

Do People Remember What is Prototypical?

The Role of Accent-Religion Intersectionality for Individual and Category Memory

Tamara Rakić¹, Melanie C. Steffens², and Atena Sazegar³

¹Psychology Department, Lancaster University, Lancaster, UK

²Faculty of Psychology, Department of Social, Environmental, and Economic Psychology,
University of Koblenz-Landau, Landau, Germany

³University Medical Center Hamburg-Eppendorf, Department of Medical Psychology,
Research Group on Psychosocial Migration Research, Germany

Author Note

Tamara Rakić, ORCID: <https://orcid.org/0000-0002-8591-923X>

Melanie Steffens, ORCID: <https://orcid.org/0000-0002-7915-3629>

The writing of this paper was supported by German Research Foundation grants (awarded to Melanie C. Steffens and Tamara Rakić, DFG, Ste 938/10-2, and to Tamara Rakić, DFG, RA 2203/1).

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. The authors would like to thank the Prof. Pete Neal for advice on homogeneity analyses and Daniel Heck for help with the set-up of the TreeBUGS package. Data used in the current article are available under license: CC BY-NC at doi:10.17635/lancaster/researchdata/103.

Correspondence should be addressed to Tamara Rakić, Psychology Department, Lancaster University, LA1 4YF Lancaster, UK. Email: t.rakic@lancaster.ac.uk

Abstract

Evidence suggests that accents can be typically more powerful in activating ethnicity categorization than appearance. Concurrently, some social categories, such as ethnicity, can be linked with other categories, such as religion. We investigate how people categorize those who belong to a (mis)matching pair of categories? In the present study we investigated Germans' categorization of women either wearing a headscarf (Muslim religious symbol), or not, and speaking either standard German or German with an Arabic accent. The "Who Said What?" paradigm and multinomial modelling yielded that category memory, indicative of subtyping, was best for non-prototypical targets (i.e., headscarf and standard German accent, no headscarf and Arabic accent). In contrast, ingroup targets (no headscarf and standard German accent) were individually remembered better than all other targets, whereas non-prototypical targets (no-headscarf and Arabic accent) were not remembered individually at all. These findings are discussed in terms of intersectionality and category prototypicality.

Keywords: classification, ethnic identity, religion, intersectionality, prototypicality, accent, WSW-paradigm

Do People Remember What is Prototypical?

The Role of Accent-Religion Intersectionality for Individual and Category Memory

People simplify the vast amount of information around them through categorization. One of the main interests of social psychology is to understand how people process multiple categories others belong to in order to form an impression of them (e.g., Fiske & Neuberg, 1990). Which categories will be activated depends on the context (or fit) as well as previous experience; however, chronically accessible categories (i.e., gender, age, ethnicity) are likely to be activated across different situations (Bruner, 1957; Oakes, 1987). Whereas these categories can be identified through facial cues (Hugenberg & Wilson, 2013), recent empirical evidence suggests that language (accent) often plays a more important role for ethnic categorization. In fact, ethnicity may be much better represented with accent than with typical appearance (Rakić et al., 2011). This demonstrates the complexity of social categories and how they are activated.

One of the major concerns related to social categories is their strong association to simplified knowledge structures, or stereotypes (Fiske, 1998). Indeed, stereotypes are powerful in determining attitudes and expectations about social groups. A particularly puzzling aspect of stereotypes is their maintenance even in the presence of disconfirming information, in other words: counter-stereotypical exemplars. In some cases, stereotypes about one social category, such as ethnicity or nationality, can be strongly associated to other social categories, for example religion. At the intersection of two social categories, we refer to individuals who belong to two associated categories as prototypical (e.g., person wearing a headscarf and speaking with a nonnative accent). The aim of this article is to investigate the outcome of accent/ethnicity-religion cross-categorization with a focus on the prototypicality of targets. We test whether targets are categorized based on their accent (as proxy for

ethnicity) and religion (here: headscarf present vs. absent) separately or concurrently and whether they are individually remembered.

Social Categorization and Stereotyping

People tend to use categorical representations of the (social) world around them. This means that the huge amount of information present in individual differences is converted into simplified categories. Both the process and consequences of social categorization have received much attention in social psychology (e.g., Macrae & Bodenhausen, 2000). There seems to be an almost infinite number of possible social categories that can in principle be perceived concurrently. Indeed people belong simultaneously to multiple social categories, and depending on their relative salience, they may be perceived differently in different contexts (Crisp & Hewstone, 2007).

