



HAL
open science

Interactional discrimination against hijab-wearing women in public places: field experiments

Martin Aranguren

► **To cite this version:**

Martin Aranguren. Interactional discrimination against hijab-wearing women in public places: field experiments. 2021. hal-03094580v2

HAL Id: hal-03094580

<https://hal.science/hal-03094580v2>

Preprint submitted on 6 Jan 2021 (v2), last revised 17 Oct 2022 (v3)

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.

**Interactional discrimination against hijab-wearing women in public places:
field experiments**

Martin Aranguren

Centre National de la Recherche Scientifique

Unité de Recherches Migrations et Sociétés

Université de Paris

martin.aranguren@cnrs.fr

Abstract. The notion of interactional discrimination is proposed to designate a so far unexplored mechanism of ascriptive inequality. The notion is illustrated through three randomized field experiments in public places showing that women who wear the islamic headscarf or hijab are exposed to interactional discrimination. Two experiments in the Paris metro relying on sensibly different procedures indicated that men decrease eye contact when they hear a woman speak to them if the woman wears the hijab. A third experiment in the same setting exposed a sample of women to this pattern of gaze, which was systematically replayed by a male confederate while women spoke to him. The experiment reveals that women who are the target of the “hijab-gaze” hold the floor for shorter periods of time, experience more negative affect, perceive that the interlocutor pays less attention, and are more easily persuaded.

Acknowledgements

The analyses presented in this paper are preliminary and have not been peer-reviewed. The data from Paris and Brussels have been collected as part of the MIDI project, funded by MITI-Centre National de la Recherche Scientifique with a Momentum grant to Martin Aranguren. The funder had no role in the conduct of this research.

The present document briefly reports a series of field experiments that jointly demonstrate the existence of a rarely investigated mechanism of ascriptive social inequality for which the designation *interactional discrimination* is proposed. Discrimination in general is unfavorable treatment explained by the ascription of its target to a specific social group¹. Discrimination may be said to be interactional when the mechanism that brings about the unfavorable outcome is to be found in the details of an interpersonal exchange.

What marks off the notion of interactional discrimination from other germane concepts available in the literature is the emphasis on social interaction as the mechanism (or mediating process) and the neutrality with regard to the outcome. The closely related idea of “interpersonal discrimination”², for example, focuses on the way in which stigmatized identities affect the course of interaction episodes, but the outcome of interest is restricted to the subtle communication and/or perception of negativity, assigning economic outcomes to the opposite pole of “formal discrimination”. Similarly, research on “microaggressions”³ accords primary importance to face-to-face encounters as their main *locus* of occurrence but the outcome is limited to subjectively felt offenses to dignity. The disadvantage caused by interactional discrimination may be economic, psychological, or of any other type of interest.

The experiments discussed below focus on one detail of dyadic social interaction, namely the way in which the listener gazes at the speaker, and demonstrate that specific differences in the pattern of gazing bring about differences on various outcomes. The stigmatized trait on focus is the wearing of the islamic headscarf or *hijab*. The backdrop of the studies are metro platforms and the participants randomly selected passengers. In each case, the experimental treatment involves a confederate who initiates a scripted interaction with a passenger. The present section offers a selective summary of these three experiments. While the first experiment has been described at length elsewhere, the rest of the present document reports the second and third experiments.

In collaboration with an Austrian team, the first experiment was carried out in the metros of Brussels, Paris, and Vienna, and relied on a helping behavior scenario⁴. On metro platforms, a confederate actress asked passengers for help either with uncovered hair or wearing a hijab. The second and third experiments focused on the Paris metro and used a sensibly different procedure involving an experimental game. For the second experiment, on the same Parisian platforms as before, an experimenter invited passengers to play an ultimatum game with another passenger (actually, the confederate), and players were allowed to interact before making a final decision. As in the first experiment relying on the helping scenario, in the second study involving the experimental game the confederate appeared either with uncovered hair or wearing a hijab. The first and second experiments, in spite of the marked difference in procedure, replicate the same finding:

male passengers, when they hear the confederate speak, decrease eye contact if the confederate wears the hijab.

