

## Gender-inclusive language as a Rational Speech Act in Spanish

Jeremy Yeaton, María Muelas-Gil & Gregory Scontras\*

**Abstract.** Amidst social changes in gendered language use, there is pushback from institutions such as the Spanish Royal Academy, which claims that the use of the generic masculine (e.g., *bomberos* ‘firemen’) in describing a mixed-gender group is equally inclusive of both men and women (Bosque 2012). By contrast, speakers of Spanish have increasingly adopted gender-inclusive alternatives to the generic masculine (e.g., *bomberos o bomberas*; Bengoechea 2015). Across two behavioral tasks, we investigated whether gender-inclusive forms actually lead to more inclusive interpretations. We found that the use of the inclusive form (by contrast to the generic masculine) indeed yields more inclusive interpretations, increasing the inferred femaleness of stereotypically male professions, but also decreasing the inferred femaleness of stereotypically female professions. In an attempt to explain the reasoning that delivers inclusive interpretations, we developed a computational cognitive model of the reasoning process. Our model treats the phenomenon as an instance of a markedness implicature: speakers use the longer, inclusive form to guide listeners away from their prior expectations. This work highlights the need for further research into the use of gender-inclusive language cross-linguistically, as well as for pushback against prescriptive institutions perpetuating stereotypes.

**Keywords.** gender stereotypes; Spanish; disjunction; Rational Speech Act framework; markedness implicature

**1. Introduction.** Consider the following headline from a news piece reporting on a group of Spanish health researchers who developed a virtual coach to prevent depression (*La Gaceta*, 10/04/2021, p. 4):

- (1) Médicos salmantinos crean un entrenador virtual que previene la  
 doctor-M-PL from-Salamanca-M-PL create a coach virtual that prevents the  
 depression laboral.  
 depression from-work  
 ‘Doctors from Salamanca create a virtual coach that prevents depression at work.’

In example (1), there is no explicit indication of the gender of the doctors involved in the project when observing the English translation; however, this gender information appears in the original text in Spanish in the form of masculine plural—the so-called “masculine generic”. Spanish has grammatical gender, where nouns are typically marked for gender (mainly final *-o* for masculine, *-a* for feminine) and can therefore carry an additional cue to referential gender (Reali et al. 2015). Following the prescriptive norm that sets that the masculine form is to be used as the generic one when both genders are included or when the distribution is unknown (*Real Academia de la Lengua Española* ‘Spanish Royal Academy’, RAE henceforth), there is no overt indication of whether the people involved in the project described in (1) were male or female, since the masculine generic must be used to include both genders when applicable. What made this headline

\* This work was supported by Ayudas UAM-SANTANDER para la movilidad de jóvenes investigadores awarded to MMG. Authors: Jeremy Yeaton, UC Irvine ([jyeaton@uci.edu](mailto:jyeaton@uci.edu)), María Muelas-Gil, Universidad Autónoma de Madrid ([maria.muelas@uam.es](mailto:maria.muelas@uam.es)) & Gregory Scontras, UC Irvine ([g.scontras@uci.edu](mailto:g.scontras@uci.edu)).

relevant here is the fact that, out of the four people involved in the team of doctors, three of them were female and only one of them was male.

While this example may seem to invoke a simple linguistic norm, what it reflects is of great importance for linguistic studies. In spite of the institutional calls to follow the norm as it is inclusive of all genders (Bosque 2012), many studies have shown that the use of the masculine generic in fact implies a lower proportion of women among the referent (Stahlberg et al. 2001, 2007; Sczesny et al. 2016; Menegatti et al. 2017). With this consideration in mind, scholars have begun to investigate the consequences of using the masculine generic in comparison to other more inclusive alternatives, such as slash forms (*médicos/as*), disjunctions (*médicas o médicos*), or even emerging options like the ending vowel *-e* (*médiques*) or the consonant *-x* (*médicxs*).

With the intention of continuing this line of research, this study carried out a behavioral study to confirm whether speakers do indeed project a higher proportion of women when using the alternative, gender-inclusive forms in comparison to the normative masculine generic. We then developed a computational cognitive model of the reasoning process that captures the observed behavior.

The results of our preliminary study show that a higher proportion of women is indeed imagined when using the gender-inclusive form compared to the masculine generic option, but this effect is limited to stereotypically masculine jobs (e.g., trucker, firefighter); however, this trend is reversed for the jobs that are stereotypically feminine in society (such as nurse or librarian), meaning fewer women (or more men) are included via the inclusive forms. This finding is in line with previous studies claiming that the use of inclusive language mitigates the effect of stereotypes regardless of which gender is involved; thus, the study supports previous claims that gender-inclusive language offers a useful tool for counteracting stereotypes.

Moreover, in an attempt to explain the inclusive effect of gender-inclusive language, we also use a computational cognitive model formulated within the Bayesian Rational Speech Act framework (Frank & Goodman 2012; Goodman & Stuhlmüller 2013; Scontras et al. 2018) to implement our hypothesis regarding the reasoning that yields inclusive interpretations. In our model, inclusive forms are more costly to produce relative to the masculine generic, and listeners reason that a speaker who produces the costlier inclusive form meant to convey a marked interpretation that deviates from the stereotypes involved. Finding a reliable qualitative fit and a moderate quantitative fit between our model's predictions and the behavior in our experiment, we have evidence in favor of this view of inclusive language use.

In what follows, we describe the experiments and model in more detail, after first introducing additional background in the next section.

**2. Background.** Stereotypes can be defined as a preconceived mental generalization that humans have and use to simplify, categorize, and somehow facilitate the perception of everything that surrounds them (Hamilton & Trolie 1986). Stereotyping is a cognitive process and an ability developed from an early age, influenced and shaped by the familial, educational, and social context (Jussim et al. 1996; Martin & Ruble 2010; Signorielli 2001). More importantly, stereotypes also influence our actions and beliefs towards the world from which they arise. In fact, research has demonstrated that all kinds of stereotypes (age, gender, group belonging, race, etc.) confer fast and simple assumptions that affect humans' behaviour towards other members of social groups (Schneider 2005; Yzerbyt & Demoulin 2010). They are also an inherent part of culture in its broadest sense, and, as Guerra et al. (2021) put it, stereotypes "are in fact culturally trans-

mitted and thus they reflect social biases rather than plausibility” (Guerra et al. 2021:2; Brown 2011).

