



HAL
open science

Interactional discrimination partly explains why Muslim women are more depressed than average: integrating survey data with field experiments in the study of minority stress

Martin Aranguren

► To cite this version:

Martin Aranguren. Interactional discrimination partly explains why Muslim women are more depressed than average: integrating survey data with field experiments in the study of minority stress. 2021. hal-03815399

HAL Id: hal-03815399

<https://hal.science/hal-03815399>

Preprint submitted on 14 Oct 2022

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

Interactional discrimination partly explains why Muslim women are more depressed than average: integrating survey data with field experiments in the study of minority stress

Author

Martin Aranguren

Centre de Recherche sur les Inégalités Sociales (CRIS), Sciences Po

martin.aranguren@sciencespo.fr

Abstract

Hundreds of studies indicate a robust correlation between perceived discrimination and emotional distress within stigmatized groups. However, about 80% of the related samples are from the US, and support for a within-group correlation is not yet evidence that discrimination contributes to disparities between groups. The article uses the minority stress model to examine the role of discrimination in European Muslim's surplus depression. The novel idea of interactional discrimination is proposed to integrate under an explicit and positive (vs. residual) concept some of the concrete forms of discriminatory behavior that are generally acknowledged to have a negative impact on mental health. Further, to measure interactional discrimination not only as a perception but also as a "distal" or objective event in the environment, the article supplements a study based on epidemiological survey data (European Social Survey, n=18,021) with two field experiments (n=370) involving live interactions with confederates in real-life situations.

Acknowledgments

I thank Francesco Madrisotti and Eser Durmaz for their help in performing the experiments and Carlo Barone, Hugues Lagrange, Allen Leblanc, Léa Pessin, Ettore Recchi, Mirna Safi, Lou Safra and Mike Vuolo for their comments on earlier versions of the manuscript.

Funding statement

The research was conducted with a Momentum grant to the author (MIDI project) funded by the MITI division of Centre National de la Recherche Scientifique, France.

Distinction

The manuscript was selected finalist for the Rinus Penninx Best Paper Award at the 19th Annual conference of the International Migration Research Network (IMISCOE), held in Oslo on June 29-July 1 2022.

Over the last two decades, hundreds of studies have investigated the correlation between discrimination and emotional distress, showing that the association is robust within stigmatized groups (reviewed in Williams et al. 2019). About 80% of the studies rest on samples from the United States, and Black Americans represent the stigmatized group that has concentrated most of the attention (Paradies et al. 2015). Further, as critics have observed (Schwartz and Meyer 2010), support for a discrimination-distress association within a stigmatized group (e.g. Black Americans) is not evidence that discrimination contributes to disparities in mental health *between* groups (e.g. between Black Americans and White Americans). Is the robust connection between discrimination and distress generalizable to countries other than the US? Does discrimination play a role in the generation of mental health inequities between groups?

The present article uses the minority stress model (Meyer 2003) to understand how discrimination negatively affects the mental health of Muslims in European countries. The minority stress model posits that members of minority groups such as European Muslims (but also racial minorities: Hughes et al. 2015, or the LGBTQ: Thomeer et al. 2018, or the intersection of both: Oyarvide Tuthill 2021) are disproportionately exposed to characteristic stressors which, in turn, explain why they are more emotionally distressed than average. Within this model, discrimination is acknowledged as a “distal” or objective stressor, to be distinguished from more “proximal” or intrapsychic stressors such as heightened vigilance or the internalization of stigma.

While the model posits that discrimination is objective, most research in this field actually relies on retrospective reports of *perceived* discrimination, creating a mismatch between the prescriptions of theory and the realities of measurement. Relatedly, although there is general agreement that the phenomena targeted by these measures include episodes such as being treated rudely by strangers, we lack a concept for the type of objective discrimination that is relevant in the generation of mental health disparities. In other fields, attempts to theorize these phenomena include the ideas of interpersonal discrimination (Hebl et al. 2002), micro-aggression (Sue et al. 2007), and assault on self-worth (Lamont et al. 2016). But, despite their merits, these ideas provide

less than a perfect fit for the minority stress model essentially because stress is not part of their definitions.

I propose the novel idea of interactional discrimination to integrate under an explicit and positive (vs. residual) concept some of the concrete forms of discriminatory behavior that are generally acknowledged to have a negative impact on mental health. Within the minority stress model, interactional discrimination designates an objective micro-social mechanism that explains why minority position, as a macro-social predictor, leads to poorer mental health, as a macro-social outcome. Further, to measure interactional discrimination not only as a perception but also as an event in the environment, I use a multi-method strategy in which a study based on epidemiological survey data is supplemented with two field experiments involving (manipulated) live interactions with confederates in real-life situations. Interactional discrimination is operationalized as a self-reported experience in the survey, but it is measured as an observable pattern of behavior in the experiments.

Using depression as a form of emotional distress, the survey data indicate that European Muslims are more depressed than average, and that the excess in depression is partly mediated by greater exposure to interactional discrimination. This analysis showcases highly religious Muslim women residing in France as the sub-sample that most neatly confirms the predictions of the minority stress model. The field experiments concentrate on this group. Corroborating the analysis of the survey data, the experiments show that Muslim women of high religiosity, as signaled by the use of an Islamic headscarf, are exposed to subtle but stressful forms of conduct in their everyday dealings with strangers in public.

The findings demonstrate the heuristic value of using the concept of interactional discrimination as an intervening variable within the minority stress model, as well as the increased validity enabled by the integration of survey and experimental data. To motivate this contribution, in what follows I define the concept of interactional discrimination and explain how this concept fits within the larger minority stress model as an intervening variable. Then I review the existing

work on the relationship between Muslim religious affiliation, discrimination, and mental health. Finally, I discuss the integration of epidemiological survey data with experimental measures.

1. Interactional discrimination: justification and definition

Before dealing with its interactional variety, it is useful to start with a definition of discrimination in general. It may be said that actor X discriminates actor Y when X treats Y differently from Z, the cause of the difference in treatment is the fact that Y and Z are (seen as) members of different social groups, and the consequence of the difference in treatment is an outcome less favorable to Y than it would had been the case had X treated Y as Z (adapted from Lippert-Rasmussen 2006). Discrimination is a causal and comparative concept. Its causal story is made up of two stages: first, a difference in group membership (e.g. Muslim vs non-Muslim) causes a difference in treatment; second, the difference in treatment causes a relative harm or disadvantage. Comparison is necessary at each stage: first, to establish a contrast between the way in which Y and Z are respectively treated; second, to establish an ordering (in terms of more and less, better and worse) between the consequences of that difference in treatment for Y and Z.

The typical “correspondence study” in labor economics and other related fields perfectly conforms to this definition (Bertrand and Duflo 2017). To investigate discrimination in the job market, the researcher creates at least two CVs that are identical in all respects save for a piece of information that allows the employer to assign the (fake) applicant to one social group or another. To assess if the two candidates are treated differently, the researcher records whether a given CV sent to a given employer (the experimental stimulus) prompts a reply, e.g. an invitation to a job interview (the response). By aggregating these dichotomous responses across the two experimental conditions, the researcher obtains a callback rate by condition – more rarely, researchers go on to see whether the callback translates in a job offer. If these turn out to differ, the researcher concludes that discrimination is at work.

Correspondence studies have been remarkably successful in demonstrating discrimination on various grounds across a variety of contexts. But one unintended consequence of their success, I surmise, has been to lead researchers and the public to restrict the practical applicability of the concept of discrimination to domains in which the detrimental outcome of receiving unequal treatment can be inferred from the act of applying, as in correspondence studies. Thus the concept of discrimination has been extended to other areas in which, analogous to the labor market, the obtention of a desired object goes through an application process, such as access to housing, credit or education (Pager and Shepherd 2008). Because the objects for which applications are relevant tend to coincide with goods and services that can be assigned a monetary value, another unintended bias present in public debates as much as in scientific practice has been to restrict the outcomes that may be affected by discrimination to the sole economic dimension (Lamont et al. 2016).

Given this unintended but unnecessary restriction of the concept of discrimination, stress research is well poised to widen the view. Stress is generally undesirable. If stress is the consequence of a difference in treatment caused by assignment to different social groups, it follows that stress must be admitted, alongside economic utility, as an independent dimension of life upon which discrimination may impinge.

However, to the best of my knowledge in the area of stress the objectivistic definition of discrimination that I have spelled out has not been used. In its lieu, what has been measured is *perceived* discrimination, as self-reported (Williams et al. 2019). Perceived discrimination raises the vexing methodological problems of unreliable recall, absence of an external criterion and confounding with the outcome. At the time of responding, does the respondent remember all the experienced occurrences of discrimination? Did discrimination really occur? Is the respondent's report of perceived discrimination an indication that discrimination occurred or a symptom of the respondent's mental health condition?

Discrimination may be said to be interactional when its constituent components (assignment to a group, difference in treatment, detrimental outcome) occur within a spatially and temporally

bounded unit of social interaction. The idea of interactional discrimination operates entirely at the “micro” level of interpersonal encounters and leaves the nature of the adverse consequence unspecified. The detrimental outcome may manifest itself on the economic dimension, on that of stress, on both or on any other. Further, interactional discrimination primarily designates an objective episode of social exchange, a pattern of observable behavior, and only secondarily a subjective report of that episode.

