The type of interval considered here is known as the “central posterior interval” (Gelman et al., 2013), and provides the equivalent of a two-tailed test. Unless otherwise indicated, the default alpha level of all the reported central posterior

intervals is the standard 5%. In the Results section, the quantities reported between square brackets make reference to the 95% central posterior interval of the parameter under consideration. Sample sizes are not provided in separate tables but incorporated to the Figures plotting the estimated parameters, facilitating the reader's access to the sample size underlying the estimation of every single reported parameter. While Figures 2-4 graphically report all the relevant main, simple, and interaction effects, only those that credibly differ from zero are verbally highlighted in the text.

Results

Descriptive statistics

The confederate interacted with a total of 309 passengers, completing the script with 298 of them. Of those who completed the script, 270 also gave consent to process the collected identifying information. Figures 1 and 2 decompose these aggregate numbers for males and females in the control and mosque conditions. Of the 270 passengers who completed the script and gave consent, 246 stated their age. On the basis of the collected videos, the age of the remaining 24 passengers was assessed as the mean of the ratings provided by three independent coders. The audiovisual recordings of six consenting passengers were accidentally lost. Only 111 passengers completed the questionnaire.

[Insert Figure 2 about here]

[Insert Figure 3 about here]

Figures 2 and 3. Passengers who completed the script and gave consent to have their identifying data processed in the control and mosque conditions.

Table 1 describes the population of each station during the time of the day at which sampling proceeded.

covariate/ station	Michel Bizot	Boucicaut	Notre- Dame-de- Lorette	Pyrénées	Ranelagh	Riquet
-----------------------	-----------------	-----------	-------------------------------	----------	----------	--------

age (median)	31	29	30	32	35	32
education (median)	4 years of higher education (HE)	3 years HE	5 years HE	4 years HE	5 years HE	3 years HE
monthly income (median)	2000€- 2500€	1800€- 2000€	2000€- 2500€	1800€- 2000€	1800€- 2000€	1500€- 1800€
% muslims	12 %	12 %	4 %	10 %	5 %	19 %

Table 1: Average characteristics of passengers by station. Age, education and income report medians. Education and income medians describe only participants aged 25 or more, to better portray the tendencies among those who are most likely to have finished their schooling and to be/have been in the workforce. The source data aggregate the questionnaires filled in the present experiment (n=111) performed in Fall 2018 with those collected in a previous experiment carried out in Spring 2018 (n=290, Aranguren et al., 2021) in the same stations and using the same sampling technique.

Grand means. To facilitate an appreciation of the magnitude of the reported effects, Table 1 provides, for each model, the estimated mean of means or “grand mean” (see Supplemental Material, Method, All estimated means for more details).

Outcome	Grand mean (95% central posterior interval)
helping behavior	[17%, 32%]
shoulder shrug	[1%, 6%]
interpersonal distance	[107 cm, 112 cm]
speaking time (speech rate)	[30%, 33%]
eye contact (gaze rate)	[55%, 62%]

Table 1: Outcomes and grand means

[Insert Figure 4 here]

Figure 4: Helping behavior. On the left margin, the text in bold gives the heading of different groups of parameters. The text in regular font specifies the individual parameters. The quantities on the right margin of the plot specify the number of observations on which the estimation of the corresponding parameter directly relies; the first number refers to the sample size of the mosque group, the second number to that of the control group. The x-axis quantifies the difference between the mosque and the control conditions. Within the plot area, the dashed vertical line in the middle indicates the location of the value 0, which signifies no difference between the control and the mosque conditions. The horizontal segments represent the central 95% posterior intervals of the parameters. The bolder section of the segment corresponds to the central 90% posterior interval and the solid point indicates the median of the distribution. Simple effects are differences between the treatment and the control groups within specific subgroups, e.g. by gender or age. When relevant, to indicate that the effect of the mosque condition operates within a given subgroup, the symbol “@” has been used as a connector. Thus “mosque @male” abbreviates “simple effect of the mosque condition within the subsample made up of male participants”.

[Insert Figure 5 here]

Figure 5: Incidence of shoulder shrug

[Insert Figure 6 here]

Figure 6: Interpersonal distance

[Insert Figure 7 here]

Figure 7: Speaking time

[Insert Figure 8 here]

Figure 8: Eye contact

Main effects. Confirming P1, passengers help less often the confederate in the mosque condition, [-27%, -6%] (parameter “mosque”, see Figure 4). Equally supporting P1, passengers show lower involvement in the mosque condition by shrugging more often [2%, 15%] (Figure 5) and by speaking less [-6%, -1%] (Figure 7). Of note, the mosque-to-control ratio of the incidence of helping is [0.33, 0.85] with a median at 0.54, roughly indicating that for every two passengers who offered assistance in the control condition, only one did so when the confederate mentioned the mosque.

Regardless of the experimental condition, compared to the young the middle-aged offer assistance more seldomly [-32%, -9%] (parameter “middle-aged”, see Figure 4) and show less involvement by shrugging more often [1%, 11%] (Figure 5) and by keeping larger interpersonal distances [1cm, 10cm] (Figure 4).

Simple effects. Supporting P3b, the effects of the experimental manipulation are sufficiently strong among the young to yield credible simple effects on helping, shrugging and speaking time, whereas the corresponding simple effects among the middle-aged are not credible. In the mosque condition, the young helped less often [-41%, -4%] (parameter “mosque @young”, Figure 4), and showed less involvement by shrugging more frequently [1%, 15%] (Figure 5) and by holding the floor for shorter [-10%, -1%] (Figure 7). In contrast, none of these simple effects are credible when the sub-sample in focus is the group of the middle-aged (parameter “mosque @middle-aged”, Figures 4, 5 and 7).

Although the data do not support the expectation of a condition * sex according to which men would show less involvement but women more of it, the presence of a simple effect of negative sign

among men not paralleled by a corresponding simple effect among women provides weak evidence in that direction. Thus, the simple effects of the mosque condition on the probability of shrugging are credible among male passengers [2%, 17%] (parameter “mosque @male”, Figure 5), but not among female passengers (“mosque @female”, Figure 5). Unexpectedly, the same pattern arises from the analysis of helping behavior, as men’s probability of offering assistance is credibly lower in the mosque condition [-38%, -7%] (parameter “mosque @male”, Figure 4), whereas women’s is not (“mosque @female”, Figure 4).

Post hoc models. In order to check that the detected effects were not masking important differences in effect across stations, each outcome was subjected to a single “varying intercepts, varying slopes” (Gelman & Hill, 2007) model allowing the effect of the mosque condition to vary by station. With the sole exception of eye contact, with respect to which a strong decrease at station Ranelagh clearly deviates from the average, no remarkable by-station differences arise from any of the other models (see Supplemental Material, Results, By-station effects of the mosque condition).

Discussion

We conducted a field experiment in the Paris metro to examine the differences in behavior that passengers show when they are approached by a non-immigrant bearded man who gives indirect cues of being Muslim. As predicted by P1, averaging over gender and age groups passengers helped less often (more specifically: half as often), and expressed lower involvement (or interpersonal warmth) by shrugging more and speaking less in the mosque condition. Although the data did not confirm the expectation of a condition * sex interaction, the negative effect of the mosque condition on shrugging (an index of negative involvement) was present among men but not among women, and unexpectedly the same pattern emerged with respect to helping behavior as well. Contradicting

P3a, and confirming P3b, all the credible main effects of the mosque manipulation turned out to be corresponded by equally credible simple effects among the young but not among the middle-aged.