The aim of the third experiment was to examine the average response of women to this way of being looked at by a male interlocutor while speaking. A male confederate was instructed to systematically imitate the gaze that men had spontaneously directed to the hijab-wearing woman in the previous experiments. The treatment in this experiment consisted in having the confederate replay that pattern of visual behavior when women spoke to him during the interaction preceding the decision in the ultimatum game. The results indicate that in the role of speaker, women who are looked at in this manner by a male listener tend 1) to decrease their speaking time in the context of a contradictory argumentation, 2) to experience more negative affect, 3) to perceive that the interlocutor pays less attention to what they say, and 4) to more often abandon their initial egalitarian view in favor of the opposite selfish-maximizer approach of the confederate.

In other words, in the context of a contradictory argumentation the documented decrease in eye contact while listening imposes various penalties on the interaction partner, illustrating an instance of interactional discrimination. Women who wear the hijab in Paris appear to be particularly exposed to this form of ascriptive inequality.

Ultimatum game in the Paris metro, treatment: hijab

BACKGROUND

The aim of the present experiment was to replicate the results of the study conducted in Brussels, Paris, and Vienna⁴ but using a sensibly different procedure. The latter experiment (n=840) examined if passengers standing on the metro platforms of these three European capitals treated a confederate actress differently when she wore the islamic headscarf or hijab. The procedure involved a helping scenario and measurements were based on videotapes of the resulting interactions.

The three-site study indicated that when randomly selected men, in the role of listeners, are asked for help by a woman on a platform of the Paris metro, they establish less eye contact with her if she wears an islamic headscarf or hijab. The generality of this finding is yet to be examined, however. In particular, a helping situation involves a power asymmetry between the interaction partners, to the extent that the help-seeker depends on the help-giver. Is the difference in visual behavior associated with the hijab specific to this asymmetric relationship or does it also generalize to situations in which partners interact on an equal footing?

To examine if the effect of the hijab on gaze transcends asymmetric helping situations, the present study preserves the design, population, sampling technique, treatment, outcome, and observation sites of the previous study but relies on an entirely different procedure in which passenger and confederate interact as equals.

METHOD

Experimental design. The experiment follows a 2 (hijab vs. uncovered hair) x 2 (female vs. male passengers) x 6 (metro stations) between-subjects design with random assignment of passengers to conditions and with equal sampling time devoted to each of the 24 unique factor combinations.

Sampling. Data collection was performed between May 6 and June 20, 2019. Random sampling 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, an experimenter 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 (say, a man), and then reapplying the method of systematic selection but only to passengers of the opposite sex (women). The third passenger was again selected regardless of sex, the fourth by stratifying by sex, and so on.

Procedure. The field team is composed of two experimenters and one confederate. A professional actress, the confederate was the same as the one involved in the previous experiment on the hijab. Experimenter 1 is in charge of recruiting passengers as they arrive to the platform, inviting them to participate in an experiment on decision making. If the passenger accepts, he is told that he or she will be playing an ultimatum game with another passenger. Using a portable tablet, Experimenter 1 explains the rules of the game to the passenger and checks that they have been correctly understood. In the ultimatum game⁵, a sum (in this case 10 euros) is to be shared by two players; one player, the Proposer, has the right to propose a division of the sum, whereas the other, in the role of Responder, has the right to accept or reject the proposal. If the Responder accepts the Proposer's division, both players earn the corresponding sums. If the Responder rejects the proposal, nobody gets anything. After explaining the game, Experimenter 1 asks the passenger his or her birthmonth "to randomly allocate an experimental condition to the assay". In the meantime, Experimenter 2 approaches discreetly with the confederate, pretending to be filling a questionnaire on another tablet. Experimenter 1 then introduces the passenger to Experimenter 2 and greets the confederate, presented as the second player of the ultimatum game. At this point, players are told that before the roles in the game are randomly assigned and the corresponding decisions made, the procedure requires them to have a short discussion about the game. With the passenger's consent, the ensuing interaction is videotaped with two cameras. To facilitate eye contact measurements, one of these, in the control of Experimenter 2, captures the passenger almost frontally using a narrow angle. To facilitate distance measurements, the other camera, handled by Experimenter 1, captures the profiles of passenger and confederate using a wider angle. During the interaction, following a standard script the confederate frames the game in terms of respect and disrespect, asserting that proposing an asymmetric share amounts to an assertion of the other player's inferiority. Once the interaction is complete, the roles in the game are "randomly allocated", and the passenger is invariably designated Proposer and the confederate Responder. Players are told that decisions will be made after separating and that they will be informed of the other's choice "via the tablet". Accompanied by Experimenter 2, the confederate leaves the platform pretending to exit the station. Experimenter 1 tells the passenger that it is time to reveal "the experimental condition of the assay" allegedly allocated on the basis of the passenger's birthmonth. Invariably, passengers are "randomly allocated" to the "constrained dictator (*dictateur contraint*) condition". The passenger becomes a dictator in the sense that the other player (i.e. the confederate) loses her right to reject the Proposer's division: the Proposer no longer needs the Responder for his or her division to take effect. But the passenger is a *constrained* dictator because only two options are now available to him or her. The Proposer has to choose between either taking nine and leaving one or taking one and