In this sense, the idea of interactional discrimination addresses two mostly tacit limitations of contemporary research practice in the field of discrimination. The first limitation is conceptual. Although not necessarily in theory, in practice most research on objective (vs. perceived) discrimination has restricted the application of this concept to the economic realm, focusing on access to employment, housing, credit, etc. The second limitation is methodological. Again, although not necessarily expressing a theoretical choice, in practice most research that has inquired into the non-economic adverse consequences of discrimination (e.g. stress) has done so on the basis of self-reported measures of perceived discrimination. The concept of interactional discrimination is at the same time neutral with regard to the adverse outcome of differential treatment (vs. a restricted focus on economic disadvantage) and designates primarily an observable episode of social interaction (vs. a possible confusion with the analytically distinct experience of being discriminated against).

In sum, what makes discrimination *interactional* is not any particular form of harm or disadvantage, economic, emotional or other, but the particular unit or context within which the discriminating occurs: a spatio-temporally delimited episode of social exchange in conditions of copresence – that particular social unit whose classical sociological analysis can be found in the writings of Erving Goffman (1963). In this sense, the cognate ideas of microaggression (Sue et al. 2007) and assault on self-worth (Lamont et al. 2016) point to a specific variety of interactional discrimination characterized by a particular type of harmful consequence, namely the experience of indignity. But the variety of interactional discrimination documented in the present article is

different. Its harmful consequence is stress, whether or not the stressful confrontation is interpreted as an assault on dignity.

Other advantages of the idea of interactional discrimination are discussed *a posteriori* in view of the results. To be sure, there is no claiming that the phenomena designated by this concept have been neglected by research. Especially in the field of mental health, measures of (perceived) discrimination often refer to events of differential treatment that occur within episodes of social interaction. The complaint is not that research has failed to acknowledge these phenomena, but rather that it has failed to give them a positive and explicitly formulated concept. I take up this issue in the Discussion.

2. Interactional discrimination as a variable within the social stress model

To clarify the theoretical role of interactional discrimination as an intervening variable, I recast the “social stress” model (Pearlin 1989), of which the more specific minority stress one is an outgrowth, as a mechanism-based explanation (Hedström and Ylikoski 2010). To this purpose, Figure 1 uses the graphical conventions of the famous diagram that sociologist John Coleman (1986) proposed to articulate micro-macro explanations. The schema begins by stating that a group’s position in a social hierarchy (component A on Figure 1) influences the position that the same group has in an ordering of mental health outcomes (F). A dashed, discontinuous arrow represents this macro-to-macro statement to indicate that the intervening micro-level mechanism is still in need of articulation. Descending from the macro to the micro, the schema goes on to state that the prevailing social hierarchy with its ordering of positions (A) influences the beliefs, motivations, habits, practices, etc. of individual actors. The word “bias” (B), in common usage in the contemporary psychology of prejudice (Richeson and Sommers 2016), seems apt to summarize the influence of the social structure on individual actors’ readiness to think, feel and act in specific ways in their dealings with others of varying status. The schema then states that bias (B) generates differences in treatment (C). The following link in the schema states that, for its recipient, this

difference in treatment (C) elicits stress (D), as indexed for instance by the experience of a negative emotion. By virtue of repetition over time, the schema pursues, the stress (D) that discriminated individuals experience in bounded episodes of discrimination cumulatively determines their probability of developing some form of distress (E), e.g. depressive symptoms. Ascending from the micro to the macro to complete the explanation, the last step in the schema states that the aggregation of individual distress outcomes (E) across groups will yield a similar ordering to the status hierarchy posited at the beginning of the causal chain (F), with the average level of distress rising as the cursor moves down in the hierarchy.

The concept of interactional discrimination specifies the “micro” process that encompasses components B, C and D within the model: in face-to-face encounters, individuals assigned to different groups (B) are treated differently (C), and the difference in treatment exposes them differently to stress (D). In return, the model specifies the form of harm in focus, namely stress (to be distinguished, for example, from monetary disadvantage or the experience of indignity). To derive concrete predictions about European Muslims, I now briefly review the empirical literature dealing with religious affiliation as a predictor of mental health and of discrimination.

[Insert Figure 1 about here – Figures available at the end of the document]

Figure 1: The social stress model recast as a mechanism-based explanation.

3. Muslim religious affiliation as a predictor of mental health and interactional discrimination

Muslims and mental health

Depending on the religious group under consideration and other contextual characteristics, religiousness may be associated with better or worse mental health outcomes relative to a baseline usually represented by individuals with no religious affiliation. To account for these contrasting trends, researchers have hypothesized that religious affiliation may be indexing helpful psychosocial resources, or alternatively that it may function to marginalize its incumbents.

According to the first view, having a religious affiliation decreases the risk of distress because it increases the probability of wielding a number of psychosocial buffers such as social support and the ability to appraise and cope with taxing circumstances in ways that decrease their stressful character (Koenig 2009). Mostly cross-sectional in design and based on samples from the United States, studies comparing Christians to individuals without any religious affiliation tend to confirm that religious affiliation is inversely correlated with symptoms of depression (Ellison and Henderson 2011).

According to the second view, conversely, religious affiliation may increase the risk of distress if it operates in a “dissonant religious context” (Rosenberg 1962) because individuals who are discordant in their characteristics from those that prevail in the surrounding population will more often be categorized as “outgroups” than as “ingroups” in social interaction. The well-studied phenomenon of “ingroup favoritism” (Tajfel et al. 1971) implies for them fewer rewards and more frequent exposure to suspicion or hostility in these interactions, and in the long run the difference in treatment is expected to negatively affect their mental health. For example, Muslims in Europe in general (Schnittker 2020), and South Asian Muslims in Scotland in particular (Williams and Hunt 1997), tend to report higher levels of depression than non-Muslims, as do Jews in the United States when compared to non-Jews (McCullough and Larson 1999).

Muslims in European countries of Christian tradition clearly represent a case of dissonant religious affiliation, leading to the prediction of a mental health disadvantage for Muslims. But beyond a general deficit in ingroup favoritism, do we expect European Muslims to be more exposed to interactional discrimination more specifically?

Muslims and interactional discrimination

The available cross-national evidence from Europe indicates that a sizable proportion of non-Muslims hold anti-Muslim views and, at the other end, that Muslims feel discriminated against. Depending on the survey, the measure and the country, in the ten European states with the largest

proportion of Muslim residents the incidence of anti-Muslim attitudes oscillates between 5% and 38% of the population (European Commission 2019; European Values Study 2017), and the part of Muslim respondents who report incidents of discrimination ranges from 9% to 37% (FRA 2017; Pew Research Center 2006).

Moving from subjective perceptions to objective measures of discrimination, the results depend on gender, the context of social activity and the stimulus used to signal Islamic affiliation. In the context of hiring, and typically using the mention of a religious volunteering antecedent in a fake CV as the signal, correspondence studies conducted in France (Adida, Laitin, and Valfort 2010; Pierné 2013; Valfort 2020) but also in other European countries (Di Stasio et al. 2021; Koopmans, Veit, and Yemane 2019) consistently show that Muslim applicants have a lower probability of receiving a job callback. In the context of everyday contacts in public places, typically using a perceivable sign of Islamic religiosity as the stimulus, the research design allows to know more about the perpetrators and the act of discrimination, but the results are less clear-cut. A series of field experiments in European cities have examined the effects of the Islamic headscarf or hijab on helping behavior. Some have found that people are less likely to offer assistance to a woman if she wears the hijab (Aidenberger and Doehne 2021; Choi, Poertner, and Sambanis 2020) but others indicate no difference in helping between the presence and the absence of the headscarf (MASKED; Diekmann, Jann, and Näf 2014). One of these field experiments, performed with the same procedure in Brussels, Paris and Vienna, also included measures of more subtle cues of interpersonal involvement, finding a mix of warmth and coldness in response to the hijab depending on the city and the gender of the participant (MASKED). Another field experiment in Paris, using a similar procedure to assess whether people treated a bearded man differently when he insinuated that he was going to a mosque, indicated that in this “mosque condition” participants offered help less often and showed lower interpersonal involvement (MASKED).

The mentioned differences in treatment observed in the context of interpersonal encounters in public places are candidate instances of interactional discrimination against Muslims. Depending

on how they are interpreted, they may cause stress in the recipient. Using the Daily Life Experiences subscale of a checklist instrument named RaLES (Harrell 1997) as a catalogue of acknowledged stressful experiences, not being helped when in need or being the recipient of inadequate interpersonal involvement in day-to-day contacts with strangers could be experienced as instances of “being treated rudely or disrespectfully”, “others reacting to you as if they were afraid or intimidated”, “being treated as if you were stupid”, “having your ideas ignored” or “not being taken seriously”. Even an intensification of interpersonal involvement could be stressful if the recipient interprets the increase in warmth as an instance of “being treated in an ‘overly’ friendly or superficial way”.

In this sense, there is some evidence that European Muslims regard instances of inadequate nonverbal involvement as offenses. According to the EU-MIDIS survey, nearly 40% of Muslim female respondents who indicated that they wear religious attire in public (mostly, the hijab) reported that they were the target of “inappropriate staring” or “offensive gestures” because they did so (FRA 2017). To the best of my knowledge, the connections between Islamic faith, interactional discrimination, stress and mental health have not been investigated in previous work.

Predictions

The proposed explanatory schema (see Figure 1), in combination with the cited literature, yields the following main prediction:

P1: Interactional discrimination mediates Muslims’ mental health disadvantage compared to non-Muslims.

As a statement about mediation (MacKinnon, Fairchild, and Fritz 2007), P1 requires prior confirmation of the following predictions:

P2: Muslims exhibit poorer mental health than non-Muslims.

P3: Muslims are more exposed to interactional discrimination.

In turn, drawing on the definition of interactional discrimination and its specification within the social stress model, P3 can be analyzed into two distinct expectations:

P3a: Muslims are treated differently from non-Muslims in face-to-face encounters.