A possible limitation of the present experiment is that the baseline or control condition was embodied by that particular type of White middle-class man known as the hipster. It could be objected that the reported differences in treatment may be describing better a privilege accorded to hipsters than a penalty imposed on Muslims. The objection would be compelling in the presence of unequivocal evidence indicating that the hipster, as a social type, commands particular respect or admiration. But the available studies suggest rather that the hipster may be as much the object of considerable denigration, both for being dismissive of others' tastes and for a superficial concern to appear "cool" (Arsel & Thompson, 2011; Cronin et al., 2014; le Grand, 2020).

Muslim men are the target of strong anti-Muslimism

If the present results are considered together with those of the experiment that we previously performed on the effects of the hijab (Aranguren et al., 2021), it appears that, at least in the context of a polite interaction between strangers in Paris, Muslim men and women who give signs of their religious affiliation are not exposed to anti-Muslimism to the same degree. Whereas the hijab was met with a mix of positive and negative behaviors, authored respectively by female and male passengers, the bearded confederate of the present experiment encountered exclusively negative responses when he mentioned the mosque.

The difference is all the more striking that the hijab unequivocally signals that its carrier is a follower of Islam (and presumably a conservative one), whereas the outer cues provided by the bearded confederate were ambiguous. The confederate's beard qualified equally well as an indication of orthodox Islam and as a signal of hipster fashion. Mentioning the Grande mosquée as a landmark could be taken as evidence that the confederate was going to attend a prayer there, but given that almost nobody in Paris ignores the location of one of the city's touristic attractions, it

could also be seen as a convenient choice for facilitating the search. If at least some participants in the mosque condition adopted these alternative interpretations, the reported effects underestimate the decrease in cooperation and interpersonal warmth that is to be expected from a non-Muslim when he or she interacts with a bearded man *effectively* perceived to be a Muslim, and not only *possibly* as in the present experiment.

One plausible reason why non-Muslims respond so differently to hijab-wearing women, on the one hand, and to bearded Muslim men, on the other, has to do with the common images or stereotypes through which the former and the latter are usually perceived in France and in other European countries. To account for the apparently contradictory reasoning of European judges in court cases, researchers in the area of law have suggested that hijab-wearing women are commonly seen through the lens of two opposite stereotypes, namely those of “aggressor” and “victim” (Edmunds, 2012; Evans, 2006). In light of the aggressor stereotype, the covered woman appears as a dangerous proselytiser of fundamentalist views, an aggressor, that is, to Western values, democracy, gender equality, and so on. In light of the victim stereotype, in contrast, she appears as passive, unable to help herself, needing protection from violent male relatives, a victim, that is, of Muslim men and more generally of a religion oppressive of women. Now, the turning of aggressors into victims works in the case of hijab-wearing women because of their gender identity. The stereotype of the victim disidentifies women from Muslims by picturing Islam as a male religion, thus limiting the scope of the aggressor stereotype to Muslim men. But these, in turn, do not benefit from any duality of stereotypes to temper the negative views that non Muslims hold of them. The fact that hijab-wearing women may be pictured either as aggressors or victims, but bearded Muslim men only as aggressors, might explain why the former elicit a mix of positive and negative behaviors (among women and men, respectively), whereas the responses that the latter motivate turn out to be univocally negative (both among women and men).

As could be expected from the stereotype that women are warm and friendly but men aggressive and potentially dangerous, overall passengers helped more often and were more involved

in interaction with the female confederate in the hijab experiment than with the male confederate in the above reported mosque experiment. But it is important to note that the effect of seeing the confederate as a Muslim (vs. non-Muslim) depended not only on the gender identity of the confederate, but more specifically on the gender composition of the interacting dyad. When both confederate and passenger were women, the Muslim condition elicited more positive behavior. In all the other cases, the effect was negative.

Explaining why the young are more cooperative and friendly but also more likely to discriminate

Although the moderating effect of age confirmed prediction P3b, the results contradict the causal story from which the expectation was derived. We predicted a weakening of the effect of the mosque condition among the middle-aged drawing on the hypothesis that with the years people become more agreeable (Allemand et al., 2008; Donnellan & Lucas, 2008; McCrae et al., 1999; Roberts et al., 2006; Soto et al., 2011; Srivastava et al., 2003) and trusting (Bailey & Leon, 2019; Li & Fung, 2013) and/or avid for social approval (Köötts-Ausmees et al., 2020; Soubelet & Salthouse, 2011), dispositions that appear to be negatively correlated with prejudice (Batson et al., 1978; Dhont & van Hiel, 2011; Ekehammar & Akrami, 2003). If the middle-aged were more agreeable and trusting than the young, we would have expected the baseline levels of involvement and helping behaviors to be higher among the middle-aged. But it is exactly the opposite that the data indicate: regardless of the experimental condition, the middle-aged offer assistance less often and show lower interpersonal involvement. The middle-aged discriminate less not because, but rather in spite of being less agreeable and trusting, and/or less concerned with social approval.

Thus far we have only considered the “horizontal” or “love-hate” dimension, assuming that a positive/negative evaluation of the interaction partner precipitates approach/avoidance, and consequently expecting the negatively evaluated confederate in the mosque condition to elicit lower

levels of helping and involvement. A way to solve the puzzle is to bring into the analysis the “vertical” (Hall et al., 2005) or “dominance-submission” (Burgoon & Hale, 1984) dimension of human relations. Verticality refers to position in a low-to-high continuum, an idea underlying the somewhat distinct constructs of “power”, “status” and “dominance” (Hall et al., 2005).

Operationally, in this study as in others the young are those under 30 years of age. At the time of performing the experiment, the confederate was in his early thirties, meaning that the young were also participants who happened to be younger than the confederate. Now, age is an obvious source of status. On the vertical dimension, being older than person X implies being *senior to X*, i.e. possessing a quality with the potential to command respect or privilege from X. In this sense, the confederate was senior to, of higher age status than, the participants in the young group.

Now, while status involves possession of a valued quality, in the context of a social interaction higher rank does not necessarily translate into interpersonal dominance, i.e. successful control of others or their resources (Burgoon et al., 1998). In other words, whereas higher status is a potential source of dominance (and its complement, deference), the effective occurrence of status-based dominance/deference depends on a number of moderators. We propose to regard the difference between the control and the mosque condition as one moderator of the status-to-dominance translation when it comes to deferring to the confederate’s higher age rank.

With these elements at hand, it is possible to explain why, although they were more helpful and friendlier overall, the young were nonetheless more likely to discriminate in the mosque condition than were the middle-aged. In this “vertical” perspective, regardless of the experimental condition the young deferred to the higher age status (to the seniority) of the confederate, whereas this influence was absent among the middle-aged, for whom the confederate was either an equal or a junior. At least in part, it was *in deference to* the confederate’s seniority that the young helped more often and showed higher interpersonal involvement, compared to the middle-aged. That is, to some extent the interactions between the confederate and the young (vs. the middle-aged) involved the expression of age-based dominance/deference.

The experimental condition, in turn, may have acted as a moderator of age-based dominance, leading the young to defer less to the higher age status of the confederate in the mosque condition than in the control condition. That is, the association with Islam or Muslims may have weakened the expression of age-based deference, as though the senior Muslim had been less entitled than the senior non-Muslim to have his higher age status recognized. This weakening effect, in turn, can be analyzed in “horizontal” or “vertical” terms. In the horizontal reading, the fact that the confederate was of higher age status but evaluated negatively as a Muslim affectively may have inhibited the expression of deference. In the vertical interpretation, the confederate’s overall status resulted from a combination of various positions, including that in the age continuum, where older age meant higher rank, and that in the Muslim vs. non Muslim dichotomy, where Muslim meant lower status. In this perspective, the weaker expression of deference in the mosque condition among the young may have reflected the acknowledgement of a smaller overall status gap vis-à-vis the confederate.