leaving nine. After the passenger makes his or her choice known and receives the chosen sum, Experimenter 1 explains the true aims of the experiment and invites the passenger to answer to a short sociodemographic questionnaire.

Experimental treatment. In one condition, the confederate wears an islamic headscarf or hijab. In the other, she appears with uncovered hair. The rest of the clothing is identical between the conditions. The clothing, the headscarf, and the actress are the same as in the previous experiment.

Measurements and outcome variables. The demographic variables that were measured with the questionnaire are age, educational achievement, income, and religion. The outcome variables include, on the one hand, the passenger's response to the game and, on the other, a set of nonverbal behaviors indicative of interpersonal involvement such as speaking time, eye contact, and interpersonal distance. The responses to the game were directly recorded on the tablet as they were given but the nonverbal measures were taken later on in the laboratory on the basis of the video and audio materials collected, using the program Elan Linguistic Annotator⁶. Regarding the nonverbal outcomes, intercoder reliability was assessed by computing Krippendorff's alpha⁷ for interval data for a subsample of independently coded assays, yielding in all cases coefficients higher than 0.7. The present report limits itself to the analysis of eye contact measurements.

The intensity of eye contact is operationalized as a proportion known as the "gaze rate", which equals the time spent looking at the other's eyes or face divided by the total observation period. Two observation periods were considered in this connection, namely the time spent in the role of the listener and alternatively the time spent in the role of the speaker. The script is organized in such a way that the confederate speaks at length twice but the passenger once. Consequently, each passenger contributes two measures of gaze rate while listening and one measure of gaze rate while speaking.

Sample description

	female passengers	male passengers
uncovered hair	54	51
hijab	44	49

Table 1: ultimatum game-treatment: hijab, passengers contributing usable gaze measurements, by condition and sex

Statistical analyses. The outcomes were analyzed through Anova-like models estimated with Bayesian inference in the framework of the Generalized Linear Model⁸, using the programming language R⁹ and the sampler Jags¹⁰. The unit of analysis in this model, indexed i , are by-passenger proportions of gaze rate while speaking or of gaze rate while listening. Each passenger contributes a maximum of two gaze rate measurements when listening and one gaze rate measurement when speaking. The inputs to the model are the gaze rates (outcome y), the experimental condition j , the sex of the passenger k , the station where the interaction took place l , the role in the conversation m , and the passenger's unique identifier o . The model assumes that the data follows a Beta distribution, and states that

$$\text{logit}(\hat{y}_i) = \beta_0 + \beta_1 \text{condition}_{j[i]} + \beta_2 \text{sex}_{k[i]} + \beta_3 \text{station}_{l[i]} + \beta_4 \text{role}_{m[i]} + \beta_5 \text{passenger}_{o[i]} + \beta_6 \text{condition.sex}_{j[i], k[i]} + \beta_7 \text{condition.role}_{j[i], m[i]} + \beta_8 \text{sex.role}_{k[i], m[i]} + \beta_9 \text{condition.sex.role}_{j[i], k[i], m[i]}, \text{ for } i=1, \dots, n=593.$$

All the parameters were given noninformative priors. The variance parameter in the beta distribution, known as the concentration, was estimated from the data and allowed to differ by conversational role.

RESULTS

[Insert Figure 1 here. Figures at the end of document]

Across conversational roles, among men the hijab decreases eye contact by between 0.2% and 14% (parameter “men”), but the same effect is not credible among women (“women”). In the role of the listener in particular, among men the effect of the headscarf is a decrease in eye contact ranging roughly from 1% to 10% (“menListener”). None of the remaining parameters estimate a credible effect of the hijab.