P3b: The differences in treatment to which Muslims are exposed in face-to-face encounters cause them stress.

4. Integrating epidemiological and experimental data

Within the mechanistic reformulation of the social stress model presented above, interactional discrimination plays the role of a micro-level process that causally connects differences in social position (the “macro” antecedent) with inequalities in mental health (the “macro” consequent). Thus specified, interactional discrimination has the theoretical status of a stressor, i.e. an outer or environmental demand that tends to cause stress (this is what Prediction 3b states). The objectivity of interactional discrimination creates the methodological desideratum of measuring this process objectively, going beyond the customary perceived discrimination scales, which rely on self-reports. A pattern in these scales is that questions are usually about mundane interpersonal events, their occurrence (yes/no) or frequency (from never to very often), and their likely cause or motive (e.g. the respondent’s race or ethnicity). For example, “How many times in the past year/in your entire life have you been treated unfairly by strangers because you are Black ?” (Schedule of Racist Events, Landrine and Klonoff 1996)

To satisfy this desideratum, the article proposes an original mix of methods.

Epidemiological data are collected with surveys and describe a population in terms of health outcomes. Often, surveys that measure epidemiological variables also include a battery of sociodemographic variables, and sometimes even questions about the experience of being discriminated (e.g. the European Social Survey). Examining the relationships between these variables is indispensable for testing the hypothesis that discrimination mediates the influence of social position on disparities in mental health. But this is not enough. First, because of the logical

mismatch between the research design (observational and very often cross-sectional) and the research objective (a causal explanation). Second, because of the unknown validity of self-reported experiences of being discriminated as indirect measures of actually occurring discrimination. These shortcomings have been repeatedly acknowledged and extensively analyzed by reviews of the discrimination-health association (Lewis et al. 2015) and more generally in the social-psychological literature concerned with stressors (Dohrenwend 2006).

In this context, the methodological innovation consists in supplementing the indispensable epidemiological data with experimental evidence collected in the field. The field experiments to be reported below involve brief face-to-face exchanges between a trained confederate and a randomly selected person in a real-life situation, using measures of behavior as the outcome variables. Randomization provides the desired basis for causal inference. Behavioral outcomes measure discrimination objectively. But note that for all its objectivity and explanatory power the experimental paradigm alone cannot be used to study the full social stress process. Consider the example of depression as a mental health outcome. Beyond its technical feasibility, it is morally unjustifiable to expose other human beings to an experimental treatment that is likely to sink them in a state of emotional distress as severe as depression. Inequalities in depression cannot be manipulated, only observed, and epidemiological surveys provide the necessary observational information.

Overview of the empirical research

The specific mental health outcome used in the present investigation is depression. P1, P2 and P3 were tested by analyzing data from the 7th round of the European Social Survey (ESS). This epidemiological study confirmed Muslims' expected surplus in depression and provided *prima facie* evidence in favor of interactional discrimination as the intervening mechanism.

To examine if the successful mediation analysis using ESS data had a correlate in objective occurrences of interactional discrimination against Muslims, two field experiments involving face-

to-face exchanges with trained confederates were conducted in urban public places. Studies of social interaction cannot possibly aspire to have the same demographic coverage as a cross-national survey representative of the European population. To connect the results of the ESS study with the field experiments, the focus had to be narrowed down. To this purpose, the group by gender, religiosity and country of residence that best confirmed the mediation hypothesis in the ESS study was selected for further investigation. This most illustrative group is the one composed by highly religious Muslim women based in France.

Using the Islamic headscarf or hijab as a public signal of high Muslim religiosity, and platforms of the Paris metro as the setting, the first field experiment involved interactions between randomly selected passengers and a female confederate who wore a hijab in the treatment condition but appeared with uncovered hair in the control condition. Supporting P3a, male passengers were found to establish less visual contact when the confederate was in the hijab condition.

But this difference in behavior cannot be immediately treated as an instance of discrimination. What if the evasive look functioned to convey respect? To examine whether the evasive look acted as a stressor, as suspected by P3b, or rather as a token of regard, as it may be alternatively hypothesized, the second field experiment used the same setting and procedure as the previous one but changed the gender of the confederate, the experimental stimulus and the target population. Not a female, but a male confederate served in this experiment. He was trained to imitate the pattern of gaze that male passengers had directed to his female counterpart when she played in the hijab condition in the previous experiment. This “hijab-gaze”, so to name it, acted as the stimulus of the second experiment. Passengers were sampled from the same places as before, but only females were eligible as experimental participants (any female passengers, not only hijab-wearing women). The experiment thus sought to inquire into the response of the average woman to the evasive look that men had been found to direct to the hijab-wearing confederate. As anticipated by P3b, the results revealed that the “hijab-gaze” aroused strong negative emotions in the recipient,

bolstering the expectation that this subtle but objective difference in treatment does function as a stressor.

Triangulating the results of the ESS epidemiological study and those arising from the field experiments, the evidence suggests that interactional discrimination carried out by men, as exemplified by evasive visual behavior, partly explains the surplus in depression that is observable among highly religious Muslim women in France.

Method

Epidemiological study: Predictions 1, 2 and 3

(For more details, see Supplementary Materials, Method, Epidemiological study)

Measures. The analysis relies on data from the seventh round of the European Social Survey, conducted in 2014. The mental health outcome in this analysis is depressive symptoms, as measured on the basis of the Center for Epidemiological Studies Depression Scale (CES-D, Radloff 1977). The variables of interest used as predictors of that outcome are Islamic religion, level of religiosity, perceived discrimination, age, educational achievement, income, gender and country. A three-level nominal variable was created by distinguishing Muslims from non-Muslims, then more religious from less religious Muslims. Drawing on estimations of the likely proportion of Muslims in the total population (Pew Research Center 2015), only European countries whose proportion of Muslims was assessed to be 5% or higher were kept in the analysis, yielding the following list of ten: Austria, Belgium, Denmark, France, Germany, Netherlands, Norway, Sweden, Switzerland and the UK.

Statistical analyses. The standard three-step method for testing mediation was implemented, estimating hierarchical regression models with Bayesian inference (Gelman et al. 2013; Kruschke 2015). These models estimate at the same time pan-European differences between Muslims and

non-Muslims and more local differences by country, by gender and even by country-gender combination.

Note on Bayesian inference. In all the analyses reported in this document, null-hypothesis tests are accomplished by examining whether the “central posterior intervals” (Gelman et al. 2013) estimating the parameters of interest credibly differ from zero. The chosen alpha level is the standard 5%. The null hypothesis of no difference between groups is rejected if zero is excluded from the values comprised between the 2.5th and 97.5th percentiles of the posterior probability distribution under examination. In frequentist statistics, “statistically significant” is the expression used to flag differences that, according to some distribution, differ from zero with some usually high probability. The Bayesian equivalent is the adjective “credible”.

Field Experiment 1: Prediction 3a

(for more details, see Supplementary Materials, Method, Field Experiment 1)

Experimental design. The experiment follows a 2 (headscarf 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 proceeded between May 6 and June 21, 2019. The sampling technique consisted in approaching the first passenger arriving to the platform after the departure of the last train and before the arrival of the following one, alternating between the sexes to ensure equal representation.

Procedure. Experimental assays took place on metro platforms and involved a team of three: Experimenter 1, Experimenter 2 and a female confederate. Experimenter 1 recruited passengers as they arrived at the platform, inviting them to participate in an experiment “on decision making.” If the passenger accepted, her or she was told that he or she would be playing an ultimatum game

(Camerer 2003) with another passenger. While Experimenter 1 explained the rules of the game to the recruited passenger, Experimenter 2 approached discreetly with the confederate, pretending to be filling a questionnaire. Experimenter 1 then introduced the passenger to Experimenter 2 and greeted the confederate, presented as the second player of the ultimatum game. At this point, players were told that before the roles in the game were randomly assigned and the corresponding decisions made, the procedure required them to have a short discussion about the game. With the passenger's (and the confederate's pretended) consent, the ensuing interaction was videotaped with two cameras. The confederate's contributions to the conversation followed a script.

Experimental treatment. In one condition, the confederate wore an Islamic headscarf or hijab. In the other, she appeared with uncovered hair. The rest of the clothing was identical between the conditions. The confederate is the same woman in both conditions.

Measurements. The demographic variables probed with the questionnaire were age, educational achievement, income, and religion. The outcome variables included, on the one hand, the passenger's response to the game and, on the other, a set of nonverbal behaviors indicative of interpersonal involvement, including eye contact. The responses to the game were directly recorded on the tablet as they were given but the nonverbal measures were taken in the laboratory on the basis of the collected video and audio materials, using to this effect the program Elan Linguistic Annotator (Sloetjes and Wittenburg 2008). Regarding the nonverbal outcomes, intercoder reliability was assessed by computing Krippendorff's alpha (Hayes and Krippendorff 2007) for interval data for a subsample of independently coded assays (at least 20%), yielding in all cases coefficients higher than 0.7.

Eye contact : measurement details. The intensity of eye contact was 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 (Exline, Ellyson, and Long 1975). 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. In this experiment, eye contact is used to examine differences in treatment between conditions (not whether differences in eye contact are stressful, which is the aim of the second experiment).

Statistical analyses. The outcomes were analyzed using regression models estimated with Bayesian inference. The gaze rate measurements, continuous but restricted to the interval $[0, 1]$, were treated with a beta-distributed logistic regression. The dichotomous choice in the game was analyzed with a linear probability model.