Conclusion

The above reported field experiment clearly indicates that bearded Muslim men are the target of more or less subtle expressions of anti-Muslim racism in everyday interactions. This result was reached using a minimal stimulus: a non-immigrant bearded confederate, speaking native French and wearing standard urban clothing hinted that he was Muslim by mentioning a mosque as a landmark of his destination. There is no reason to doubt that the same effects would be observed if the bearded Muslim man signaled an extra-European origin, or spoke with a foreign accent, or wore religious attire, or explicitly said that he was going to a prayer, or all or some of these in combination. In contrast, considering that most Muslim men residing in France do possess some of these characteristics, there is every reason to think that the reported effects underestimate the incidence and strength of the forms of anti-Muslim behavior that they actually encounter in

everyday life. Our behavioral experiment corroborates Muslim men's reported experience of discrimination.

References

- Adida, C. L., Laitin, D. D., & Valfort, M.-A. (2010). Identifying barriers to Muslim integration in France. *Proceedings of the National Academy of Sciences*, *107*(52), 22384–22390. <https://doi.org/10.1073/pnas.1015550107>
- Adida, C. L., Laitin, D. D., & Valfort, M.-A. (2016). *Why Muslim Integration Fails in Christian-Heritage Societies*. Harvard University Press.
- Aidenberger, A., & Doehne, M. (2021). Unveiling everyday discrimination. Two field experiments on discrimination against religious minorities in day-to-day interactions. *British Journal of Sociology*, *00*, 1–19. <https://doi.org/10.1111/1468-4446.12811>
- Allemand, M., Zimprich, D., & Hendriks, A. A. J. (2008). Age differences in five personality domains across the life span. *Developmental Psychology*, *44*(3), 758–770. <https://doi.org/10.1037/0012-1649.44.3.758>
- Aranguren, M., Madrisotti, F., Durmaz-Martins, E., Gerger, G., Wittmann, L., & Méhu, M. (2021). Responses to the islamic headscarf in everyday interactions depend on sex and locale: A field experiment in the metros of Brussels, Paris, and Vienna on helping and involvement behaviors *PLoS ONE (forthcoming)*. <https://hal.archives-ouvertes.fr/hal-03107103>
- Arsel, Z., & Thompson, C. J. (2011). Demythologizing Consumption Practices: How Consumers Protect Their Field-Dependent Identity Investments from Devaluing Marketplace Myths. *Journal of Consumer Research*, *37*(5), 791–806. <https://doi.org/10.1086/656389>
- Bailey, P. E., & Leon, T. (2019). A systematic review and meta-analysis of age-related differences in trust. *Psychology and Aging*, *34*(5), 674–685. <https://doi.org/10.1037/pag0000368>
- Batson, C. D., Naifeh, S. J., & Pate, S. (1978). Social Desirability, Religious Orientation, and Racial Prejudice. *Journal for the Scientific Study of Religion*, *17*(1), 31–41. <https://doi.org/10.2307/1385425>
- Bond, M. H. (1972). Effect of an impression set on subsequent behavior. *Journal of Personality and Social Psychology*, *24*, 301–305. <https://doi.org/doi.org/10.1037/h0033716>
- Brinbaum, Y., Safi, M., & Simon, P. (2018). Discrimination in France: Between perception and experience. In C. Beauchemin, C. Hamel, & P. Simon (Eds.), *In Trajectories and Origins: Survey on the Diversity of the French Population* (pp. 195–222). Springer.
- Burgoon, J. K., & Hale, J. L. (1984). The fundamental topoi of relational communication. *Communication Monographs*, *51*, 193–214.
- Burgoon, J. K., Johnson, M. L., & Koch, P. T. (1998). The nature and measurement of interpersonal dominance. *Communication Monographs*, *65*(4), 308–335. <https://doi.org/10.1080/03637759809376456>
- Choi, D. D., Poertner, M., & Sambanis, N. (2020). Linguistic assimilation does not reduce discrimination against immigrants: Evidence from Germany. *Journal of Experimental Political Science*, *1*(12). <https://doi.org/10.1017/XPS.2020.20>
- CNCDH. (2019). *La lutte contre le racisme, l'antisémitisme et la xénophobie: Rapport 2018* (p.

345). Commission Nationale Consultative des Droits de l'Homme; La documentation française. https://www.cncdh.fr/sites/default/files/23072019_version_corrige_rapport_racisme.pdf

Cornelis, I., Hiel, A. V., Roets, A., & Kossowska, M. (2009). Age Differences in Conservatism: Evidence on the Mediating Effects of Personality and Cognitive Style. *Journal of Personality*, 77(1), 51–88. <https://doi.org/10.1111/j.1467-6494.2008.00538.x>

Coutts, S. M., Schneider, F. W., & Montgomery, S. (1980). An investigation of the arousal model of interpersonal intimacy. *Journal of Experimental Social Psychology*, 16, 545–561.

Cronin, J. M., McCarthy, M. B., & Collins, A. M. (2014). Covert distinction: How hipsters practice food-based resistance strategies in the production of identity. *Consumption Markets & Culture*, 17(1), 2–28. <https://doi.org/10.1080/10253866.2012.678785>

Dhont, K., & van Hiel, A. (2011). Direct contact and authoritarianism as moderators between extended contact and reduced prejudice: Lower threat and greater trust as mediators. *Group Processes & Intergroup Relations*, 14(2), 223–237. <https://doi.org/10.1177/1368430210391121>

Diekmann, A., Jann, B., & Näf, M. (2014). Wie fremdenfeindlich ist die Schweiz? Fünf Feldexperimente über prosoziales Verhalten und die Diskriminierung von Ausländer in der Stadt Zürich und in der Deutschschweiz [How xenophobic is Switzerland? Five field experiments on prosocial behavior and discrimination of foreigners in Zürich city and in German-speaking Switzerland]. *Soziale Welt*, 65, 185–199.

Di Stasio, V. D., Lancee, B., Veit, S., & Yemane, R. (2021). Muslim by default or religious discrimination? Results from a cross-national field experiment on hiring discrimination. *Journal of Ethnic and Migration Studies*, 47(6), 1305–1326. <https://doi.org/10.1080/1369183X.2019.1622826>

Donnellan, M. B., & Lucas, R. E. (2008). Age differences in the big five across the life span: Evidence from two national samples. - PsycNET. *Psychology and Aging*, 23(3), 558–566. <https://doi.org/10.1037/a0012897>

Eberhard, M., & Simon, P. (2016). *Expérience et perception des discrimination en Ile-de-France* [Survey report]. Observatoire régional des discriminations (ORDIS).

Edmunds, J. (2012). The limits of post-national citizenship: European Muslims, human rights and the hijab. *Ethnic and Racial Studies*, 35(7), 1181–1199. <https://doi.org/10.1080/01419870.2011.593640>

Ekehammar, B., & Akrami, N. (2003). The relation between personality and prejudice: A variable- and a person-centred approach. *European Journal of Personality*, 17(6), 449–464. <https://doi.org/10.1002/per.494>

Ekehammar, B., & Akrami, N. (2007). Personality and Prejudice: From Big Five Personality Factors to Facets. *Journal of Personality*, 75(5), 899–926. <https://doi.org/10.1111/j.1467-6494.2007.00460.x>

European Commission. (2019). *Special Eurobarometer 493. Report on Discrimination in the European Union*. European Union; ec.europa.eu.

European Values Study. (2017). *European Values Study 2017: Integrated Dataset (EVS 2017)*. GESIS Data Archive. <https://doi.org/10.4232/1.13560>

Evans, C. (2006). The islamic scarf in the European Court of Human Rights. *Melbourne Journal of International Law*, 7(1), 52–74.