DISCUSSION

Using a helping scenario and involving passengers on metro platforms who were unaware of participating in a field experiment, the previous study found that men decrease eye contact with a woman who speaks to them if the woman wears the islamic headscarf. Preserving those platforms as the backdrop but explicitly inviting passengers to participate in an experimental game, the present study replicates that finding: men gaze less in general, and in particular in the role of listener, when the interlocutor is a hijab-wearing woman.

Ultimatum game in the Paris metro, treatment: hijab-gaze

BACKGROUND

Two different experiments conducted on platforms of the Paris metro with randomly selected passengers converge in the finding that men decrease eye contact with a woman speaking to them if the woman wears the islamic headscarf or hijab. In one case, the hijab-wearing woman asked for help to passengers who were unaware of participating in an experiment. In the other, passengers accepted to participate in an experiment on decision making involving an experimental game (but they were not aware until the end of the assay that the other player was actually a confederate actress).

The aim of the present experiment is to examine the average impact, for the recipient, of this way of being looked at while speaking. How do women in the role of speaker react to being looked at by a male listener as men turned out to gaze at the hijab-wearing woman in the previous experiments?

METHOD

Experimental design. The experiment follows a 2 (experimental condition) x 6 (metro stations) between-subjects design with random assignment of passengers to conditions and with equal sampling time devoted to each of the 12 unique factor combinations.

Pilot work. During October 2019 a series of field visits were scheduled to calibrate the procedure. The experimenters served as confederates and played a simplified version of the script that a professional actor was to perform in full-fledged version for the main study. This pilot work was instructive in several respects. First, it confirmed that most women spontaneously frame the ultimatum game in terms of an “obvious” egalitarian division. Second, it helped to design an adequate stimulus, i.e. a credible and effective imitation of the pattern of gaze that men had directed to the hijab-wearing confederate in the previous experiments. Third, it contributed to phrase the questions meant to capture women’s experience of this look in relevant everyday terms. Last, it revealed that the reactions to this gaze were very different among younger and more mature women. Women’s self-reported experience of this look, as captured by earlier versions of the Lickert-type questions eventually used, suggested that its effect was potent among the younger but weaker, if not inexistent, as age increased. To take into account this potentially important source of variability in the analysis of the main study, it was decided that age would be included in the models and allowed to interact with the experimental condition.

Sampling. Data collection proceeded between November 4 and December 4, 2019. The sampling technique was identical to the one used in the previous experiment, except that only female passengers were approached.

Procedure. With minor differences, the unfolding of the task is identical to the procedure used in the previous experiment until the discussion between the passenger and the confederate. During the discussion the confederate defends, with everyday words, the view prescribed by selfish rationality. He (the confederate is a male in this experiment) maintains that if he is given the role of Proposer, he will propose to take nine euros for him and to leave one euro for the passenger (i.e. the most unequal division that the game contemplates), whereas if he is allocated the alternative role of Responder, he will accept all offers, including one for him and nine for the other, because in all cases he earns something. The purpose of having the confederate follow this script was to maximize the passenger's speaking time. As indicated by the previous experiment and the pilot work preceding the present one, for most female passengers the spontaneous way of framing the game was in terms of an "obvious" equal division. By framing the game in terms of selfish interest, it was expected that the script would motivate female passengers to try to persuade the confederate that an egalitarian division was the best choice. When the discussion is complete and roles are "randomly" allocated, the passenger is invariably given the role of the Responder. After the confederate has left the scene with Experimenter 2, Experimenter 1 "receives" his offer on the tablet. As anticipated at the time of bargaining, he proposes nine euros for himself and one euro for the passenger. The passenger makes her choice, receives the corresponding sum, and is then informed of the true purpose of the research. She is then invited to answer to a questionnaire consisting in 9-point Lickert-type questions inquiring into the experience of the confederate's look, followed by sociodemographic questions.

Experimental treatment. In the control condition, the male confederate is instructed to look at the interaction partner in a "normal" way when he is in the role of the listener. More precisely, he must gaze continuously, with minor interruptions such as occasional blinks, while the passenger speaks to him. In the treatment condition, his gaze follows a "programme", i.e. an explicit set of rules, when the passenger holds the floor. The aim of the programme is to imitate the pattern of gaze of the men who interacted with the hijab-wearing woman in the previous experiment. The programme is not a fixed repetitive routine but a set of responses contingent on the passenger's visual and verbal behavior, in such a manner that its performance does not result in an unnatural "robotic" gaze, partly because the underlying rules are not easy to infer. To assess stimulus equivalence and adequacy, the visual behavior of the confederate was measured for all assays. The confederate's

gaze rate resulting from the performance of the programme mostly corresponds to the second quartile of the distribution of gaze rate measurements in the hijab condition in the previous experiment. For clarity, in what follows the gaze resulting from the performance of the programme is made reference to using the term *hijab-gaze*.