Field Experiment 2: Prediction 3b

(for more details, see Supplementary Materials, Method, Field Experiment 2)

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 confirmed that most women spontaneously framed the ultimatum game in terms of an “obvious” egalitarian division. It also 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 Experiment 1. Last, it contributed to phrase in relevant everyday terms the questions meant to capture women’s experience of this look.

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 was identical to the procedure used in the previous experiment.

Experimental treatment. In the control condition, the male confederate was instructed to look at the interaction partner in a “normal” way when he was in the role of the listener. More precisely, he had to gaze continuously, with minor interruptions such as occasional blinks, while the passenger spoke to him. In the treatment condition, his gaze followed a “program”, i.e. an explicit set of rules, which were active as long as the passenger held the floor. The goal of the program was to generate a natural-looking imitation of the pattern of gaze of the men who had interacted with the hijab-wearing woman in Experiment 1. The program was 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 did not result in an awkward “robotic” gaze, partly because the underlying rules were not easy to infer. To assess stimulus equivalence and adequacy, the visual behavior of the confederate was measured for all assays. In performance of the program, the confederate’s gaze rate turned out to fall within the second quartile of passengers’ original gaze rates in the hijab condition in Experiment 1. The term “hijab-gaze” refers to the pattern of visual behavior that resulted from the performance of this program.

Measurements. The demographic variables measured with the questionnaire were age, educational achievement and income. The outcomes were measured using 9-point Likert-type questions concerned with i) stress operationalized as negative affect (four items), ii) attributions of interpersonal involvement (two items) and iii) judgments of conformity with rules of politeness (two items). The negative affect items inquired into the degree to which the confederate’s gaze elicited negatively-valenced emotional states (*mal à l’aise, gênée, agacée, perturbée*). The involvement items sought to capture the extent to which the confederate’s gaze indicated that he was attentive to, and interested in, the participant’s speech (*attentif, intéressé*). The politeness items asked about the degree to which the confederate’s gaze was judged to be correct and respectful (*correct, respectueux*). Internal reliability was satisfactory for all sets of questions related to the

same construct, as assessed by Cronbach's alpha (negative affect, $\alpha=0.82$; involvement, $\alpha=0.86$; politeness, $\alpha=0.83$).

Statistical analyses. The answers to the 9-point scale questions were analyzed with ordered probit regressions estimated with Bayesian inference.

Results

Epidemiological study: European Social Survey, 7th round data

Descriptive statistics

Figure 2 displays the by-country proportional weights for the whole sample, for Muslim women and for Muslim men. Representing over 75% of the total weight, Germany, France and the UK dominate the overall sample, reflecting the size of their respective populations. Within countries, the weight of Muslim men is approximately proportionate with the weight of the national sample. In contrast, the weight of Muslim women in the whole sample is carried at nearly 50% by the subsample from France, and at 30% by the sub-subsample of highly religious Muslim women residing there. Conversely, the weight of the subsample of Muslim women from Germany is well below the weight of the national sample.

[Insert Figure 2 about here]

Figure 2: Epidemiological study; proportional weights by country.

P2: Muslims exhibit poorer mental health than non-Muslims.

On a 24-point scale, the average level of depression is 5.56 at the 2.5th percentile of the estimated posterior distribution, and 6.11 at its 97.5th. In short, the average level of depression is [5.56, 6.11]. Confirming P2, overall Muslims (vs. non-Muslims) present a credible excess in depression in the interval [0.25, 0.92], with no difference between males and females (see Figure 3, pane “Muslim vs. non-Muslim”, rows “average”, “male” and “female”). At the country level, that excess is sufficiently strong to stand as a credible effect in France, the Netherlands and the UK. At the country-gender level, the gap remains credible among males in the Netherlands and the UK, and among females in France.

Neither overall nor within gender groups do more and less religious Muslims differ in surplus depression (see Figure 3, pane “interaction Muslim*religiosity”, row “average”). A single, very marked exception stands out, however. In France, highly religious Muslims report more depression than the less religious, and the difference between these two groups of Muslims is in the range [0.94, 2.6]. The pattern is the same when French male and female respondents are considered separately (see Figure 3, pane “interaction Muslim*religiosity”, rows “France”, “France, m” and “France, f”).

Looking in particular at the excess in depression among highly religious Muslims in France, the difference relative to non-Muslims is estimated to be in the interval [0.94, 2.12] (see Figure 3, pane “interaction Muslim & high religiosity vs. non-Muslim”, row “France, f”).

[Insert Figure 3 about here]

Figure 3: Epidemiological study; Muslims’ surplus depression

P3: Muslims are more exposed to interactional discrimination.

In the analysis of ESS data, interactional discrimination was measured with a proxy, namely the answer to the question whether the respondent belongs to a group discriminated against in the country. The overall proportion of the population that belongs to a discriminated group is estimated

to be in the range [22%, 27%]. Confirming P3, overall Muslims are [20%, 28%] more likely to check this question than non-Muslims (see Figure 4, pane “Muslim vs. non-Muslim”, “average”). No overall or by-gender difference in surplus discrimination arises from the comparison between more and less religious Muslims (see Figure 4, “interaction Muslim*religiosity”, “average”, “male” and “female”), but substantial differences emerge between groups of respondents by country and gender.

[Insert Figure 4 about here]

Figure 4: Epidemiological study; Muslims’ surplus discrimination

P1: Interactional discrimination mediates Muslims’ mental health disadvantage compared to non-Muslims.

Confirming the mediation hypothesis entailed in P1, for the entire sample coefficients a (association between predictor Muslim/religiosity and mediator discrimination), b (association between mediator discrimination and outcome depression) and c (association between predictor Muslim/religiosity and outcome depression) are positive and c' (association between predictor Muslim/religiosity and outcome depression when mediator discrimination is controlled for) is null, indicating that Muslims’ higher level of depression is at least partially mediated by interactional discrimination (see Figure 5, pane “Muslim vs. non-Muslim”, row “average”; see Supplementary Material, Epidemiological study, Method, Prediction 1, Statistical analyses for a description of classical mediation analysis). The result remains when more and less religious Muslims are contrasted, separately, to non-Muslims.

When subsamples are considered, mediation does not remain credible for all the contrasts within all country, gender, and country-gender groups. Averaging over countries, mediation is upheld for females in all contrasts but for males only when Muslims of low religiosity are compared to non-Muslims. At the country level, France is the only country where mediation is supported when the

Muslim vs. non-Muslim contrast is considered (see Figure 5, pane “Muslim vs. non-Muslim”, row “France”). At the country-gender level, the only group where mediation is upheld concerns highly religious women residing in France (see Figure 5, pane “Muslim & high religiosity vs. non-Muslim”, row “France, f”). Within the latter group, considering the influence of discrimination as a mediator brings the gap in depression between highly religious Muslim women and non-Muslim women from [0.94, 2.12] as per coefficient c down to [0.24, 1.55] as per c' .

[Insert Figure 5 about here]

Figure 5: Epidemiological study; discrimination mediates Muslims’ surplus depression

Field Experiment 1: female confederate wears the hijab

P3a: Muslims are treated differently from non-Muslims in face-to-face encounters.

In view of the results from the epidemiological study, P3a was thus specified: In France, highly religious Muslim women will be treated differently from non-Muslim women.

On average, passengers gaze at the confederate [78%, 83%] of the time. Confirming P3 for men, in the hijab condition male passengers reduce by [-12%, -2%] the intensity of eye contact (see Figure 6, pane “eye contact”, row “hijab @male”). Additionally, a treatment*gender interaction in the interval [-18%, -3%] indicates that the effect of the hijab differs by gender, suggesting opposite effects of the garb for men and women. In contrast, the responses to the game do not appear to differ between the experimental conditions, neither overall nor within gender.

[Insert Figure 6 about here]

Figure 6: Field experiment 1; differences in treatment elicited by the Islamic headscarf

Field Experiment 2: male confederate replays the “hijab-gaze”

P3b: The differences in treatment to which Muslims are subject in face-to-face encounters will cause them stress.

Building on the results of Field Experiment 1, P3b reads more precisely: The pattern of gaze that male passengers directed to the female confederate when she was wearing the Islamic headscarf is stressful for the recipient.

Regardless of the experimental condition, on a 9-point scale the average score is [0.17, 1.28] for negative affect, [6.25, 7.22] for involvement and [7.78, 8.75] for politeness. Confirming P3b (see Figure 7), being the target of the hijab-gaze increases women's negative affect by [0.81, 2.63], with effect size $d=[0.42, 1.37]$. Similarly, the hijab-gaze leads to a [-1.94, -0.08] decrease in the level of interpersonal involvement attributed to the confederate, with $d=[-1.56, -0.06]$. Although only credible at a less stringent $\alpha=0.10$, participants similarly reported that the hijab-gaze was on average less polite than the control gaze.

[Insert Figure 7 about here]

Figure 7: Reaction of the average woman to the different pattern of gaze to which headscarf-wearing women are exposed

Discussion

This article focused on a particular stressor: interactional discrimination, and a particular social group: Muslims in Europe. Drawing on the minority stress model, the prediction under examination was that Muslims' surplus depression is at least partly imputable to interactional discrimination in everyday life. The concept of interactional discrimination provides an explicit and positive formulation for the "distal" or objective discrimination component within the minority stress model.

Using hierarchical models estimated with Bayesian inference, the ESS data inspected in the epidemiological study showed that Muslims are more depressed than average and that this excess in

distress is partly mediated by interactional discrimination. Incorporating varying intercepts and slopes, the models also indicated that highly religious women living in France, who cover nearly 30% of the total sample of Muslim women, represent the group by gender and country that best exemplifies the hypothesis that Muslims' excess in depression is partly mediated by interactional discrimination.