Federico, C. M., & Sidanius, J. (2002). Sophistication and the Antecedents of Whites' Racial Policy Attitudes: Racism, Ideology, and Affirmative Action in America. *Public Opinion Quarterly*, 66(2), 145–176. <https://doi.org/10.1086/339848>

Ford, R. (2008). Is racial prejudice declining in Britain? *The British Journal of Sociology*, 59(4),

609–636. <https://doi.org/10.1111/j.1468-4446.2008.00212.x>

FRA. (2017). *Second European Union Minorities and Discrimination Survey. Muslims-Selected findings*. European Union Agency for Fundamental Rights.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis* (3rd ed.). CRC Press.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. Cambridge University Press.

Gonsalkorale, K., Sherman, J. W., & Klauer, K. C. (2009). Aging and prejudice: Diminished regulation of automatic race bias among older adults. *Journal of Experimental Social Psychology*, 45(2), 410–414. <https://doi.org/10.1016/j.jesp.2008.11.004>

Hall, J. A., Coats, E. J., & Smith Lebeau, L. (2005). Nonverbal behavior and the vertical dimension of social relations: A meta-analysis. *Psychological Bulletin*, 131(6), 898–924.

Halliday, F. (1999). “Islamophobia” reconsidered. *Ethnic and Racial Studies*, 22(5), 892–902. <https://doi.org/10.1080/014198799329305>

Harteveld, E., & Ivarsflaten, E. (2018). Why Women Avoid the Radical Right: Internalized Norms and Party Reputations. *British Journal of Political Science*, 48(2), 369–384. <https://doi.org/10.1017/S0007123415000745>

Hebl, M. R., Bigazzi-Foster, J., Mannix, L. M., & Dovidio, J. F. (2002). Formal and interpersonal discrimination: A field study of bias toward homosexual applicants. *Personality and Social Psychology Bulletin*, 28(6), 815–825. <https://doi.org/10.1177/0146167202289010>

Ickes, W., Patterson, M. L., Rajecki, D. W., & Tanford, S. (1982). Behavioral and cognitive consequences of reciprocal versus compensatory responses to preinteraction expectancies. *Social Cognition*, 1(2), 160–190. <https://doi.org/10.1521/soco.1982.1.2.160>

King, E. B., & Ahmad, A. S. (2010). An experimental field study of interpersonal discrimination toward muslim job applicants. *Personnel Psychology*, 63, 881–906. <https://doi.org/10.1111/j.1744-6570.2010.01199.x>

Koopmans, R., Veit, S., & Yemane, R. (2019). Taste or statistics? A correspondence study of ethnic, racial and religious labour market discrimination in Germany. *Ethnic and Racial Studies*, 42(16), 233–252. <https://doi.org/10.1080/01419870.2019.1654114>

Kööts-Ausmees, L., Kandler, C., Realo, A., Allik, J., Borkenau, P., Hřebíčková, M., & Möttus, R. (2020). *Age Differences in Personality Traits and Social Desirability: A Multi-Rater Multi-Sample Study* [Pre-print]. psyarxiv.com. 10.31234/osf.io/bmv9r

Kruschke, J. K. (2015). *Doing Bayesian data analysis*. Academic Press.

le Grand, E. (2020). Representing the middle-class ‘hipster’: Emerging modes of distinction, generational oppositions and gentrification. *European Journal of Cultural Studies*, 23(2), 184–200. <https://doi.org/10.1177/1367549418772168>

Lemasson, A., Toutain, M., Madrisotti, F., & Aranguren, M. (2021). *Effects of the Islamic headscarf on vocal arousal and intimacy: A field experiment in the Paris metro* [Pre-print]. hal.archives-ouvertes.fr. <https://hal.archives-ouvertes.fr/hal-03140246v2>

Li, T., & Fung, H. H. (2013). Age Differences in Trust: An Investigation Across 38 Countries. *The Journals of Gerontology: Series B*, 68(3), 347–355. <https://doi.org/10.1093/geronb/gbs072>

Littleford, L. N., Wright, M. O., & Sayoc-Parial, M. (2005). White Students’ Intergroup Anxiety During Same-Race and Interracial Interactions: A Multimethod Approach. *Basic and Applied Social Psychology*, 27(1), 85–94. https://doi.org/10.1207/s15324834basp2701_9

- McCrae, R. R., Costa Jr., P. T., de Lima, M. P., Simoes, A., Ostendorf, F., Angleitner, A., Marušić, I., Bratko, D., Caprara, G., Barbaranelli, C., Chae, J. H., & Piedmont, R. L. (1999). Age differences in personality across the adult life span: Parallels in five cultures. - *PsycNET. Developmental Psychology*, *35*(2), 466–477. <https://doi.org/10.1037/0012-1649.35.2.466>
- Patterson, M. L. (1982). A sequential functional model of nonverbal exchange. *Psychological Review*, *89*(3), 231–249. <https://doi.org/10.1037/0033-295X.89.3.231>
- Pew Research Center. (2006, August 17). *The French-Muslim Connection*. Pew Research Center. <https://www.pewresearch.org/2006/08/17/the-frenchmuslim-connection/>
- Pew Research Center. (2017, November 29). *Europe's Growing Muslim Population*. Pew Research Center's Religion & Public Life Project. <https://www.pewforum.org/2017/11/29/europes-growing-muslim-population/>
- Pierné, G. (2013). Hiring discrimination based on national origin and religious closeness: Results from a field experiment in the Paris area. *IZA Journal of Labor Economics*, *2*(4), 1–15.
- Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin*, *132*(1), 1–25. <https://doi.org/10.1037/0033-2909.132.1.1>
- Soto, C. J., John, O. P., Gosling, S. D., & Potter, J. (2011). Age differences in personality traits from 10 to 65: Big Five domains and facets in a large cross-sectional sample. *Journal of Personality and Social Psychology*, *100*(2), 330–348. <https://doi.org/10.1037/a0021717>
- Soubelet, A., & Salthouse, T. A. (2011). Influence of social desirability on age differences in self-reports of mood and personality. *Journal of Personality*, *79*(4), 741–762. <https://doi.org/10.1111/j.1467-6494.2011.00700.x>
- Srivastava, S., John, O. P., Gosling, S. D., & Potter, J. (2003). Development of personality in early and middle adulthood: Set like plaster or persistent change? *Journal of Personality and Social Psychology*, *84*(5), 1041–1053. <https://doi.org/10.1037/0022-3514.84.5.1041>
- Stewart, B. D., von Hippel, W., & Radvansky, G. A. (2009). Age, Race, and Implicit Prejudice: Using Process Dissociation to Separate the Underlying Components. *Psychological Science*, *20*(2), 164–168. <https://doi.org/10.1111/j.1467-9280.2009.02274.x>
- Truett, K. R. (1993). Age differences in conservatism. *Personality and Individual Differences*, *14*(3), 405–411. [https://doi.org/10.1016/0191-8869\(93\)90309-Q](https://doi.org/10.1016/0191-8869(93)90309-Q)
- Valfort, M.-A. (2020). Anti-Muslim discrimination in France: Evidence from a field experiment. *World Development*, *135*, 105022. <https://doi.org/10.1016/j.worlddev.2020.105022>
- von Hippel, W., Silver, L. A., & Lynch, M. E. (2000). Stereotyping Against Your Will: The Role of Inhibitory Ability in Stereotyping and Prejudice among the Elderly. *Personality and Social Psychology Bulletin*, *26*(5), 523–532. <https://doi.org/10.1177/0146167200267001>