Measurements and outcome variables. As in the previous experiment, the demographic variables that were measured with the questionnaire are age, educational achievement, income, and religion. The outcome variables reported here are the time that the female passenger spent speaking in each part of the interaction (the script involved three main slots) and the Lickert questions. These measured the degree to which the confederate’s gaze elicited negative affect (*mal à l’aise, gênée, destabilisé*), breached informal social norms (*respectueux, correct*: reverse-coded), and indicated that the confederate lacked interest or attention (*intéressé, attentif*).

Sample description

	below 25 years of age	25 years of age or older
“normal” gaze	35	33
hijab-gaze	33	40

Table 2: ultimatum game-treatment: hijab-gaze, female passengers contributing speaking time measurements. The age categories reflect a median split.

Statistical analyses. Using Anova-like models computed with Bayesian inference in the GLM framework, the speaking time data were handled with a linear model and the answers to the 9-point scale questions analyzed with ordered probit regressions.

Linear model on speaking time. The unit of analysis in the linear model, indexed i , are by-passenger durations of speaking time (in seconds) in one of the three parts of the script. The inputs to the model are speaking time (outcome y), the experimental condition j , the station where the interaction took place k , the part of the script l , the age group of the passenger m , and the passenger’s unique identifier m . The model assumes that the data follows a t-distribution (to make the estimations robust to outliers), and states that

$$\hat{y}_i = \beta_0 + \beta_1 \text{condition}_{j[i]} + \beta_2 \text{station}_{k[i]} + \beta_3 \text{partScript}_{l[i]} + \beta_4 \text{ageGroup}_{m[i]} + \beta_5 \text{condition.ageGroup}_{j[i], m[i]} + \beta_6 \text{condition.partScript}_{j[i], l[i]} + \beta_7 \text{ageGroup.partScript}_{m[i], l[i]} + \beta_8 \text{condition.partScript.ageGroup}_{j[i], l[i], m[i]}, \text{ for } i=1, \dots, n=402.$$

All the parameters were given noninformative priors. The variance of the outcome variable was estimated from the data and allowed to differ by part of the script.

Ordered probit regression on 9-point questions. The unit of analysis in the ordered probit regressions, indexed i , are by-passenger answers to the 9-point Lickert-type questions. A separate ordered probit regression was computed for each of the eight questions. Each passenger contributes a single measurement. In each case, the inputs to the model are the answer (outcome y), the experimental condition j , the station where the interaction took place k , and the age group of the passenger l . The model states that

$$\hat{y}_i = \beta_0 + \beta_1 \text{condition}_{j[i]} + \beta_2 \text{station}_{k[i]} + \beta_3 \text{ageGroup}_{l[i]} + \beta_4 \text{condition.ageGroup}_{j[i], l[i]}, \text{ for } i=1, \dots, n.$$

All the parameters were given noninformative priors.

RESULTS

[insert Figure 2 here. Figures at the end of document]

Speaking time. The grand mean, that is the intercept of this Anova-like model, is estimated to lie between 14 seconds and 16s. The main effect of the hijab-gaze (i.e. the gaze that results from the performance of the programme) on the time that female passengers speak is a decrease ranging from less than a second to more than 4s (parameter “mainEffect”). However, when age is considered the effect appears to vary widely. The average speaking time among women younger than 25 lies between 11s and 14s, whereas that of women aged 25 or more ranges from 16s to 19s. The simple effect of the hijab-gaze for the latter group of women is a decrease lying between 2 and 8 seconds (“age2”), whereas the simple effect of that gaze among younger women is not credible (“age1”). Moreover, the difference between these differences yields a credible gap (that is, a condition*age interaction) of between 2 and 10 seconds. The third and last part of the script, whose duration is estimated to lie between 6 and 8 seconds, is the one where the difference between the conditions is most apparent (“part3”), ranging from 1 to 5 seconds. The difference corresponding to the last part of the script is even larger when attention is restricted to the group of women aged 25 or more (“age2part3”). Women in that age group, during the third part of the script, speak on average between 8 and 12 seconds. But when they are in the hijab-gaze condition, that quantity drops by between 4 to 11 seconds. This effect of the hijab-gaze on speaking time among women

aged 25 or more is credibly different from its counterpart among younger women (“age1part3”), and the difference in differences is estimated to lie between 4 and 13 seconds.