Moving from self-reported experiences of discrimination to behavioral measures in real-life settings, two field experiments examined this causal hypothesis within the realm of everyday interpersonal encounters. Using a version of the "ultimatum game" in a public place of Paris, Experiment 1 assessed whether a female confederate is treated differently in a brief face-to-face exchange with a stranger when she signals high Islamic religiosity by wearing a headscarf. Replicating the result of another experiment (MASKED), the study found that male passengers do diminish the intensity of eye contact in response to the headscarf.

Using the same procedure in the same setting, Experiment 2 investigated the average effect, for women, of being the target of this less intense pattern of gaze, as performed by a male confederate. The results revealed that this look causes stress and the effect size statistic (whose most probable value nears one standard deviation) underscored that this seemingly harmless pattern of gaze actually functions to elicit a powerful negative reaction in the recipient. Together, the studies support the hypothesis that Muslim women in France are more depressed than baseline partly because they are more exposed, in situations of day-to-day living, to interactional discrimination on the part of men.

Three analytical benefits of the concept of interactional discrimination

Whenever types of discrimination have been distinguished in the past, there has been a tendency to dichotomize the phenomenal field into two categories, setting discrimination in the labour market (especially as it can be measured with correspondence studies) as the point of reference. The other

category has been traditionally constructed as a residue, that is to say, as the type of discrimination that is left once discrimination in the labor (or credit, or real estate, etc.) market has been taken out. This presupposes a prior definition of what is essential to this paradigmatic type of discrimination, and the usual approach has been to locate this essence in the economic nature of the disadvantage at stake. By opposition, the residual category has tended to cover forms of discrimination whose detrimental consequence is non-economic. Some prominent instances of discrimination types constructed with this residual logic are the ideas of “interpersonal discrimination” (Hebl et al. 2002), “microaggression” (Sue et al. 2007) and “assault on self-worth” (Lamont et al. 2016). The latter two in particular are centrally concerned with the experience of indignity, as the counterpart of economic disadvantage.

The notion of interactional discrimination is not predicated on the opposition between the economic and the non-economic. It arises rather from the acknowledgement of a simple fact: a significant portion of the phenomena covered by those residually constructed categories manifests in the realm of face-to-face interaction. Whatever else may mark out these phenomena, they are also characterized by the setting or context within which the act of discrimination is embedded, namely a spatio-temporally bounded episode of social interaction.

When the criterion for constructing types of discrimination is shifted from the nature of the adverse consequence (economic vs. non-economic) to the setting or context of discrimination (interactional vs. non-interactional), two implications follow. First, nothing prevents an act of discrimination embedded in an episode of social interaction from being economic in its detrimental consequence. Bargaining, for instance, represents a form of face-to-face interaction where the occurrence of discrimination may clearly impose economic disadvantage on the recipient. Second, the form of economic disadvantage widely documented by correspondence studies is notoriously non-interactional: when the decision is made not to call back the discriminated applicant, the decision is obviously not made in the presence of the applicant.

But in the research reported here the form of disadvantage under scrutiny, namely stress, is non-economic. Why bother to introduce new concepts if already established categories such as “micro-aggression” or “assault on self-worth” can surely do an equivalent job? The answer is that they actually do not. These categories are not meant to cover *every* form of disadvantage arising in the course of an interactional encounter, but only a quite specific sub-class: acts by others that victims interpret as attacks on their dignity. This is a very demanding definition indeed. It is not difficult to argue that a person may well be discriminated in an interpersonal encounter without realizing it, or experience the discriminatory act yet without invoking the concept of dignity.

In the realm of experience, stress manifests as negative emotion. Interpreting others’ behavior as an attack on one’s dignity is not necessary for that behavior to arouse a negative emotion. In other words, the conceptually mediated experience of indignity is not necessary for stress to occur, and more generally conceptual thought is not necessary for emotions to be elicited (Lambie and Marcel 2002). The essentially non-reflective experience of negative emotion does not necessarily have a correlate in the essentially reflective experience of indignity.

The concept of interactional discrimination thus brings three immediate benefits: 1) it subsumes “micro-aggressions” and “assaults on self-worth” as specific sub-types concerned with lived indignity, 2) it makes theoretical room for discrimination that is interactional *and* economic (e.g. bargaining), and 3) it carves out an analytical niche for stress as a detrimental consequence of discrimination that is *neither* economic *nor* necessarily concerned with the experience of indignity (e.g. the stressful experience of the “hijab-gaze”).

These three types of interactional discrimination have been empirically investigated in the research presented above. The “hijab-gaze” could have been construed as a micro-aggression if it had been rated by its recipients as impolite, disrespectful, or improper. But it was not. After bargaining, participants in the first experiment could have discriminated the hijab-wearing confederate on the economic dimension by taking advantage of their privilege as “dictators” in the simplified ultimatum game. But they did not. Nonetheless, male participants did discriminate the

hijab-wearing confederate in that they looked at her in a way that on average elicits strong negative emotions in the recipient. Without the concept of interactional discrimination, this subtle differential treatment and its powerful detrimental outcome would have gone unnoticed.

On mixing survey and field experimental data

One classical way of mixing methods in sociology consists in combining surveys with fieldwork (Sieber 1973). The entrenchment of the qualitative-quantitative divide in the discipline suggests an association between fieldwork and a non-quantitative focus on the actor's perspective, but when consideration of methods is divorced from ontological-epistemological debates (Pearce 2012) nothing prevents sociological fieldwork from consisting in the quantification of objective measures of behavior.

In the present investigation, the survey and experimental data were integrated in three distinct ways (Greene, Caracelli, and Graham 1989). First, both sources were triangulated to corroborate the common hypothesis that highly religious women in France are subject to interactional discrimination. Second, the strengths of one source were used to offset the weaknesses of the other. Thus survey data were used to produce representative estimates at the pan-European level, which is beyond the reach of a field experiment, and also to measure depressive symptoms as they naturally occur, which it would have been unethical to provoke experimentally. Symmetrically, experimental data were used to test causal statements rigorously, which is beyond the capabilities of a dataset relying on cross-sectional observational measurements, and also to produce measures of objective discriminatory behavior, which is out of reach when information is acquired with a questionnaire. Third, the analysis of the data stemming from the European Social Survey sequentially informed the design of the field experiments, by showcasing highly religious women residing in France as the subsample that provided the strongest *prima facie* evidence in support of the hypothesized causal pathway.

Conclusion

The present article demonstrates the heuristic value of the concept of interactional discrimination and of the epidemiological-experimental mix of methods. Interactional discrimination represents a concept that is at once precise enough to designate a spatio-temporally bounded “micro” process and flexible enough to accommodate any form of disadvantage, economic, emotional, or other. As for the mix, the empirical analysis shows that, in explaining mental health disparities via discrimination, surveys and field experiments can be fruitfully combined for corroborating hypotheses, making strategic design decisions, and offsetting the weaknesses of one method with the strengths of the other.

References

- Adida, Claire L., David D. Laitin, and Marie-Anne Valfort. 2010. “Identifying Barriers to Muslim Integration in France.” *Proceedings of the National Academy of Sciences* 107(52):22384–90. doi: 10.1073/pnas.1015550107.
- Aidenberger, Amelie, and Malte Doehne. 2021. “Unveiling Everyday Discrimination. Two Field Experiments on Discrimination against Religious Minorities in Day-to-Day Interactions.” *British Journal of Sociology* 00:1–19. doi: 10.1111/1468-4446.12811.
- Bertrand, Marianne, and Esther Duflo. 2017. “Field Experiments on Discrimination.” in *Handbook of field experiments*. Vol. 1, *Handbooks in economics*, edited by A. V. Banerjee and E. Duflo. Elsevier.
- Camerer, Colin F. 2003. *Behavioral Game Theory: Experiments in Strategic Interaction*. Princeton (New Jersey): Princeton University Press.
- Choi, Donghyum Danny, Mathias Poertner, and Nicholas Sambanis. 2020. “Linguistic Assimilation Does Not Reduce Discrimination against Immigrants: Evidence from Germany.” *Journal of Experimental Political Science* 1(12). doi: 10.1017/XPS.2020.20.
- Coleman, James S. 1986. “Social Theory, Social Research, and a Theory of Action.” *American Journal of Sociology* 91(6):1309–35. doi: 10.1086/228423.
- Di Stasio, Valentina, Bram Lancee, Susanne Veit, and Ruta Yemane. 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):1305–26. doi: 10.1080/1369183X.2019.1622826.
- Diekmann, Andreas, Ben Jann, and Matthias Näf. 2014. “Wie Fremdenfeindlich Ist Die Schweiz? Fünf Feldexperimente Über Prosoziales Verhalten Und Die Diskriminierung von Ausländer in Der Stadt Zürich Und in Der Deutschschweiz” *Soziale Welt* 65:185–99.
- Dohrenwend, Bruce P. 2006. “Inventorying Stressful Life Events as Risk Factors for Psychopathology: Toward Resolution of the Problem of Intracategory Variability.” *Psychological Bulletin* 132(3):477–95. doi: 10.1037/0033-2909.132.3.477.
- Ellison, Christopher G., and Andrea K. Henderson. 2011. “Religion and Mental Health: Through the Lens of the Stress Process.” Pp. 11–44 in *Toward a sociological theory of religion and*