The authors declare no conflict of interest

Figure 1



Figure 2

**male
passengers**

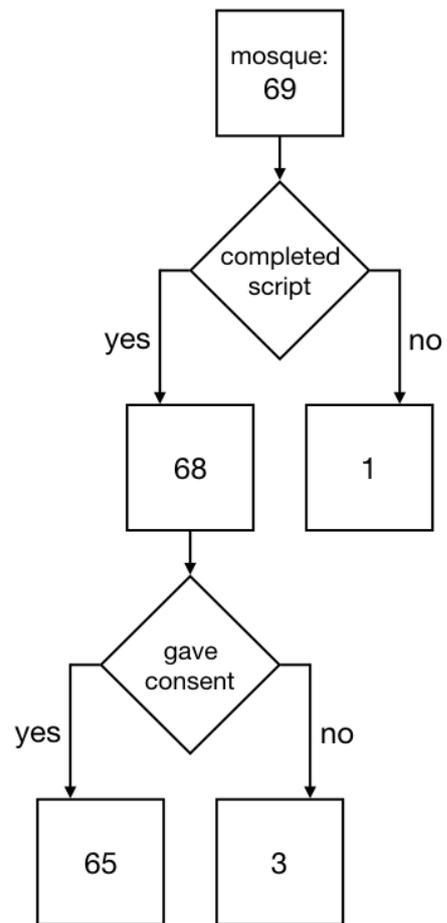
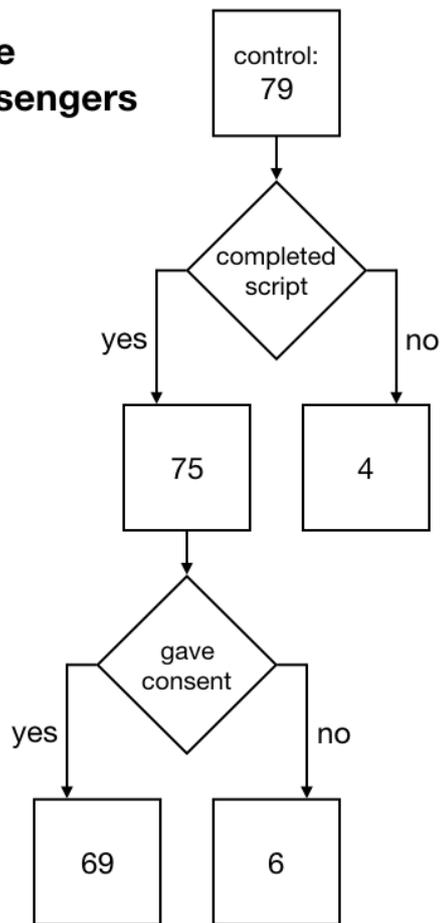