[insert Figure 3 here. Figures at the end of document]

Negative affect (Figure 4, panes a, b, c, and d). The hijab-gaze produces a credible main effect on three of the four affect-related outcomes (*mal à l’aise*, *gênée*, *destabilisée*), indicating this pattern of visual behavior increases negative affect. The size of the main effect is worth of note, ranging roughly from 1 to 3 points on a 9-point scale. In the three cases, the effect is stronger among women younger than 25, for whom the upper limit of the 95% central posterior interval reaches a 4-point difference. However, women do not report being irritated (*agacée*) to a different degree between experimental conditions.

Judgments of normative conformity (Figure 4, panes e and f). Women judge on average that the hijab-gaze is less “correct” or appropriate than the control gaze, and the estimated difference lies between slightly more than zero and 2 points ($\alpha=0.10$). Women do not credibly rate the hijab-gaze less respectful, although most of the posterior distribution is of negative sign.

Perceived involvement (Figure 4, panes g and h). Female passengers perceive that the hijab-gaze indicates less attentiveness on the part of the confederate than the control gaze, with a credible difference lying between 1 and 3 points. In contrast with the affective reaction data, the difference appears to be stronger among women aged 25 or more. Although weaker ($\alpha=0.10$), a credible negative difference also arises in the perceived level of interest that the programmed look tends to convey.

Post hoc analysis of speech content. The planned model on speaking time estimates that women hold the floor for shorter periods of time when the confederate replays the hijab-gaze as they speak. To better interpret the meaning of this brute decrease in duration, the transcripts of the conversations between the confederate and the female passengers were content analyzed. An analyst blind to the experimental condition read the entire corpus in search of relevant differences in content across the assays. It may be recalled that most women frame the experimental game in terms of an “obvious” egalitarian division, whereas the confederate was instructed to adopt the opposite point of view of a selfish maximizer. The analyst noted that in one part of the assays women stuck to their egalitarian position from the beginning to the end of the interaction (“persuaded no”), whereas in another part of the assays women started off with the egalitarian view

but eventually gave in to the opposite selfish view of the confederate (“persuaded yes”). The full set of conversations was coded with regard to this dichotomy. A second analyst, also blind to the experimental condition, following the instructions given by the first analyst recoded a random sample representing half of the transcripts. The first and second coders agreed on all codings.

This dichotomous outcome was analyzed by means of a logistic regression. The unit, inputs and linear function are identical to the one of the ordered probit regressions reported above, except that outcome y is the persuaded-no/persuaded-yes dichotomy and the predicted outcome \hat{y} operates on the logit scale. The model estimates that the hijab-gaze increases the probability that the woman will give in to the selfish view of the confederate by between 10% and 30%. No differences arise between the two age groups.

DISCUSSION

In this experiment, a confederate actor gazed at female metro passengers performing a systematic imitation of the gaze that men had spontaneously directed to the hijab-wearing confederate in the previous experiments. The aim was to examine the effect of the hijab-gaze from the point of view of its target. The sample is exclusively composed of female passengers.

The analysis shows that in the context of a contradictory argumentation the hijab-gaze imposes several penalties on its recipient. In the role of speaker, women who are looked at in this manner by the interlocutor tend 1) to decrease their speaking time, 2) to experience more negative affect, 3) to perceive that the interlocutor pays less attention to what they say, and 4) to more often abandon their initial egalitarian view in favor of the selfish approach of the confederate.

References

1. Lippert-Rasmussen, K. The badness of discrimination. *Ethical theory and moral practice* 9, 167–185 (2006).
2. Hebl, M. R., Bigazzi-Foster, J., Mannix, L. M. & Dovidio, J. F. Formal and interpersonal discrimination: a field study of bias toward homosexual applicants. *Personality and Social Psychology Bulletin* 28, 815–825 (2002).
3. Sue, D. et al. Racial microaggressions in everyday life: implications for clinical practice. *American psychologist* 62, 271–286 (2007).
4. Aranguren, M., Madrisotti, F., Martins-Durmaz, E., Gerger, G., Wittmann, L. & Méhu, M. 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.
under review.