When referring to culturally-transmitted values and patterns in society, one of the means through which they get carried forward is language. Humans’ first contact with language is right after birth (or perhaps even before; Moon et al. 1993) when children encounter adults’ use of language. Although our tendency to socially categorize the world around us allows for a quick and efficient understanding and organization of the environment, this quick categorization does tend to foster generalizations, prejudices, and discrimination on the basis of stereotypes. It is precisely this type of discrimination on the basis of gender that motivated the present study. Social role theory proposes that gender stereotypes reflect the distribution of men and women into social roles (Bosak et al. 2012; Eagly & Steffen 1986; Eagly et al. 2000; Wood & Eagly 2002) and affect speakers throughout their lives and in all spheres, including job choices. This effect has been claimed to be even more noticeable in societies whose languages carry an extra load of grammatical gender marking. In other words, languages that mark gender, or *grammatical gender languages* like Spanish, tend to be linked to a higher degree of gender discrimination in their cultures, in comparison to *natural gender languages*, where gender is not morphologically marked (UNESCO 2011; Prewitt-Freilino et al. 2012).

However, there is still resistance from both society in general and normative institutions in particular to acknowledging and addressing such influence and the potential relation between gender marking in a language and gender inequalities. Among such institutions, the RAE states that the masculine generic suffices to include all genders equally, adding that using alternatives like disjunctive forms is redundant and goes against language economy. There is already an extensive literature supporting the idea that gender-inclusive language strongly acts to prevent gender discrimination and should therefore be encouraged in languages such as English (Hamilton 1988; Gastil 1990), Italian (Cacciari & Padovani 2007), German (Braun et al. 2007; Stahlberg et al. 2001), or French (Gygax & Gabriel 2008). The present study aims at expanding the existing literature by examining gender-inclusive language in Spanish.

**3. Experiment 1: Assessing stereotypes.** For the first stage of the present study, a list of stereotypical jobs was drawn up based on previous studies that asked participants to rate the stereotypicality of professions (Carreiras et al. 1996; Kennison & Trofe 2003; Irmen & Kurovskaia 2010; Pyykkönen et al. 2009; Su et al. 2016, among others). Although these studies were extensive and some covered hundreds of nouns, we considered that social advancements in the last decade were significant enough to warrant updating the gender stereotypes for jobs. Moreover, only one recent study collected stereotype ratings for Spanish (Su et al. 2016).

3.1. PARTICIPANTS. Seventy-eight university students (64 females, 12 males, 1 non-binary and 1 would rather not say; aged 18-30, mean age 22.6) who spoke Spanish as a first language were asked to complete the task in exchange for course credit.

3.2. DESIGN AND METHOD. We created an initial list of 200 nouns representing different jobs. To select these jobs, we followed previous studies (Gabriel et al. 2008; Misersky et al. 2014) and Spanish job-terminology guides (Lledó 2006). However, after two rounds of consultation with experts, 8 jobs were removed from the initial list as they were either synonyms or almost synonyms of another member of the list, obsolete, or not expected to be recognized by the participants. The final list of professions contained 192 items, which were presented in singular and in

**Lista de profesiones**

A continuación vas a ver una lista de oficios o profesiones.  
 Por favor, indica en la escala si crees que en cada oficio hay un mayor número de trabajadores de género masculino (♂) o femenino (♀), siendo 1 = solo hombres, y 9= solo mujeres.  
 Importante: la respuesta no debe indicar "cuántos hombres o cuántas mujeres creo yo que DEBE O DEBERÍA haber", sino "cuántos o cuántas me consta que HAY a día de hoy".

---

arquitecto/a \*

	1	2	3	4	5	6	7	8	9	
♂	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	♀

Figure 1. Materials from Experiment 1: instructions (top) and an example trial (bottom).

the slash form (*profesor/a, médico/a...*). By means of a privately shared Google Forms link with the participants, they were asked to fill out the survey evaluating the listed jobs using a Likert scale ranging from 1 to 9, saying whether they thought that each job was carried out by mostly men (1 = only men), mostly women (9 = only women) or equally by both (5 = 50/50). Participants were reminded that they should answer giving the value they consider to be a reflection of the reality in Spain and not their desired or ideal distribution. Figure 1 provides a screenshot of the instructions and an example trial. Participants provided a total of 192 ratings, one for each profession.

3.3. RESULTS. Figure 2A plots a histogram of mean stereotype scores for each of the 192 professions, calculated by averaging across participants' ratings. There we see that the mean stereotype scores for the surveyed professions cover nearly the full range of the Likert scale (min = 1.31, very masculine; max = 8.44, very feminine) with  $\mu = 4.38$  and  $\sigma = 1.69$ . Figure 3 plots the distribution of raw ratings for six representative example professions: two stereotypically masculine (trucker and firefighter), two stereotypically feminine (nurse and librarian), and two without a clear bias (baker and reporter). We use average stereotype scores to characterize our experimental items in the next task, and we use the response distributions later in our model to serve as prior expectations for gender breakdown by profession.

**4. Experiment 2: Evaluating gender-inclusive language.** Having collected baseline stereotypes for a list of professions, the next question to address was whether gender-inclusive forms—specifically, a disjunctive form as opposed to the masculine generic—indeed lead to more inclusive interpretations. We re-used the rating methodology from Experiment 1, but this time participants either provided responses for the noun in the masculine generic or as a disjunction between masculine and feminine plural forms.