Figure 3

**female
passengers**

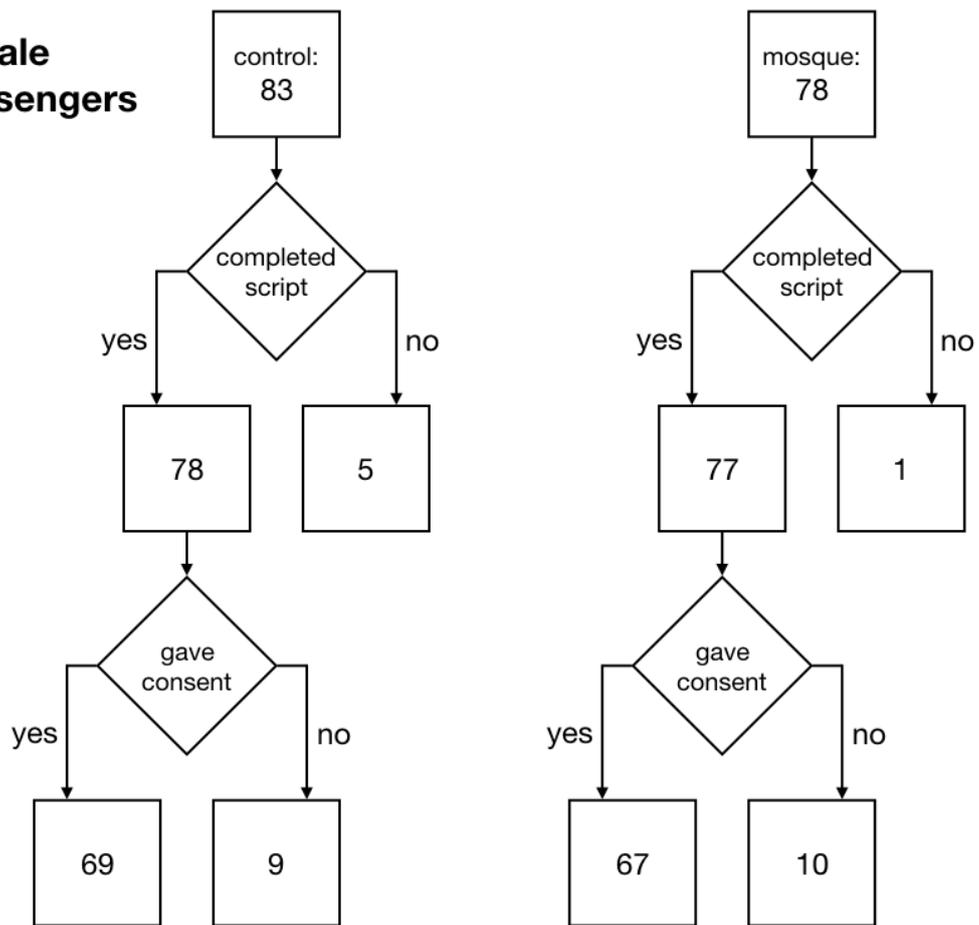


Figure 4

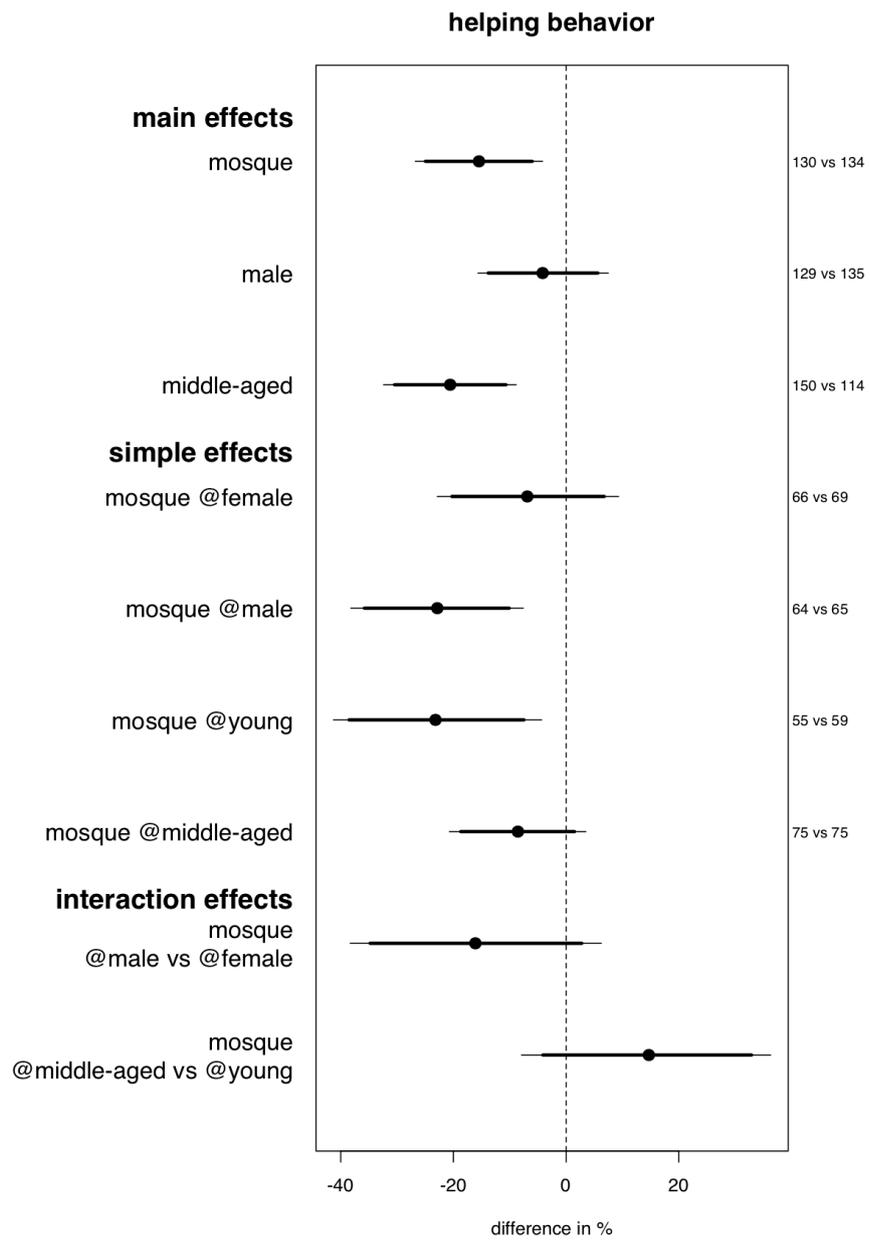


Figure 5

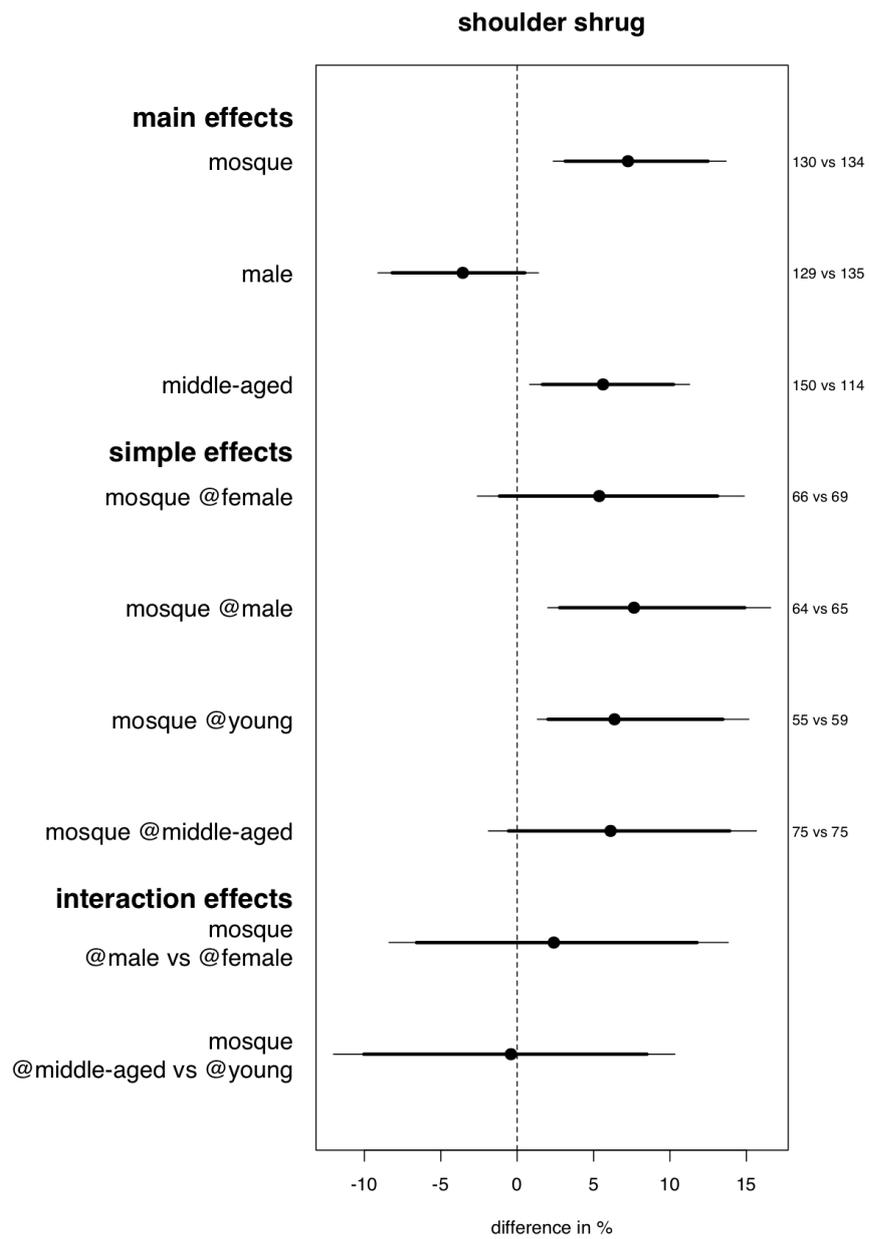


Figure 6

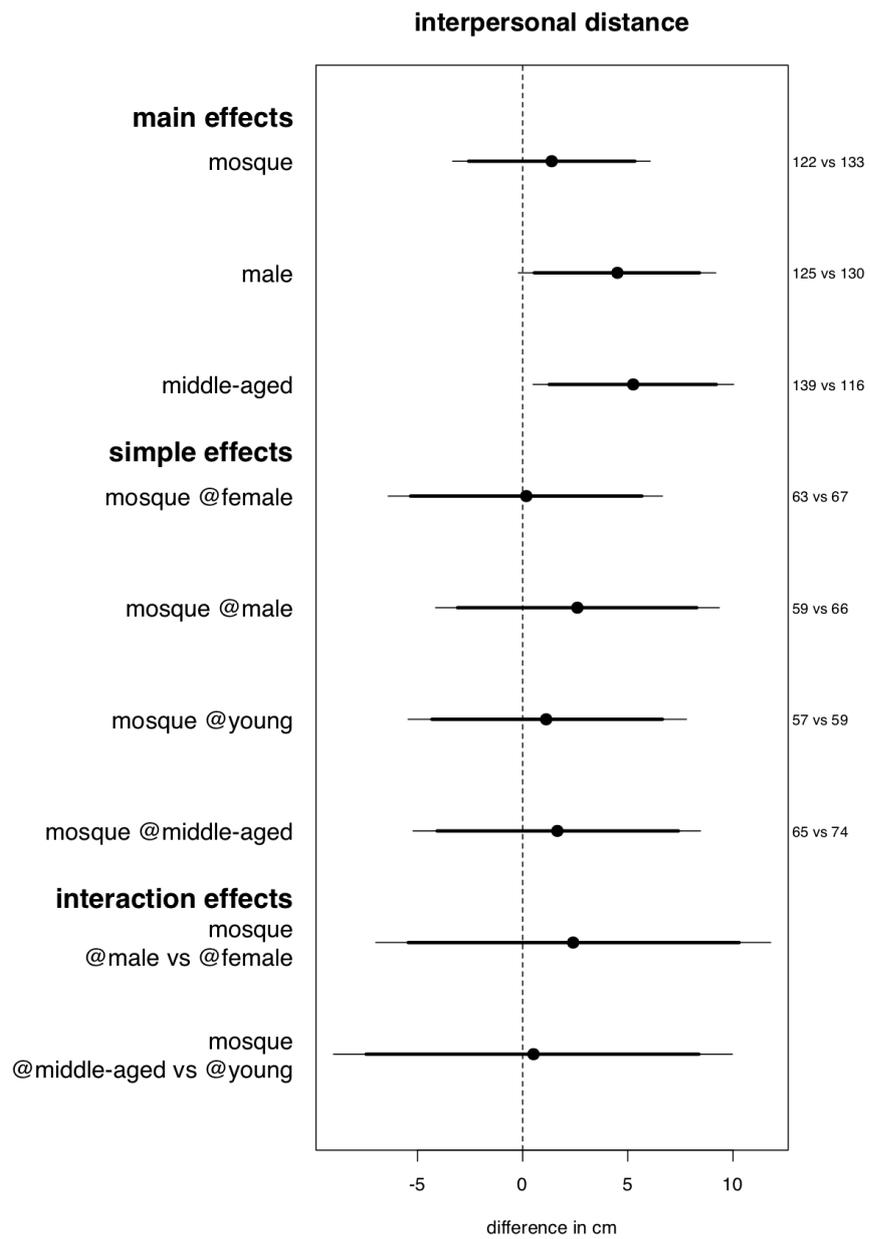


Figure 7

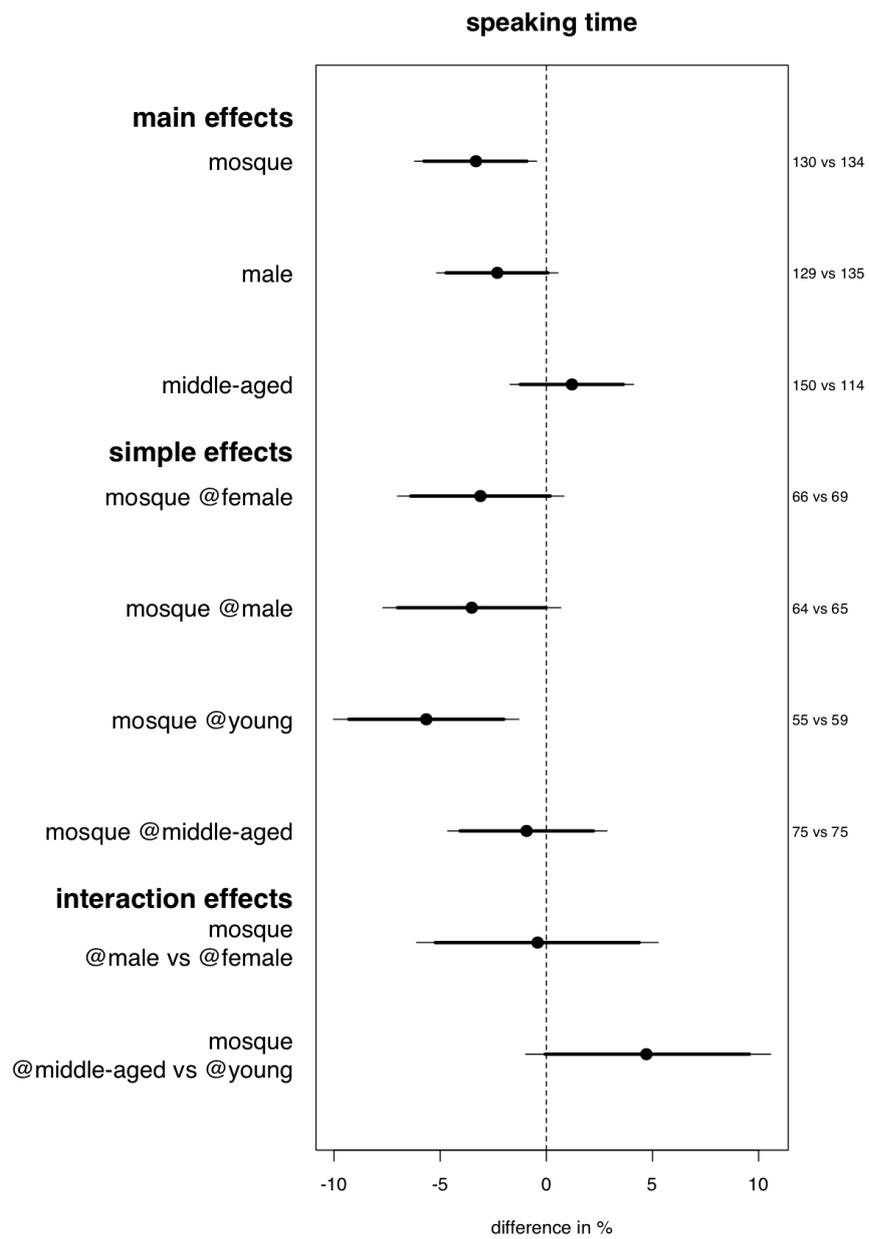
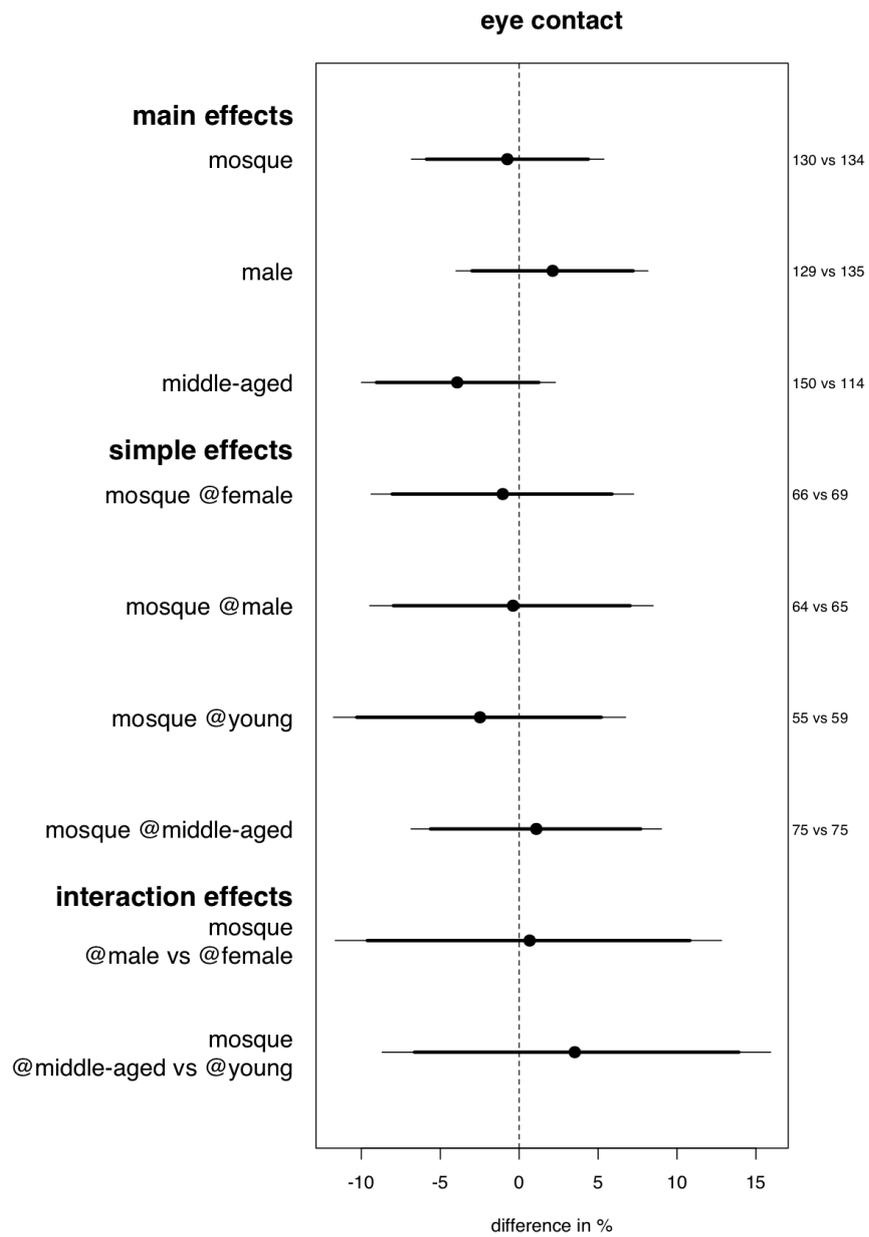


Figure 8



Supplemental Material, Method

Procedure



Figures i (left) and ii (right). In the first stage of the script the confederate (in white sweater, giving the back to the camera) and the passenger adopt a side-by-side postural arrangement to locate items on a portable map (left image). In the second stage of the script, the confederate shifts to a face-to-face position (right image).