Which category or categories will be activated depends on their salience and fit (Bruner, 1957; Oakes, 1987). The categories gender, age, and ethnicity are often referred to as the Big Three social categories. They are characterized by high salience making them likely to be perceived and used in categorization and, consequently, in stereotype activation (e.g., Stangor et al., 1992). Other “less salient” categories may pass unnoticed. However, all categories can be more or less salient because of the intersectionality of categories and different contexts (e.g., pertaining to research, by using visual or auditory stimuli).

Research Methods in Social Categorization

Pertaining to research methods, the cues that are used to indicate a given category can impact how and when people use those categories. For example, every category can be presented through labels (e.g., “man”, “black”), and every literate person will at once understand the meaning of these words and categories; consequently, respective stereotypes could be immediately activated (Zarate & Smith, 1990). However, people encountered in everyday life do not wear labels to indicate social categories. People extricate social-category

membership from what they see and/or hear (i.e., from the available information). Recently, there has been a greater awareness in the choice of methodology when studying categorization and stereotyping. It is well known that language contains powerful social cues (Giles & Rakić, 2014); indeed (spoken) language is a key aspect of ethnicity (as indicated by ethnolinguistic identity theory; Giles et al., 1977; Giles & Johnson, 1981, 1987). Still, for a long time language only held a peripheral position in studies of social phenomena (including the “Who Said What?” paradigm we describe below, ironically, in which often only a written statement indicates what someone said). Recent empirical evidence demonstrated the dominant role of language (or accents) in ethnic categorization: On its own, a typical appearance (i.e., facial visual information) was an equally strong cue for ethnic categorization as an accent (i.e., auditory information). However, when these two modalities were combined, people relied significantly more on accent than on looks in ethnic categorization (Rakić et al., 2011).

Category Intersections

Not only the cues researchers use affect categorization, but so do category intersections. Different social categories, such as ethnicity or nationality on the one hand and religion on the other, can have a rather complex relationship. For example, being Polish is also highly associated with being Catholic (i.e., prototypical combination). Still, minorities exist such as Polish Tatars, for whom their Muslim religion is important in their Polish identity (Verkuyten, 2007). Similar concepts can be observed also in other European countries such as the Netherlands or Germany, both comprising a significant proportion of national minorities of Muslim religion (e.g., of Arabic or Turkish origin). Correlations between categories do not only affect individuals’ identities but can influence how other people perceive and categorize them (e.g., as members of their ingroup or as members of an outgroup).

Investigating and understanding this intersectionality of social categories is important if we are to truly understand why some groups are more or less likely to be discriminated against (e.g., Goff et al., 2008; Schug et al., 2015). For example, in terms of race and gender, Black women seem to be invisible compared to the prototypical combinations of Black men or White women (Sesko & Biernat, 2010). The invisibility in this case refers to the lack of individuation or lack of differentiation between group members, who go unnoticed. Therefore, looking at these two categories separately would not provide an adequate understanding due to the incomplete information, because they are perceived as welded together and therefore inseparable. Researchers have only begun to understand the invisibility of category intersections. For example, it has recently been shown that both similarity to the category prototype and a focus on individual differences increase individual visibility (Sesko & Biernat, 2018). A prototype can be understood as “the ideal-type member of a category that best represents its identity in a given context” (Wenzel et al., 2007, p. 335). This has consequences when it comes to understanding immigrant health (e.g., Viruell-Fuentes et al., 2012) or how, for example, German Muslim feminists are perceived in feminist communities (Weber, 2015).

Category intersections could also affect the susceptibility of stereotypes to change. Indeed, a lot of research has investigated how people react to stereotype-(in)consistent information or stereotype-inconsistent members of a given social group (e.g., Dolderer et al., 2009). Two main patterns are possible: *subtyping* the stereotype-inconsistent individuals as the exceptions to the generally correct rule (or stereotype), which implies stereotype maintenance; or *subgrouping* by creating a more differentiated picture of a given social group while accounting for different and potentially stereotype-inconsistent members. The latter process leads to stereotype change (Richards & Hewstone, 2001). Paralleling stereotype-(in)consistency, understanding how exemplars belonging to prototypical and non-prototypical

category intersections are mentally processed could also be important for understanding people's tendency for stereotype change and maintenance and consequently in preventing discrimination.