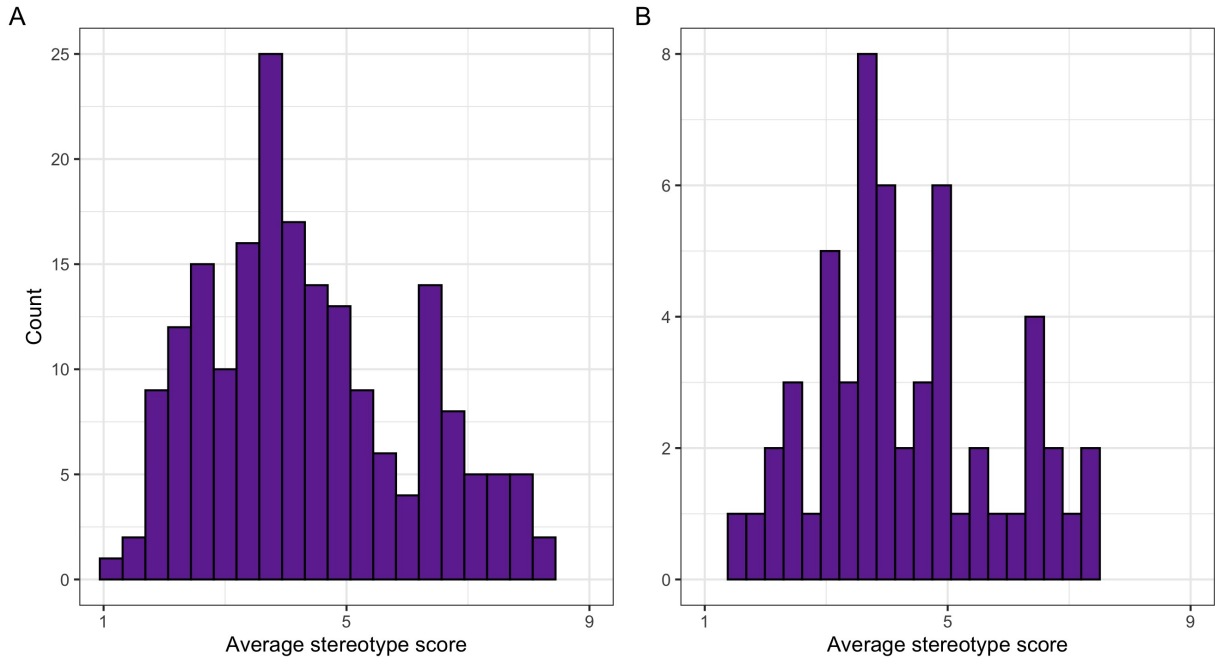


Figure 2. Distribution of average stereotype scores for professions surveyed in Experiment 1. The x-axis plots mean stereotype scores ranging from 1 (completely male) to 9 (completely female). A: All 192 professions surveyed; B: Only the 55 professions which were used in Experiment 2.

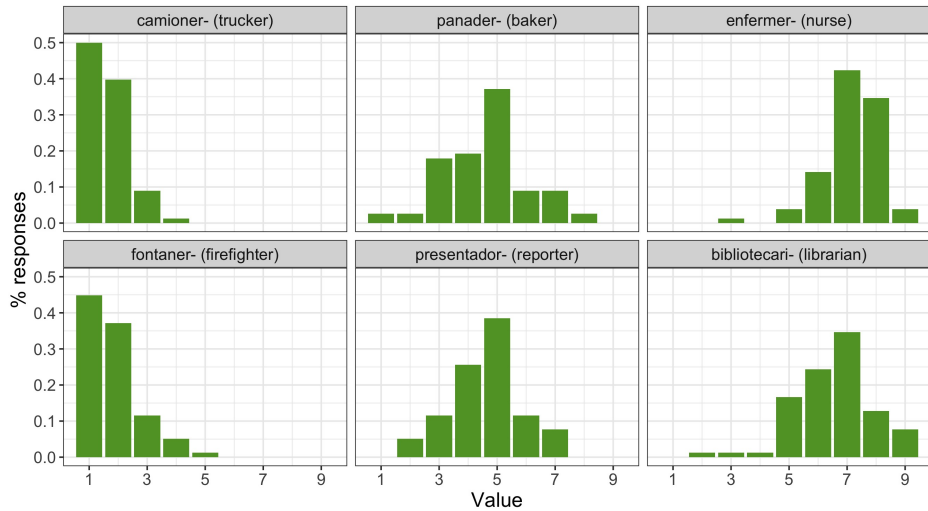


Figure 3. Response distributions for representative professions (two male-stereotyped, two neutral, and two female-stereotyped) in Experiment 1.

fontaneros \*

1 2 3 4 5 6 7 8 9

♂ ○ ○ ○ ○ ○ ○ ○ ○ ○ ♀

enfermeros \*

1 2 3 4 5 6 7 8 9

♂ ○ ○ ○ ○ ○ ○ ○ ○ ○ ♀

fontaneras o fontaneros \*

1 2 3 4 5 6 7 8 9

♂ ○ ○ ○ ○ ○ ○ ○ ○ ○ ♀

enfermeras o enfermeros \*

1 2 3 4 5 6 7 8 9

♂ ○ ○ ○ ○ ○ ○ ○ ○ ○ ♀

Figure 4. Four example trials from Experiment 2. Each participant saw either the masculine generic (top) or an inclusive form (bottom).

4.1. PARTICIPANTS. One hundred and three university students (84 females, 17 males, 2 non-binary; aged 18-30, mean age 20.2) who spoke Spanish as a first language were asked to complete the task in exchange for course credit. They were not informed about the final purpose of the project (i.e., analyzing the perception and processing of gender stereotypes) in order to avoid biased answers.

4.2. DESIGN AND METHOD. In order to assess whether the more-inclusive option (i.e., disjunction) did in fact foster a higher projection of women than the masculine generic form, 60 items were randomly selected out of the 200-item original list used in Experiment 1 to create the survey. Once participants had added information on their gender, age, level of English as an L2, level of studies of parent(s) and knowledge of any other languages, the participants had to randomly select among three options (A, B, C) to start the survey; depending on which option they selected, they would be taken to the experiment itself (with option A leading to masculine generics, B to the slash form and C to the disjunction). All three options contained exactly the same 60 jobs presented in alphabetical order. Then, a similar process to Experiment 1 was followed: they were asked to fill out the survey evaluating the listed jobs using a Likert scale ranging from 1 to 9, saying whether they thought that each job was carried out by mostly men (1 = only men), mostly women (9 = only women) or equally by both (5 = 50/50). Participants were reminded that they should answer giving the value they consider to be a reflection of the reality in Spain and not their desired or ideal distribution. In the end, 38 participants completed option A, 41 did option B, and 24 option C. For the present study, only answers to option A (masculine generics) and C (disjunction) were considered, leaving the slash form for further research. Finally, although all participants rated the same 60 professions, five professions were removed from the final list prior to analysis due to potential ambiguity of the words (for instance, *química* might mean both ‘chemist’ and ‘chemistry’, or *cajera* might mean ‘cashier’ or ‘ATM’); Figure 2B plots the stereotype distribution as measured in Experiment 1 for these 55 professions.