5. Camerer, C. F. Behavioral game theory: experiments in strategic interaction. (Princeton University Press, 2003).
6. Sloetjes, H. & Wittenburg, P. Annotation by category - ELAN and ISO DCR. in Proceedings of the 6th International Conference on Language Resources and Evaluation (2008).
7. Hayes, A. F. & Krippendorff, K. Answering the call for a standard reliability measure for coding data. *Communication methods and measures* 1, 77–89 (2007).
8. Kruschke, J. K. *Doing Bayesian data analysis*. (Academic Press, 2015).
9. R Core Team. *R: a language and environment for statistical computing*. (2017).
10. Plummer, M. JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. in vol. 124:10 1–10 (2003).

Figure 1: ultimatum game-treatment: hijab, model on gaze rate proportions, central posterior intervals. On the y axis, the headings on the left-hand side of the plot specify the individual parameters. The quantities on the right-hand side specify the number of observations on which the estimation of the corresponding parameter directly relies; the first number specifies the size of the sample in the hijab condition, while the second number specifies that of the control condition. The x-axis represents the estimated value of the parameter. Within the plot area, the dashed vertical line indicates the location of the value 0, which signifies no difference and consequently the absence of an effect imputable to the hijab. 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. Our decision rule is to reject the null hypothesis of no difference if the 95% or 90% posterior interval of the parameter estimating the difference excludes the value zero. In graphical terms this implies that the thin (95%, $\alpha=0.05$) or bold (90%, $\alpha=0.10$) segment representing the parameter does not intersect the dashed vertical line. The differences in color are only meant to facilitate reading.