The Present Research

In the present experiment, we investigated how women in Germany are categorized based on their ethnicity/accent (presented with either standard German vs. Arabic accent¹) and Muslim religion (wearing a headscarf or not). We chose the modified “Who Said What?” paradigm (Taylor et al., 1978), as it allows for testing spontaneous categorization of people in a group setting while controlling for guessing (Klauer & Wegener, 1998). We did not use group labels because we aim to investigate information processing more closely to real-life – where more often than not people need to extract categories from available pieces of information. Whereas objectively the two categories (i.e., ethnicity/accent and religion) are independent, there can be strong intersectionality between them where one seems to imply the other. Alternatively, one could argue that both potentially are sufficient as cues for foreignness (i.e., Arabic accent, headscarf, or both) by non-Muslim German participants.

Given the association of both an Arabic accent and a headscarf with Muslim religion, and given the group setting implemented in our investigation, in which headscarves and foreign accents were presented often, we hypothesized that our targets would be categorized based on prototypicality in terms of category intersectionality; namely, women wearing a headscarf and speaking with an Arabic accent (or no headscarf and speaking with a standard German accent) should be perceived as prototypical, whereas those wearing a headscarf but speaking with a standard German accent (or no headscarf but speaking with an Arabic accent) should be perceived as non-prototypical. Our main dependent variables were memory for each individual and memory for the (sub-)category the individual belongs to. We

¹ We use the term Arabic accent throughout this article; this term was used in the pre-test (see SOM, Appendix A) and it was the most comprehensible for our participants.

expected that participants will remember individual targets better if they are prototypical (e.g., Sesko & Biernat, 2010); in contrast, we expected an effect similar to outgroup homogeneity/invisibility where non-prototypical targets will not be remembered individually.

In terms of subgroup memory (containing both categories), we expected category memory to reflect prototypicality, with all (non)prototypical targets being mixed up with each other, indicating subtyping.

Method

Ethics Statement

Ethical approval for this study was obtained from the Ethics Commission, Friedrich-Schiller University of Jena, Faculty of Social and Behavioral Sciences (number FSV 12/02).

“Who Said What?” Paradigm

We used the modified “Who Said What?” paradigm (WSW-paradigm; Taylor et al., 1978), as it allows testing spontaneous category activation. The WSW-paradigm is particularly useful as it is sensitive to implicit information organisation (i.e., categorization) that can hardly be changed during later explicit judgements (Carlston & Schneid, 2015). Participants were instructed to observe parts taken from a discussion among eight different students. Their only task was to form an impression of the group as a whole. During the discussion each target was presented with a picture and audio extract uttering a given statement. At the end of the “discussion” part participants were asked to match written statements to the speakers, with the photographs of all speakers being shown.

The original WSW-paradigm is based on the principle of accentuation (e.g., Corneille et al., 2004). Once a given category is activated people automatically tend to decrease the perceived differences within the category and increase those between categories. In terms of the WSW-paradigm this is revealed in the type of errors participants make during the matching task. For example, in case of accent-based categorization, if a statement made by a

speaker with a standard German accent is falsely assigned to another standard-German-accent speaker rather than to an Arabic-accent speaker, this would be considered a within category error. Conversely, confusing a statement made by a speaker with standard German accent as being made by an Arabic-accent speaker would constitute a between category error. The more within compared to between category confusions participants make, the stronger the evidence for categorization on the selected dimension. The WSW-paradigm has been successfully used with single as well as two crossed categories (e.g., Klauer et al., 2003; Klauer & Wegener, 1998). The same principle as for accent-based categorization applies for cross-categorization based on two categories (e.g., ethnicity/accent and religion). In the case of cross-categorization there are within and/or between category errors based on each of two categories, resulting in a more complex assignment.

However, with this (original) procedure it is impossible to differentiate between participants' actual memory and lucky guessing. In order to account for guessing, the adapted procedure developed by Klauer and Wegener (1998) was used: In the matching task, 48 new statements that had not been presented during the discussion phase were added next to 48 old statements. The new statements could have easily been presented in the discussion based on their content. The participants' task for each statement was first to say whether a statement was old (i.e., previously presented) or new (i.e., not previously presented). If a statement was categorized as new, then another statement appeared on the screen, whereas if the statement was recognized as old, participants were asked to indicate which target made it. Klauer and Wegener (1998) demonstrated two points: whereas participants are typically very good at identifying the new statements, they do not remember a portion of the old statements and categorized them as new. This extension made it possible to analyse the data with multinomial modelling.