4.3. RESULTS. For each of the 55 professions, we first calculated the difference between responses provided for the masculine generic and the disjunctive form: positive values indicate that participants responded with a greater proportion of women for the disjunctive relative to the

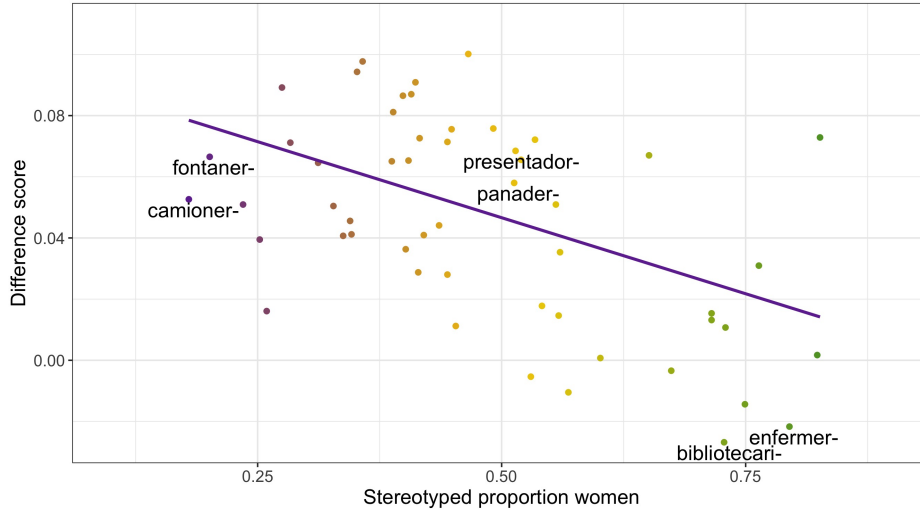


Figure 5. Behavioral difference score in Exp. 2 (difference in reported value for the masculine generic plural and the disjunctive forms) by stereotype measured in Exp. 1.

masculine generic form. Figure 5 plots these difference scores against the stereotyped proportion of women measured in Experiment 1. We found a significant negative correlation between the gender stereotype from Experiment 1 and the difference scores ( $r = -0.44, t = -3.61, df = 53, p < 0.001$ ; Figure 5). In other words, the more masculine a profession was stereotyped to be, the more women were judged to be included by the disjunctive form relative to the masculine generic. Conversely, the more feminine a profession was stereotyped to be, the more men were judged to be included by the disjunctive form relative to the masculine generic. The effect size for each profession is relatively small ( $\mu = 0.05, \sigma = 0.03$ ), but the relationship remains robust. We fit the following linear mixed-effect model to the data:

$$\text{JUDGMENT} \sim \text{STEREOTYPE} \times \text{FORM} \times (1|\text{PARTICIPANT}),$$

with the masculine generic as the baseline form. We found significant effects of STEREOTYPE, where a more feminine stereotype led to an overall higher proportion of women reported ( $\beta = 0.97, \hat{\sigma}_{\beta} = 0.02, t = 45.51, p < 0.001$ ), and FORM ( $\beta = 0.89, \hat{\sigma}_{\beta} = 0.18, t = 4.84, p < 0.001$ ), where the disjunctive form generally led to a higher proportion of women reported. We also found a significant interaction between the factors ( $\beta = -0.10, \hat{\sigma}_{\beta} = 0.03, t = -3.04, p < 0.01$ ), where a more feminine profession in the disjunctive form was reported to be *less* feminine than in the masculine generic form.

4.4. DISCUSSION. The results of Experiment 2 confirm that using the inclusive disjunctive form leads to interpretations that are more inclusive than those for the masculine generic. For masculine-stereotyped nouns, the disjunctive form led to an increased prevalence of women, and for feminine-stereotyped nouns, the disjunctive form led to an increased prevalence of men. This latter finding might at first sound counterintuitive: why should including the feminine plural form in addition to the masculine (generic) plural via disjunction (e.g., *enfermeras o enfermeros*) lead to an increased prevalence of men? We hypothesize that the disjunctive form serves to cue a markedness implicature, which invites the listener to deviate from their stereotypes. In

the case where the stereotype strongly favors women (i.e., for feminine-stereotyped nouns like *enfermeros/as*), deviating from the stereotype involves inferring more men, even though the cue for that inference is the feminine plural form of the noun included via disjunction. In the following section, we make more explicit our assumptions about how we believe this reasoning unfolds. For now we simply note that the results of the behavioral studies, especially Experiment 2, are in line with previous research highlighting greater inclusion of women with more inclusive forms in other languages; the results also demonstrate that using the inclusive forms benefit both male and female workers in stereotyped jobs, as they are more likely to be imagined as part of the referenced group.

**5. Our model of understanding gendered language.** Having documented that inclusive disjunctive forms indeed lead to more inclusive interpretations of profession nouns relative to the masculine generic, the task now is to explain how the inclusive interpretations arise. Our hypothesis is that the more cumbersome disjunctive forms trigger a markedness implicature (Levinson 2000) relative to the lighter-weight masculine generic: listeners identify a speaker as going to extra effort when using the disjunctive form, and so they infer that the interpretation the speaker means to convey is itself marked. Markedness of meaning gets operationalized relative to stereotypes: a middle-of-the-road interpretation invites listeners to draw on their knowledge of stereotypes when interpreting the referent of a noun phrase; a marked interpretation will deviate from those stereotypes. In the case of a feminine-stereotyped noun, a marked meaning would feature more men; with masculine-stereotyped nouns, marked means featuring more women.

To implement this hypothesis in a way that makes both qualitative and quantitative predictions that can be tested against the behavioral data in our experiment, we develop a computational cognitive model of the reasoning that goes into interpreting the relevant noun phrases. Our model is formulated within the RSA framework (Frank & Goodman 2012; Goodman & Stuhlmüller 2013; Scontras et al. 2018), which models communication as a process of recursive social reasoning between speakers and listeners: a listener interprets a noun phrase by imagining the referent (i.e., a collection of men and women) that led the speaker to produce it, and a speaker chooses which noun phrases to produce by reasoning about how a naive listener would literally interpret it.

We begin with a specification of our model, then we make assumptions about free parameters, and finally we explore the model’s predictions in light of the behavioral data collected in our experiment.