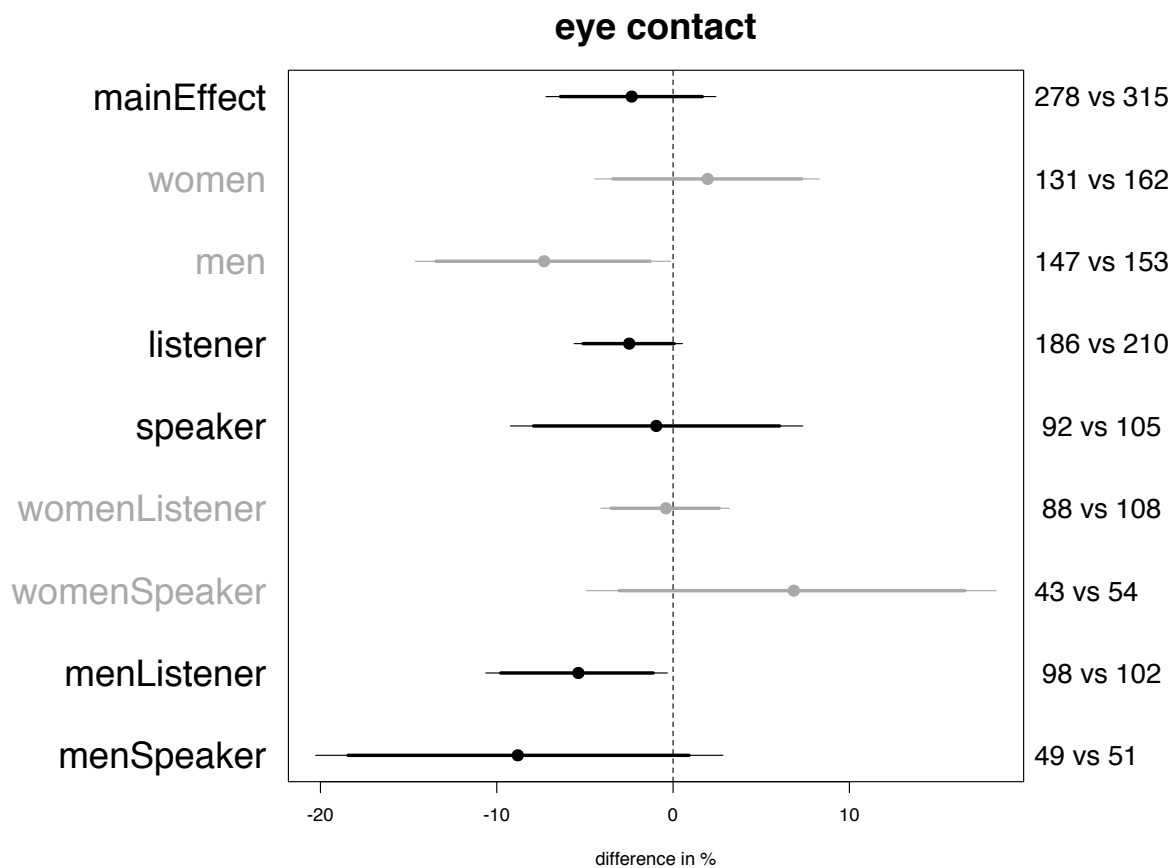


Figure 2: Ultimatum game scenario-treatment: hijab-gaze, model on speaking time in seconds, central posterior intervals.

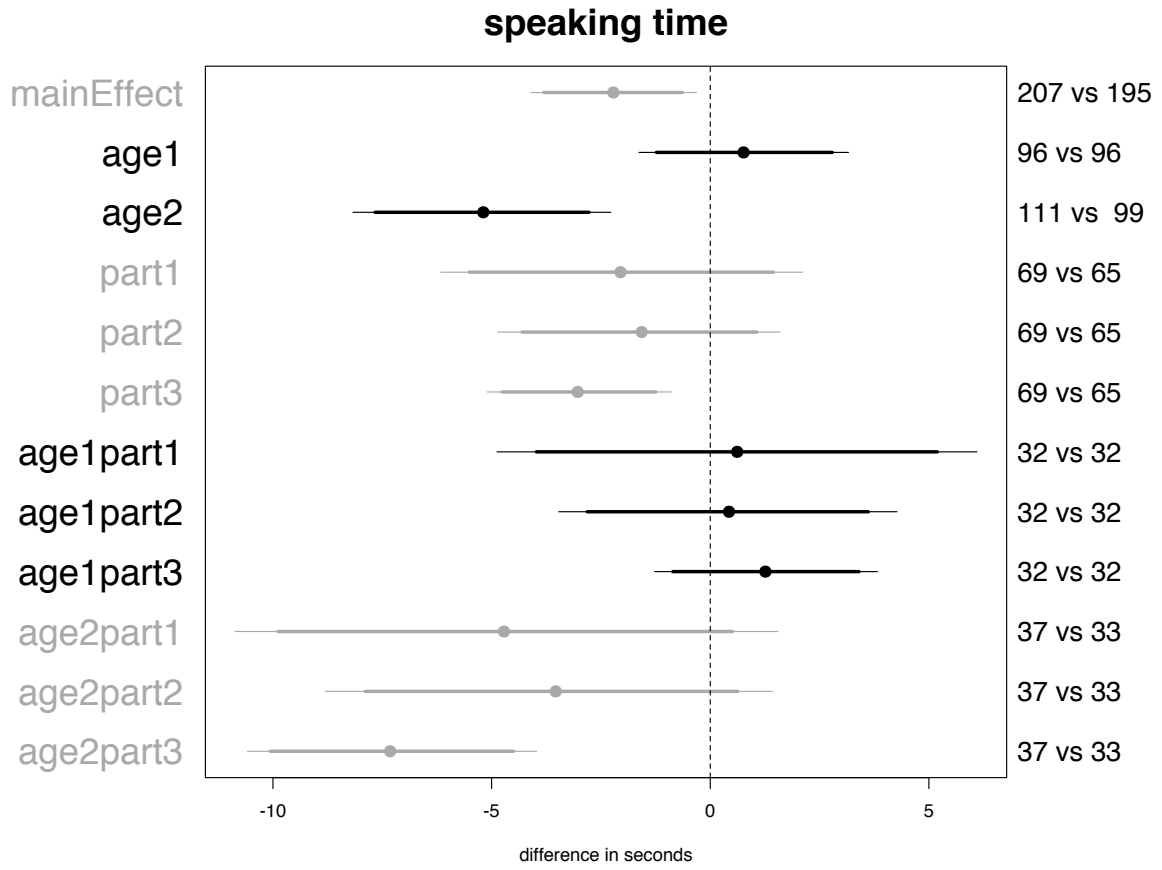


Figure 3: Ultimatum game scenario-treatment: hijab-gaze, models on answers to 9-point Likert-type questions describing the experience of being the target of the hijab-gaze. Panes a, b, c, and d: negative affect. Panes e and f: judgments of normative conformity. Panes g and h: perceptions of interlocutor's involvement.