- health*. Vol. 19, *Religion and the social order*, edited by A. J. Blasi. Leiden and Boston: Brill.
- European Commission. 2019. *Special Eurobarometer 493. Report on Discrimination in the European Union*. European Union.
- European Values Study. 2017. *European Values Study 2017: Integrated Dataset (EVS 2017)*. Cologne (Germany): GESIS Data Archive.
- Exline, R. V., S. L. Ellyson, and B. Long. 1975. "Visual Behavior as an Aspect of Power Role Relationships." Pp. 21–52 in *Advances in the Study of Communication and Affect*. Vol. 2, edited by P. Pilner, L. Krames, and T. Alloway. New York: Plenum Press.
- FRA. 2017. *Second European Union Minorities and Discrimination Survey. Muslims-Selected Findings*. European Union Agency for Fundamental Rights.
- Gelman, Andrew, John B. Carlin, Hal S. Stern, David B. Dunson, Aki Vehtari, and Donald B. Rubin. 2013. *Bayesian Data Analysis*. 3rd ed. Boca Raton, London, New York: CRC Press.
- Goffman, E. 1963. *Behavior in Public Places: Notes on the Social Organization of Gatherings*. New York: The free press.
- Greene, Jennifer, Valerie Caracelli and Wendy Graham. 1989. "Toward a conceptual framework for mixed-method evaluation designs." *Educational evaluation and policy analysis* 11(3): 255-274. doi: 10.3102/01623737011003255
- Harrell, Shelly P. 1997. "The Racism and Life Experiences Scales (RaLES): Daily Life Experiences (Racial Hassles)."
- Hayes, Andrew F., and Klaus Krippendorff. 2007. "Answering the Call for a Standard Reliability Measure for Coding Data." *Communication Methods and Measures* 1(1):77–89.
- Hebl Michelle, Jessica Bigazzi-Foster, Laura Mannix and John Dovidio. 2002. "Formal and interpersonal discrimination: a field study of bias toward homosexual applicants." *Personality and Social Psychology Bulletin* 28(6): 815-825. doi: 10.1177/0146167202289010
- Hedström, Peter, and Petri Ylikoski. 2010. "Causal Mechanisms in the Social Sciences." *Annual Review of Sociology* 36(1):49–67. doi: 10.1146/annurev.soc.012809.102632.
- Koenig, Harold G. 2009. "Research on Religion, Spirituality, and Mental Health: A Review." 54(5):283–91. doi: <https://doi.org/10.1177/070674370905400502>.
- Koopmans, Ruud, Susanne Veit, and Ruta Yemane. 2019. "Taste or Statistics? A Correspondence Study of Ethnic, Racial and Religious Labour Market Discrimination in Germany." *Ethnic and Racial Studies* 42(16):233–52. doi: 10.1080/01419870.2019.1654114.
- Kruschke, John K. 2015. *Doing Bayesian Data Analysis*. London: Academic Press.
- Lambie, John A., and Anthony J. Marcel. 2002. "Consciousness and the Varieties of Emotion Experience: A Theoretical Framework." *Psychological Review* 109(2):219–59. doi: 10.1037/0033-295X.109.2.219.
- Lamont, Michèle, G. M. Silva, J. Welburn, J. Guetzkow, N. Mizrahi, H. Herzog, and E. Reis. 2016. *Getting Respect: Responding to Stigma and Discrimination in the United States, Brazil, and Israel*. Princeton and Oxford: Princeton University Press.
- Landrine, Hope, and Elizabeth A. Klonoff. 1996. "The Schedule of Racist Events: A Measure of Racial Discrimination and a Study of Its Negative Physical and Mental Health Consequences." *Journal of Black Psychology* 22(2):144–68. doi: 10.1177/00957984960222002
- Lewis, Tené T., Courtney D. Cogburn, and David R. Williams. 2015. "Self-Reported Experiences of Discrimination and Health: Scientific Advances, Ongoing Controversies, and Emerging Issues." *Annual Review of Clinical Psychology* 11(1):407–40. doi: 10.1146/annurev-clinpsy-032814-112728.
- Lippert-Rasmussen, Kasper. 2006. "The Badness of Discrimination." *Ethical Theory and Moral Practice* 9(2):167–85.
- MacKinnon, David P., Amanda J. Fairchild, and Matthew S. Fritz. 2007. "Mediation Analysis." *Annual Review of Psychology* 58:593. doi: 10.1146/annurev.psych.58.110405.085542.
- McCullough, Michael E., and David B. Larson. 1999. "Religion and Depression: A Review of the Literature." *Twin Research and Human Genetics* 2(2):126–36. doi: 10.1375/twin.2.2.126.

- Hughes, Michael, K. Jill Kiecolt, Verna M. Keith, and David H. Demo. 2015. "Racial Identity and Well-Being among African Americans." *Social Psychology Quarterly* 78(1):25–48. doi: 10.1177/0190272514554043.
- Meyer, Ilan H. 2003. "Prejudice, Social Stress, and Mental Health in Lesbian, Gay, and Bisexual Populations: Conceptual Issues and Research Evidence." *Psychological Bulletin* 129(5):674–97. doi: 10.1037/0033-2909.129.5.674.
- Oyarvide Tuthill, Zelma. 2021. "The Intersection of Sexual and Racial/Ethnic Identity Centrality and Mental Well-Being among Black and Latinx Sexual Minority Adults." *Social Psychology Quarterly* 84(4):331–52. doi: 10.1177/01902725211037637.
- Pager, Devah, and Hana Shepherd. 2008. "The Sociology of Discrimination: Racial Discrimination in Employment, Housing, Credit, and Consumer Markets." *Annual Review of Sociology* 34:181–209.
- Paradies, Yin, Jehonathan Ben, Nida Denson, Amanuel Elias, Naomi Priest, Alex Pieterse, Arpana Gupta, Margaret Kelaher, and Gilbert Gee. 2015. "Racism as a Determinant of Health: A Systematic Review and Meta-Analysis." *PLOS ONE* 10(9):e0138511. doi: 10.1371/journal.pone.0138511.
- Pearce, Lisa. 2012. "Mixed Methods Inquiry in Sociology." *American Behavioral Scientist* 56(6): 829–848. doi: 10.1177/0002764211433798
- Pearlin, Leonard I. 1989. "The Sociological Study of Stress." *Journal of Health and Social Behavior* 30(3):241–56. doi: 10.2307/2136956.
- Pew Research Center. 2006. "The French-Muslim Connection." *Pew Research Center*. Retrieved March 12, 2021 (<https://www.pewresearch.org/2006/08/17/the-frenchmuslim-connection/>).
- Pew Research Center. 2015. "Five Facts about the Muslim Population in Europe."
- Pierné, Guillaume. 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.
- Radloff, L. S. 1977. "The CES-D Scale: A Self-Report Depression Scale for Research in the General Population." *Applied Psychological Measurement* 1(3):385–401. doi: <https://doi.org/10.1177/014662167700100306>.
- Richeson, Jennifer A., and Samuel R. Sommers. 2016. "Toward a Social Psychology of Race and Race Relations for the Twenty-First Century." *Annual Review of Psychology* 67(1):439–63. doi: 10.1146/annurev-psych-010213-115115.
- Rosenberg, Morris. 1962. "The Dissonant Religious Context and Emotional Disturbance." *American Journal of Sociology* 68(1):1–10. doi: 10.1086/223261.
- Schnittker, Jason. 2020. "Religion, Social Integration, and Depression in Europe: Evidence from the European Social Survey." *Social Science & Medicine* 267:112376. doi: 10.1016/j.socscimed.2019.112376.
- Schwartz, Sharon, and Ilan H. Meyer. 2010. "Mental Health Disparities Research: The Impact of within and between Group Analyses on Tests of Social Stress Hypotheses." *Social Science & Medicine* 70(8):1111–18. doi: 10.1016/j.socscimed.2009.11.032.
- Sieber, Sam. 1973. "The integration of fieldwork and survey methods." *American Journal of Sociology* 78(6): 1335–1359. doi: 10.1086/225467
- Sloetjes, H., and P. Wittenburg. 2008. "Annotation by Category - ELAN and ISO DCR." in *Proceedings of the 6th International Conference on Language Resources and Evaluation*.
- Sue, D., C. M. Capodilupo, G. C. Torino, J. M. Bucceri, A. Holder, K. L. Nadal, and M. Esquilin. 2007. "Racial Microaggressions in Everyday Life: Implications for Clinical Practice." *American Psychologist* 62(4):271–86.
- Tajfel, Henri, M. G. Billig, R. P. Bundy, and Claude Flament. 1971. "Social Categorization and Intergroup Behaviour." *European Journal of Social Psychology* 1(2):149–78. doi: <https://doi.org/10.1002/ejsp.2420010202>.
- Thomeer, Mieke Beth, Allen J. LeBlanc, David M. Frost, and Kayla Bowen. 2018. "Anticipatory Minority Stressors among Same-Sex Couples: A Relationship Timeline Approach." *Social*

- Psychology Quarterly* 81(2):126–48. doi: 10.1177/0190272518769603.
- Valfort, Marie-Anne. 2020. “Anti-Muslim Discrimination in France: Evidence from a Field Experiment.” *World Development* 135:105022. doi: 10.1016/j.worlddev.2020.105022.
- Williams, David R., Jourdyn A. Lawrence, Brigette A. Davis, and Cecilia Vu. 2019. “Understanding How Discrimination Can Affect Health.” *Health Services Research* 54(S2):1374–88. doi: 10.1111/1475-6773.13222.
- Williams, Rory, and Kate Hunt. 1997. “Psychological Distress among British South Asians: The Contribution of Stressful Situations and Subcultural Differences in the West of Scotland Twenty-07 Study.” *Psychological Medicine* 27(5):1173–81. doi: 10.1017/S0033291797005473.