**5.1. MODEL SPECIFICATION.** We treat the state of the world that a speaker describes using a profession noun  $k$  as a group of men and women. We indicate the prevalence  $s_k$  as the real proportion of women in the group. A prevalence of  $s = 0$  indicates all men and  $s = 1$  indicates all women. In this work, we assume that all group members are drawn from the man/woman binary. The percent of men in the group can therefore be inferred as  $1 - s$  where  $s$  represents the observed prevalence of women. The values of  $s$  were from 0.01 to 1.0 in steps of 0.01.

**5.1.1. UTTERANCES & SEMANTICS.** In selecting their utterance, a speaker chooses among four options  $u_k \in U_k$ :

$$U_k = \{\text{MASCPL}; \text{FEMPL}; \text{DISJ}; \text{NULL}\} \quad (1)$$

MASCPL is the generic masculine plural—the so-called default; it is true of a state when the group contains at least some men (i.e.,  $s < 1$ ); FEMPL is the feminine plural utterance, which



is true of a group of exclusively women (i.e.,  $s = 1$ ).<sup>1</sup> DISJ is the disjunctive utterance, which is true if either MASCPL or FEMPL is true of the state. Finally, NULL represents the null utterance, or saying nothing at all and thereby instructing the listener to rely solely on their prior knowledge; NULL is true for all values of  $s$ .

5.1.2. PRIOR BELIEFS. Prior probability of prevalence  $s_k$  for a given profession  $k$  is drawn from a discretized Beta distribution parameterized by mean  $\gamma_k$  and concentration (inverse variance)  $\delta_k$ :

$$P(s_k) \sim \text{Beta}(\gamma_k, \delta_k) \quad (2)$$

The  $\gamma_k$  and  $\delta_k$  parameters are estimated based on the behavioral data via a Bayesian graphical model in JAGS (Plummer et al. 2003) with priors:

$$\begin{aligned} \gamma_k &\sim \text{uniform}(1, 10), \\ \delta_k &\sim \text{uniform}(0, 3). \end{aligned}$$

Estimation takes place by iterating over the  $j$  responses from the behavioral studies and drawing samples from a Gaussian distribution with mean  $\gamma_k$  and precision  $\frac{1}{\delta_k^2}$ :

$$s_{k,j} \sim \text{Gaussian}(\gamma_k, \frac{1}{\delta_k^2}).$$

These values are then transformed to proportions by dividing by the number of points on the Likert scale (i.e., by 9). In effect, these prior distributions are the smoothed results from Experiment 1.

5.1.3. LITERAL LISTENER. The naive literal listener  $L_0$  hears an utterance  $u_k$  and infers the proportion of women  $s_k$  in the group the utterance is meant to describe.  $L_0$  performs this inference by reasoning about the literal semantics and updating flat prior beliefs about the prevalence in a given profession,  $P_{L_0}(s_k)$ . In this way,  $L_0$  returns a uniform distribution over values of  $s$  for which  $\llbracket u_k \rrbracket(s_k) = 1$ .

$$P_{L_0}(s_k|u_k) \propto \llbracket u_k \rrbracket(s_k) \cdot P_{L_0}(s_k). \quad (3)$$

5.1.4. SPEAKER. The speaker  $S_1$  observes the prevalence of women  $s_k$  among some group and selects an utterance  $u_k$  to communicate that prevalence to  $L_0$ . The utterance is chosen according to an optimality parameter  $\alpha$ , the presumed behavior of the naive listener  $P_{L_0}$ , and the utterance cost  $C(u_k)$ . The cost is weighted by a stereotype parameter  $\lambda$ .

$$P_{S_1}(u_k|s_k) \propto \exp(\alpha \cdot [\log(P_{L_0}(s_k|u_k) - \lambda \cdot C(u_k))]) \quad (4)$$

5.1.5. PRAGMATIC LISTENER. The pragmatic listener  $L_1$  infers the prevalence of women  $s_k$  among a group given an utterance  $u_k$  by considering which values for  $s_k$  would have been most likely to lead  $S_1$  to produce  $u_k$  (i.e., the utterance that  $L_1$  encountered).  $L_1$  uses this reasoning to update the prior beliefs about prevalences  $P(s_k)$ :

$$P_{L_1}(s_k|u_k) \propto P_{S_1}(u_k|s_k) \cdot P(s_k) \quad (5)$$

<sup>1</sup> A more complex model where the semantic thresholds  $\theta_m$  and  $\theta_f$  were inferred by a pragmatic listener returned qualitatively comparable results.

5.2. **PARAMETER SETTING.** To generate predictions from our model, we need to fix values for the various free parameters. Utterance costs were set according to their length and/or relative frequency, where  $C(\text{MASCPL}) = 1$ ,  $C(\text{DISJ}) = 6$ ,  $C(\text{FEMPL}) = 4$ , and  $C(\text{NULL}) = 1$ . The optimality parameter was set at  $\alpha = 4$ , corresponding to a cooperative speaker who is especially likely to select utterances with higher utility.

We operationalize  $\lambda$  as  $\lambda = |s_k - 0.5|$ , which weights the utterance cost according to the size of the stereotype effect—relative to a 50/50 split—as measured behaviorally in Experiment 1. Under this weighting, costly utterances used to describe states near the extremes— $s = 0.01$  (all men in our model) or  $s = 1.0$  (all women)—are costlier than costly utterances used to describe states near  $s = 0.5$  (evenly split).

5.3. **PREDICTIONS.** Figure 6 plots model-predicted posterior distributions for all 55 professions, and Figure 7 plots six representative professions in more detail. Among the stereotypically-male professions (e.g., trucker, firefighter), the model estimates a higher prevalence of women under the higher-cost DISJ utterance relative to the MASCPL utterance. For stereotypically female professions (e.g., nurse, librarian), the model predicts a higher prevalence of women under the DISJ utterance than under the MASCPL. Under the high cost utterance, therefore, the model draws posterior mass away from the extrema relative to the lower-cost utterance.

Figure 8 plots model predictions for the model against human behavior from Experiment 2. For both model predictions and human behavior, we calculate a difference score between the proportions of women given for the masculine generic form and the disjunctive form. Positive values of the difference score indicate that more men were inferred for the disjunctive form relative to the masculine generic, while negative values indicate that more women were inferred. The model accounts for 22.4% of the variance in the human behavior ( $r^2 = 0.224$ ).