Measurements

Helping was measured by considering whether the passenger assisted the confederate after the indirect request for the phone. For distance, speech rate and gaze rate, the observation period

starts as the confederate begins the shift in position and ends with the end of the confederate's indirect request. For helping, the observation period goes from the end of the confederate's indirect request to the ritual closure of the interaction (typically, reciprocal greetings). For shoulder shrug, the observation window covers the entire period from the transition to the face-to-face position to the closure of the exchange. Except for interpersonal distance, all outcome variables were measured from the collected videos and audios using the annotation software ELAN (Sloetjes & Wittenburg, 2008). For each variable a minimal 1/5 portion of the measurements were performed by two coders independently and their level of intercoder agreement quantified using Krippendorff's alpha (Hayes & Krippendorff, 2007). The latter turned out to be above 0.7 in all the cases.

Interpersonal distance (in cm). This outcome was defined as the maximal distance between the confederate's farthest heel, relative to the passenger, and the passenger's farthest heel, relative to the confederate. Operationally, we located the chronologically last annotation of a step, and we used the video frame located in the middle of this annotation to take the measurement. The measuring technique, relying on a screen ruler, consisted in counting the number of pixels comprised between the farthest heels of confederate and passenger, and then counting the number of pixels represented by the length of the confederate's shoe. The outcome variable interpersonal distance is the ratio of the first measure and the second one, multiplied by the actual length in centimeters of the confederate's shoe.

Speaking time (proportion). The outcome variable speech rate is the proportion of the observation period that the passenger spent vocalizing. Every annotation begins with any vocalization, not only articulated speech, on the part of confederate or passenger, and ends when the vocalization ends. Vocalizations may include laughter, hesitations, and fragmented

speech, for example. Speech pauses shorter than 500ms were not counted as pauses and were included in the continuous annotations.

Eye contact (proportion). Eye contact was operationalized as two distinct outcome variables:

i) the gaze rate while speaking, which is the total proportion of time that the passenger spent looking at the confederate while the passenger was speaking, and ii) the gaze rate while listening, which is the total proportion of time that the passenger spent looking at the confederate while the confederate was speaking. Only the gaze rate while listening was used in statistical analyses to avoid overlap with the speech rate, and therefore the possibility that the effect of the mosque condition on one of these outcomes may be mediated by the other.

Further, the gaze rate while listening, given that looking at the speaker is a conventional display of attention or “advertence”, is a better indicator of involvement than the gaze rate while speaking or an overall gaze rate averaging over conversational roles (Fehr & Exline, 1987). Annotations begin when the passenger starts to look at the confederate in the face and end when he or she stops doing so. If the quality of the image was insufficient to understand where exactly the passenger’s eyes were aiming, the orientation of the head was used instead.

Shoulder shrug (dichotomous). In the course of the pilot study, we realized that a considerable proportion of passengers responded to the confederate’s requests or complaints by raising their shoulders with hand palms turned upward, often accompanying the gesture by raised eyebrows and chin, and compressed lips. As a reply, this emblem, in partial or full form, conveys indifference in the face of the confederate’s predicament, and so we decided to consider it in the main experiment as an index of negative involvement. This outcome was measured dichotomously; shrug-yes indicates that during the relevant observation period the passenger shrugged his or her shoulders at least once.

Helping behavior (dichotomous). For each complete interaction, we coded helping as a dichotomy: either the passenger cooperated or not after the confederate finished the last line of the script. The category cooperated-yes includes a variety of possible helping behaviors on the part of the passenger, such as inviting the confederate to dictate an sms, asking the confederate to spell the phone number of the person he had to meet, or simply handing the mobile to him.

Statistical analyses

Each outcome was subjected to a separate model, whose units of analysis were unique participant-confederate interactions denoted i . The chosen distributional form depended on the outcome's mathematical properties.

The continuous and unbounded distance measurements were given a normal distribution with mean \hat{y} and standard deviation σ^2 . Equally continuous but bounded to the interval $[0, 1]$, the speech and gaze rates (i.e. proportions) were given a beta distribution with parameter a equal to the product of the concentration parameter κ and the mean \hat{y} , and parameter b equal to $\kappa*(1-\hat{y})$. Finally, the dichotomous helping and shrug codings were given a Bernoulli distribution equal to the average probability of observing a 1 instead of a 0. More formally,

if y_i is the i^{th} interpersonal distance measurement (in cm), then $y_i \sim \text{normal}(\hat{y}_i, \sigma^2)$, for $i = 1 \dots, n$, where $\hat{y}_i = X_i\beta$.

If y_i is the i^{th} speech or gaze rate (expressed as a proportion), then $y_i \sim \text{beta}(a = \kappa * \hat{y}_i, b = \kappa * (1 - \hat{y}_i))$, for $i = 1 \dots, n$, where $\hat{y}_i = \text{logit}^{-1}(X_i\beta)$.

If y_i is the i^{th} helping or shrug dichotomous coding (0 or 1), then $y_i \sim \text{Bernoulli}(\text{Pr}(y_i = 1))$, for $i = 1 \dots, n$, where $\text{Pr}(y_i = 1) = \text{logit}^{-1}(X_i\beta)$.

$X_i\beta$ is the same vector of predictors for all models. Aside from outcome y , the inputs in all models are the experimental condition j , the sex of the participant k , the participant's age group l , and the metro station m . The predictors are the experimental condition, the sex of the passenger, the age group of the passenger, all the two-way interactions between pairs of these factors, and the three-way interaction. An additional control predictor is the station where the interaction took place. Formally,

$$X_i\beta = \beta_0 + \beta_1\text{condition}_{j[i]} + \beta_2\text{sex}_{k[i]} + \beta_3\text{ageGroup}_{l[i]} + \beta_4\text{condition.sex}_{j[i],k[i]} + \beta_5\text{condition.ageGroup}_{j[i],l[i]} + \beta_6\text{sex.ageGroup}_{k[i],l[i]} + \beta_7\text{condition.sex.ageGroup}_{j[i],k[i],l[i]} + \beta_8\text{station}_{m[i]}.$$

In the normal- and beta-distributed models the variance parameter (σ^2 and κ , respectively) was estimated from the data. All hyperparameters were given noninformative prior distributions (Gelman et al., 2013; Kruschke, 2015).

To approximate the posterior distribution of the parameters of interest we used Markov chain Monte Carlo (MCMC) sampling as implemented by the software Jags (Plummer, 2003) via the programming language R (R Core Team, 2017). We checked that the samples were representative of the posterior distribution through visual examination of trace plots and density plots, on the one hand, and consideration of the Gelman-Rubin statistic of convergence, on the other. None of these checks gave any signs of unrepresentativeness. We additionally checked that the generated samples were large enough (and therefore accurate and stable) by considering a measure called the “effective sample size”. The estimates of all the parameters reported below rest on effective sample sizes of at least 10,000.

Reference

Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures*, 1(1), 77–89.