Figure 1

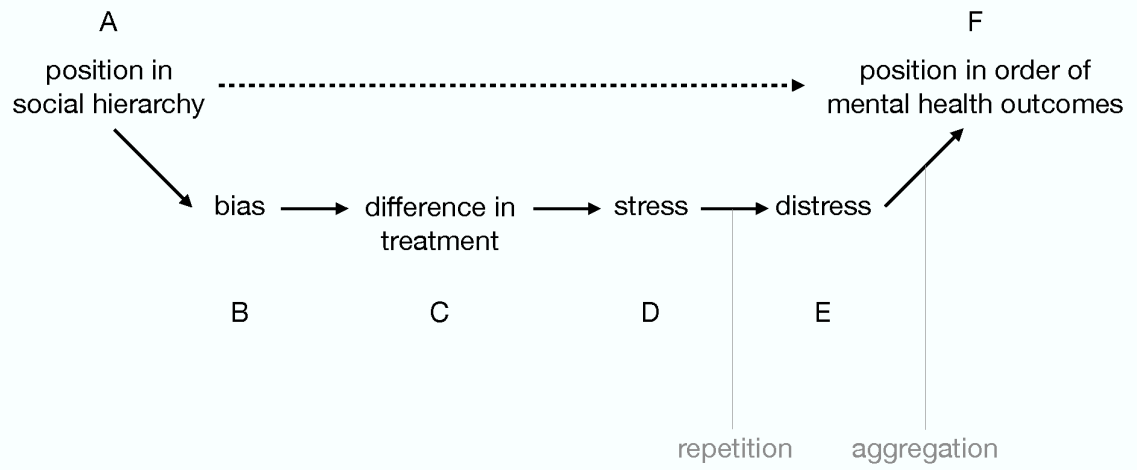


Figure 2

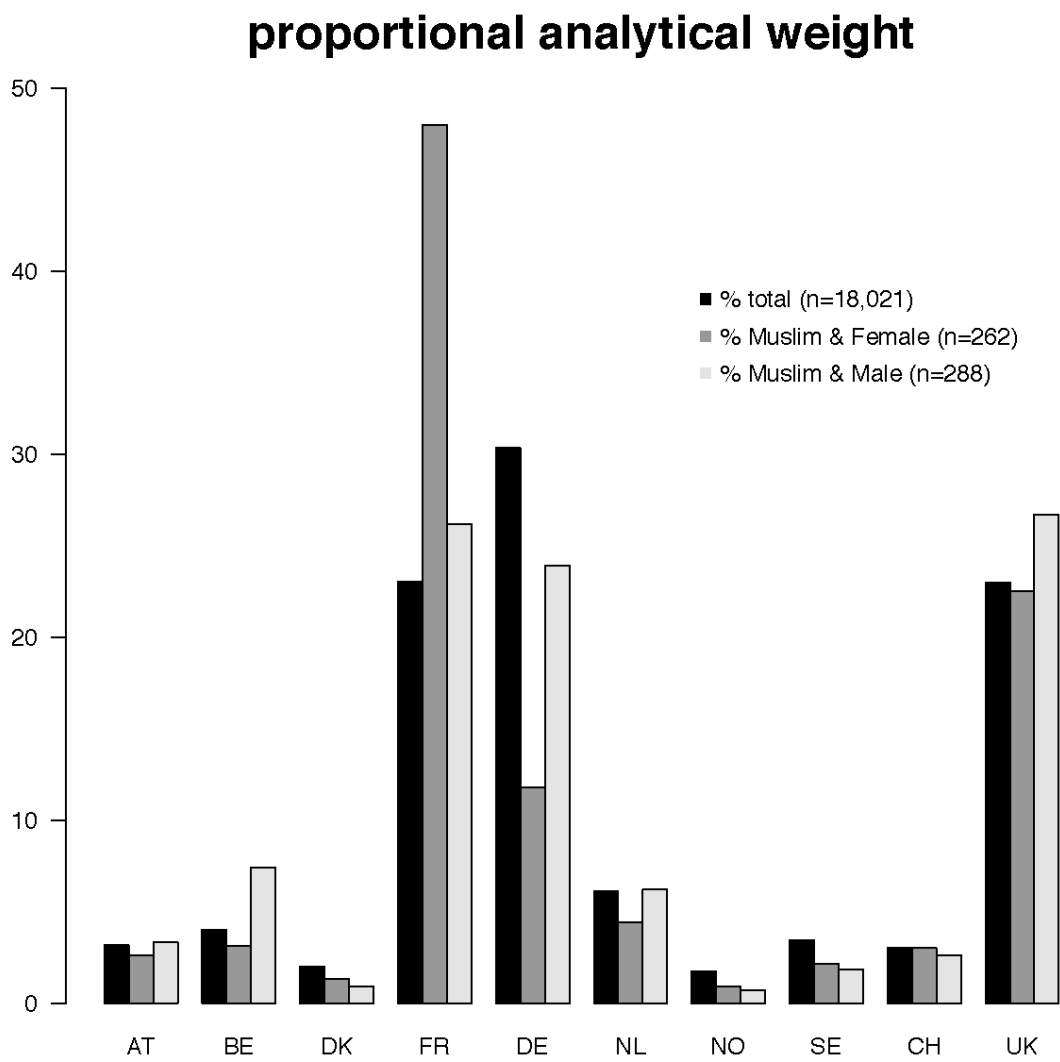


Figure 3

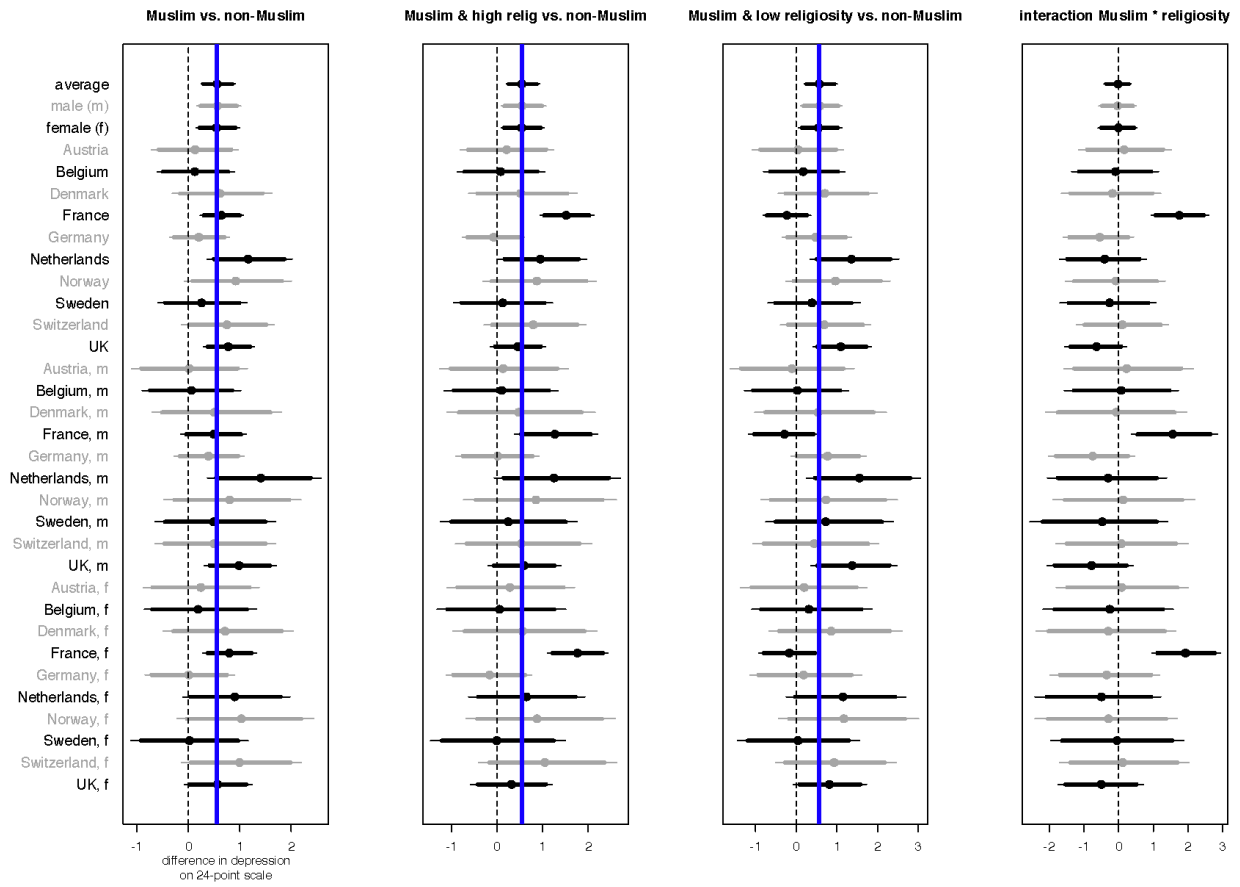