5.4. **DISCUSSION.** Overall, the model provided qualitatively human-like predictions regarding the proportion of women under the DISJ utterance relative to the MASCPL. In order to produce these behavioral patterns, we introduced two pieces into our RSA model that allowed it to function: the disjunctive semantics of the DISJ utterance, and the stereotype cost-weighting parameter  $\lambda$ . With regard to the semantics, the MASCPL and FEMPL utterances have simple interpretations as either meeting some minimum or maximum proportion of women respectively. By contrast, for the DISJ utterance, the semantics is the disjunction of the semantics of the other two non-null utterances, MASCPL and FEMPL.

With regard to the cost-weighting, one might ask whether this parameter is necessary to produce the observed behavior since the utterances have such different costs. Indeed, the use of this parameter represents a departure from more typical RSA models where the norm is to use a 1:1 mapping between form and cost. We found, however, that this cost-weighting is necessary to produce the observed behavior. When  $\lambda$  was removed, the model returned the same quantitative predictions for both the DISJ and MASCPL forms, rather than the difference observed in the behavioral data. It seems, then, that treating costly utterances (i.e., disjunction) as even costlier when they are used to describe a strongly-stereotyped profession is a necessary ingredient to the calculation markedness implicatures for our utterances.

**6. General discussion.** Contrary to the claims of the RAE, the “default” low-cost masculine generic utterance does not yield unbiased expectations over gender distribution regardless of profession, a pattern we observe in our behavioral data and reproduce in our model. Crucially,

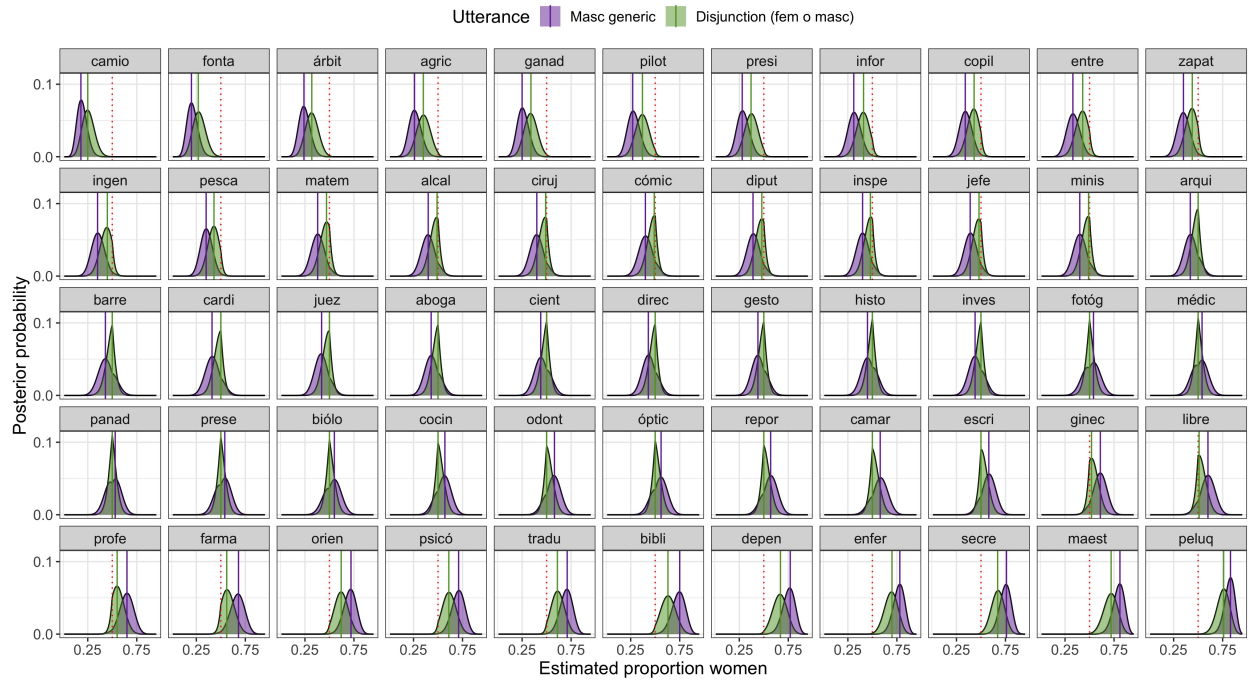


Figure 6. Posterior distributions for all professions. Solid vertical lines represent maximum likelihood estimations for each utterance. Dotted red line included at 0.5 for reference.

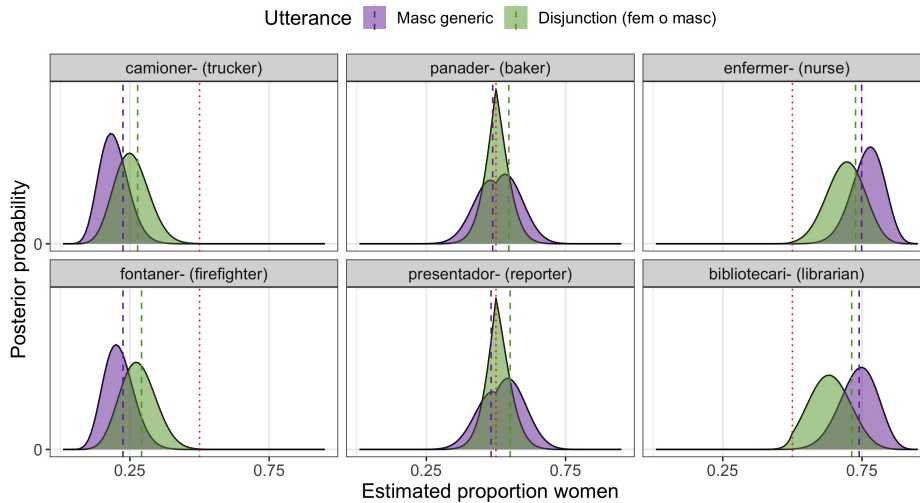


Figure 7. Posterior distributions for representative professions. Dashed vertical lines represent mean response for each utterance from Exp. 2. Dotted red line included at 0.5 for reference.

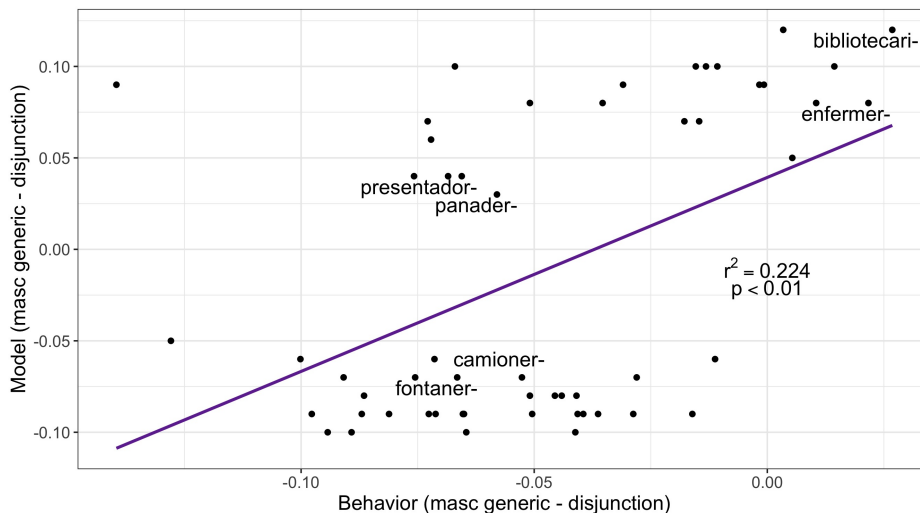


Figure 8. Comparison of behavioral effect and model prediction. We found a significant positive correlation ( $r = 0.47, p < 0.001$ ) between the effect sizes for behavior and model predictions.

the behavioral data show that the higher-cost disjunctive utterance does not necessarily signal more women than the low-cost masculine generic utterance, but can also be used to indicate more men—again, a behavioral pattern we have been able to recreate with our RSA model. We therefore posit that the high-cost utterance is not used to indicate more women but rather to contrast with the listener’s prior expectations about the gender prevalence of the profession.

In short, we have a functioning model to capture an interesting pattern of human linguistic behavior. We propose that the differences in interpretation between the masculine generic and inclusive disjunction utterances occur on the basis of a markedness implicature (Levinson 2000)—the listener infers that if the speaker was describing a group of professionals with the stereotypically-expected proportions of men and women, then the low-cost masculine generic would be used. If this assumption is held, then the listener who hears a disjunctive utterances infers that the group must somehow be different from expectations in order to warrant the use of a more costly marked inclusive utterance.

More work is needed, however, to achieve a more accurate quantitative fit to human behavior. One element of this work would be a behavioral study whose dependent measure maps more clearly onto model predictions, for example asking participants how many out of a set number of people are women given the utterances. Moreover, further research would need to include other alternatives such as the slash form or even the use of newly introduced options like the ending vowel *-e*.

**7. Conclusion.** We find that inclusive alternatives to the masculine generic in Spanish serve an important role when referring to mixed-gender groups, delivering interpretations that are indeed more inclusive. We further find that the interpretation behavior can be modeled as rational inference between listeners and speakers. One implication of this work is that the masculine generic is not so gender-neutral as the RAE would have us believe, but fortunately the Spanish language provides disjunctive forms as a ready, more inclusive alternative. The present study thus further highlights the potential benefits of expanding or normalizing the use of such alternative forms in everyday language use.

## References

- Bengoechea, Mercedes. 2015. Cuerpos hablados, cuerpos negados y el fascinante devenir del género gramatical. *Bulletin of Hispanic studies* 92(1). 1–25.
- Bosak, Janine, Sabine Sczesny & Alice H. Eagly. 2012. The impact of social roles on trait judgments: A critical reexamination. *Personality and Social Psychology Bulletin* 38(4). 429–440. <https://doi.org/10.1177/0146167211427308>.
- Bosque, Ignacio. 2012. Sexismo lingüístico y visibilidad de la mujer. Tech. rep. Madrid: Real Academia Española.
- Braun, Friederike, Susanne Oelkers, Karin Rogalski, Janine Bosak & Sabine Sczesny. 2007. “For reasons of intelligibility...”: How masculine generics and alternative forms affect the cognitive processing of a text. *Psychologische Rundschau* 58(3). 183–189.
- Brown, Rupert. 2011. *Prejudice: Its social psychology*. Hoboken, NJ: John Wiley & Sons.
- Cacciari, Cristina & Roberto Padovani. 2007. Further evidence of gender stereotype priming in language: Semantic facilitation and inhibition in Italian role nouns. *Applied Psycholinguistics* 28(2). 277–293. <https://doi.org/10.1017/S0142716407070142>.
- Carreiras, Manuel, Alan Garnham, Jane Oakhill & Kate Cain. 1996. The use of stereotypical gender information in constructing a mental model: Evidence from English and Spanish. *The Quarterly Journal of Experimental Psychology Section A* 49(3). 639–663. <https://doi.org/10.1080/713755647>.
- Eagly, Alice H. & Valerie J. Steffen. 1986. Gender and aggressive behavior: A meta-analytic review of the social psychological literature. *Psychological Bulletin* 100(3). 309. <https://doi.org/10.1037/0033-2909.100.3.309>.
- Eagly, Alice H., Wendy Wood & Amanda B. Diekmann. 2000. Social role theory of sex differences and similarities: A current appraisal. In Thomas Eckes & Hanns M. Trautner (eds.), *The developmental social psychology of gender*, 123–174. Mahwah, NJ: Lawrence Erlbaum.
- Frank, Michael C. & Noah D/ Goodman. 2012. Predicting pragmatic reasoning in language games. *Science* 336(6084). 998–998. <https://doi.org/10.1126/science.1218633>.
- Gabriel, Ute, Pascal Gygas, Oriane Sarrasin, Alan Garnham & Jane Oakhill. 2008. Au pairs are rarely male: Norms on the gender perception of role names across English, French, and German. *Behavior Research Methods* 40(1). 206–212. <https://doi.org/10.3758/BRM.40.1.206>.
- Gastil, John. 1990. Generic pronouns and sexist language: The oxymoronic character of masculine generics. *Sex Roles* 23. 629–643.
- Goodman, Noah D. & Andreas Stuhlmüller. 2013. Knowledge and implicature: Modeling language understanding as social cognition. *Topics in Cognitive Science* 5(1). 173–184. <https://doi.org/10.1111/tops.12007>.
- Guerra, Ernesto, Jasmin Bernotat, Héctor Carvacho & Gerd Bohner. 2021. Ladies first: Gender stereotypes drive anticipatory eye-movements during incremental sentence interpretation. *Frontiers in Psychology* 12. 1868. <https://doi.org/10.3389/fpsyg.2021.589429>.
- Gygas, Pascal & Ute Gabriel. 2008. Can a group of musicians be composed of women? Generic interpretation of French masculine role names in the absence and presence of feminine forms. *Swiss Journal of Psychology* 67(3). 143–151.
- Hamilton, David L. & Tina K. Troler. 1986. Stereotypes and stereotyping: An overview of the cognitive approach. In John F. Dovidio & Samuel L. Gaertner (eds.), *Prejudice, discrimination, and racism*, 127–163. Cambridge, MA: Academic Press.
- Hamilton, Mykol C. 1988. Using masculine generics: Does generic *he* increase male bias in the user’s imagery? *Sex Roles* 19. 785–799. <https://doi.org/10.1007/BF00288993>.