Figure 4

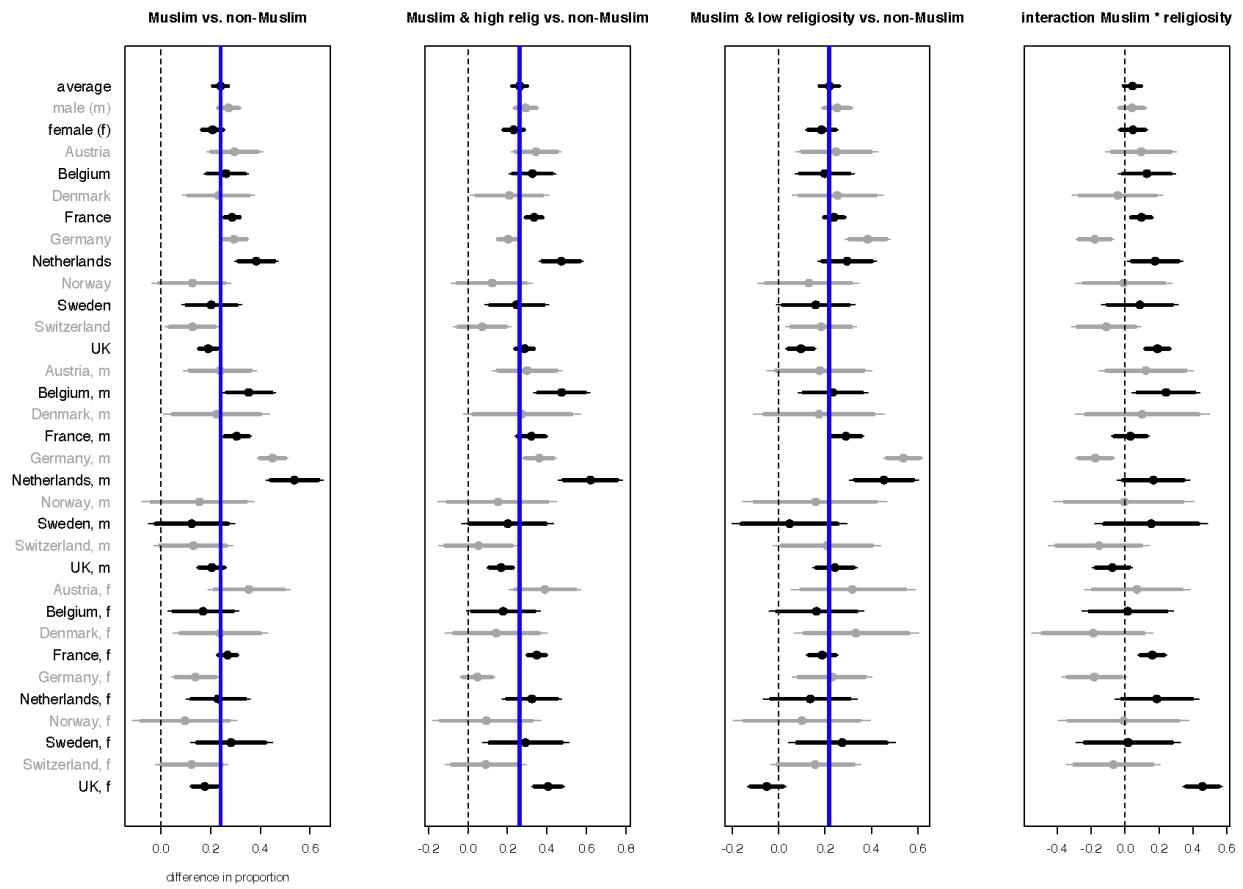


Figure 5

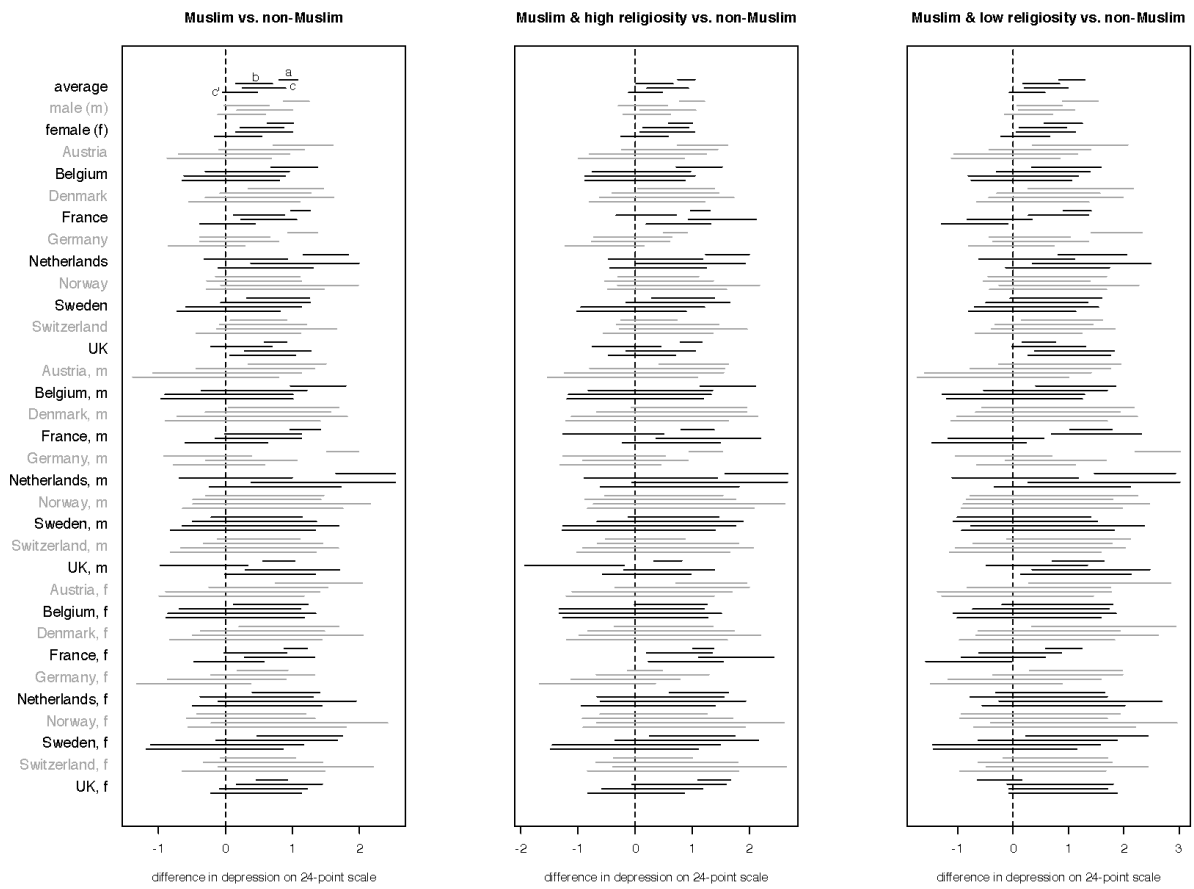


Figure 6

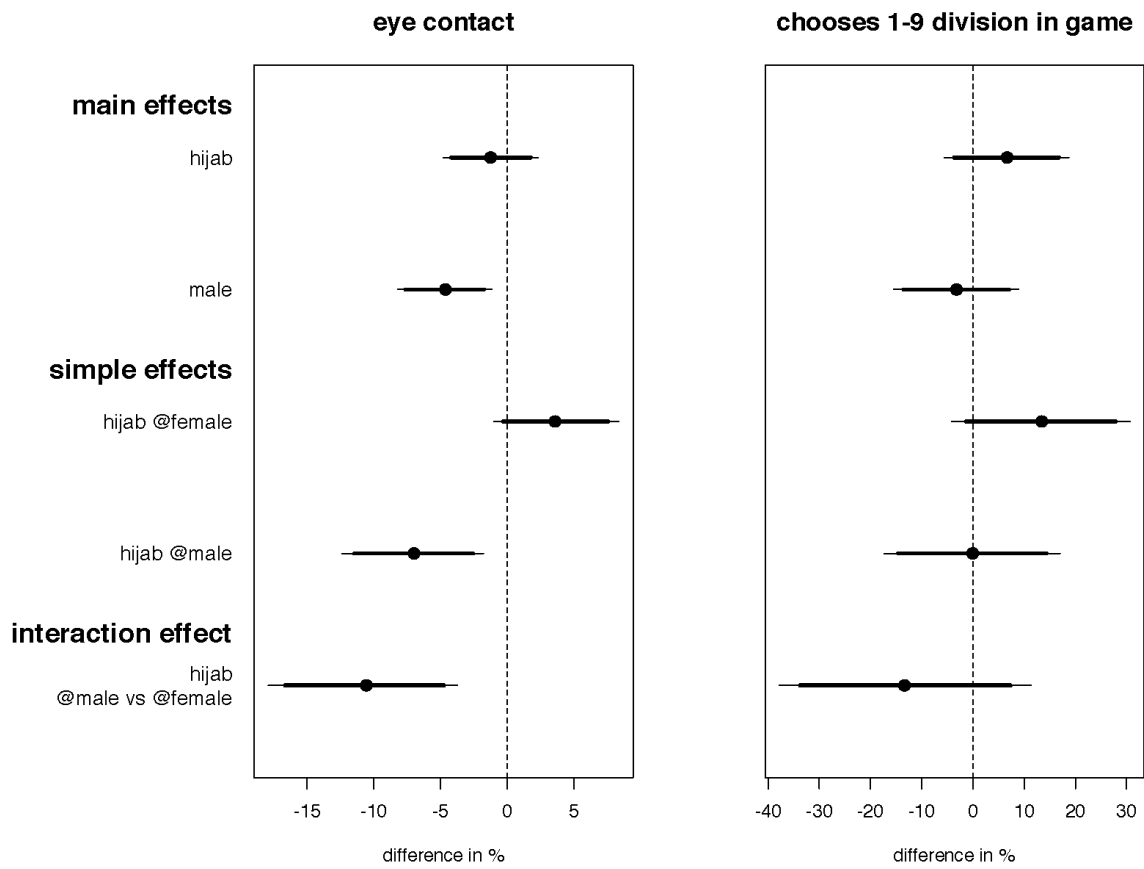


Figure 7