- Irmen, Lisa & Julia Kurovskaja. 2010. On the semantic content of grammatical gender and its impact on the representation of human referents. *Experimental Psychology* 57(5). 367–375. <https://doi.org/10.1027/1618-3169/a000044>.
- Jussim, Lee, Jacquelynne Eccles & Stephanie Madon. 1996. Social perception, social stereotypes, and teacher expectations: accuracy and the quest for the powerful self-fulfilling prophecy. In Mark P. Zanna (ed.), *Advances in experimental social psychology*, 281–388. Cambridge, MA: Academic Press. [https://doi.org/10.1016/S0065-2601\(08\)60240-3](https://doi.org/10.1016/S0065-2601(08)60240-3).
- Kennison, Shelia M & Jessie L Trofe. 2003. Comprehending pronouns: A role for word-specific gender stereotype information. *Journal of Psycholinguistic Research* 32. 355–378. <https://doi.org/10.1023/A:1023599719948>.
- Levinson, Stephen C. 2000. *Presumptive meanings: The theory of generalized conversational implicature*. Cambridge, MA: MIT Press.
- Lledó, Eulàlia. 2006. *Las profesiones de la A a la Z: En femenino y en masculino*. Madrid: Instituto de la Mujer.
- Martin, Carol Lynn & Diane N. Ruble. 2010. Patterns of gender development. *Annual Review of Psychology* 61. 353–381. <https://doi.org/10.1146/annurev.psych.093008.100511>.
- Menegatti, Michela, Elisabetta Crocetti & Monica Rubini. 2017. Do gender and ethnicity make the difference? Linguistic evaluation bias in primary school. *Journal of Language and Social Psychology* 36(4). 415–437. <https://doi.org/10.1177/0261927X17694980>.
- Misersky, Julia, Pascal M. Gygax, Paolo Canal, Ute Gabriel, et al. 2014. Norms on the gender perception of role nouns in Czech, English, French, German, Italian, Norwegian, and Slovak. *Behavior Research Methods* 46. 841–871. <https://doi.org/10.3758/s13428-013-0409-z>.
- Moon, Christine, Robin Panneton Cooper & William P. Fifer. 1993. Two-day-olds prefer their native language. *Infant Behavior and Development* 16(4). 495–500. <https://doi.org/b9xpwq>.
- Plummer, Martyn. 2003. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proceedings of the International Workshop on Distributed Statistical Computing* 3. 1–10.
- Prewitt-Freilino, Jennifer L., T. Andrew Caswell & Emmi K. Laakso. 2012. The gendering of language: A comparison of gender equality in countries with gendered, natural gender, and genderless languages. *Sex Roles* 66(3). 268–281. <https://doi.org/10.1007/s11199-011-0083-5>.
- Pyykkönen, Pirita, Jukka Hyönä & Roger P. G. Van Gompel. 2009. Activating gender stereotypes during online spoken language processing. *Experimental Psychology* 57(2). 126–133. <https://doi.org/10.1027/1618-3169/a000016>.
- Reali, Chiara, Yulia Esaulova, Anton Öttl & Lisa Von Stockhausen. 2015. Role descriptions induce gender mismatch effects in eye movements during reading. *Frontiers in Psychology* 6. 1607. <https://doi.org/10.3389/fpsyg.2015.01607>.
- Schneider, David J. 2005. *The psychology of stereotyping*. New York: Guilford Press.
- Scontras, Gregory, M. H. Tessler & Michael Franke. 2018. Probabilistic language understanding: An introduction to the rational speech act framework . Website. <https://www.problang.org>.
- Sczesny, Sabine, Magda Formanowicz & Franziska Moser. 2016. Can gender-fair language reduce gender stereotyping and discrimination? *Frontiers in Psychology* 7. 25. <https://doi.org/10.3389/fpsyg.2016.00025>.
- Signorielli, Nancy. 2001. Television's contribution to stereotyping: Past, present, future. In (eds. n/a), *Handbook of children and the media*, 341–358. Publisher n/a.

- Stahlberg, Dagmar, Friederike Braun, Lisa Irmen & Sabine Sczesny. 2007. Representation of the sexes in language. In Klaus Fiedler (ed.), *Social communication: Frontiers of social psychology*, 163–187. Hove, UK: Psychology Press.
- Stahlberg, Dagmar, Sabine Sczesny & Friederike Braun. 2001. Name your favorite musician: Effects of masculine generics and of their alternatives in German. *Journal of Language and Social Psychology* 20(4). 464–469. <https://doi.org/10.1177/0261927X01020004004>.
- Su, Jui-Ju, Nicola Molinaro, Margaret Gillon-Dowens, Pei-Shu Tsai, Denise H. Wu & Manuel Carreiras. 2016. When “he” can also be “she”: An ERP study of reflexive pronoun resolution in written Mandarin Chinese. *Frontiers in Psychology* 7. 151. <https://doi.org/j5cf>.
- UNESCO. 2011. Priority gender equality guidelines. Tech. rep. United Nations Educational, Scientific and Cultural Organization.
- Wood, Wendy & Alice H. Eagly. 2002. A cross-cultural analysis of the behavior of women and men: Implications for the origins of sex differences. *Psychological Bulletin* 128(5). 699. <https://doi.org/10.1037/0033-2909.128.5.699>.
- Yzerbyt, Vincent & Stéphanie Demoulin. 2010. Intergroup relations. In Daniel T. Gilbert, Susan T. Fiske & Gardner Lindzey (eds.), *Handbook of social psychology*. Hoboken, NJ: John Wiley & Sons. <https://doi.org/10.1002/9780470561119.socpsy002028>.