The multinomial modelling procedure enables a distinction between actual memory and guessing, both of which can result in the same outcome (e.g., correct assignment or within category confusion); this allows for a more informed interpretation of processes that contribute to the observed outcome. Such refinement is necessary because the correct match between a statement and a speaker can be arrived at by very different cognitive paths (see Figure 1 in Rakić et al., 2011, p. 6). These vary from actually remembering the statement and the speaker, through a more indirect route: guessing that the statement was old, then guessing both categories of the speaker and, finally, guessing the correct speaker. Furthermore, compared to the original WSW-paradigm the use of multinomial modelling can control for different confounds. For example, increased similarity of targets would result in a higher number of raw within-category errors, thus falsely indicate categorization. On the contrary, with multinomial modelling this affects memory parameters of individual targets but not category memory. Similarly, the original WSW-paradigm has issues with baseline probabilities, which need to be taken into account to avoid misrepresenting the actual frequencies for different types of errors; in contrast, this does not affect the parameter estimates of the multinomial model (Klauer & Wegener, 1998).

Materials

In the WSW-paradigm different statements are used during the discussion part. All statements we used were neutral and covered different aspects of university; for example, “There are too few textbook exemplars in the library” (see Klauer & Wegener, 1998, for the original set of statements). This was done to ensure that the content of statements could not be used for categorizing. Each target made statements on six different topics but the topics were held constant between targets (e.g., everyone made a remark about the library).

In terms of selection of voices and faces, a detailed description of targets, as well as respective pre-tests are outlined in the Supplementary Online Material (SOM), Appendix A.

Briefly, for voices (i.e., accents) we used the matched guise technique (MGT; Lambert et al., 1960), where all speakers were recorded speaking with Arabic and with Standard German accent. This allows to control for the influence of other vocal qualities. Independently of the speakers' accent they were always clearly intelligible. Similarly, the same "neutral" looking women (i.e., not perceived as typically German or Arabic) were photographed either with or without a headscarf. The headscarves were in neutral colours and fitted by an experimenter familiar with the use of headscarves. Women's faces were always fully and clearly visible.

Procedure

We introduced the experiment as a study on information processing. The procedure followed the modified WSW-paradigm² as described above: the discussion part was identical to the original paradigm (see Taylor et al., 1978, for the original instructions) and the matching task was adapted from Klauer and Wegener (1998). Each time a participant heard a given target speaking, the respective photograph was shown on the screen (see Rakić et al., 2011, for a similar procedure). A total of eight targets was presented in random order uttering a total of six statements each. All of the targets were women, differing on the bases of their accent (half standard German, half Arabic accent) and displayed (Muslim) religion (half wearing a headscarf, half no headscarf).

The experiment consisted of two counterbalanced conditions created to control for the specific face (with or without headscarf) and accent match. Consequently, if women in Condition A were presented with a headscarf speaking with a standard German accent, in Condition B these women were presented with a headscarf and speaking with an Arabic accent. Participants were randomly assigned to one of the two conditions. Because results

² A drawback of the WSW-paradigm and this procedure was that visual information had an advantage over auditory information because it was present both in the discussion and matching task, whereas accents were only present during the initial discussion part. Nevertheless, this is only a minor issue; any evidence of accent-memory means that the effect was in some degree stronger as it had to be sustained from the discussion phase without further hints during the matching task (see Rakić et al., 2011, for the finding that accent still trumped looks using this procedure).

indicated no differences between conditions we report analyses collapsed across this control factor.

The computer-based experiment was completed individually, in one session of about 20–25 min. At the end, participants were fully debriefed and rewarded for their participation with either a chocolate bar or partial course credit.

Participants

Given a within-subject design with response frequencies aggregated over participants as the units of analysis, statistical power was high. However, because it cannot be determined in advance how many observations will be obtained in each cell (because participants do not make identical responses to each of the multiple response categories) precise power calculations are impossible (see Klauer & Wegener, 1998, for the recommendation of $n > 5$ for each cell). Therefore, we aimed at the recommended number of at least 20 participants per condition (see Simmons et al., 2011).

During the pre-determined data collection interval, 37 participants (19 men, 18 women; $M_{\text{age}} = 23.24$, $SD = 6.16$) from a German university took part. All except two participants were native German speakers studying in a relatively small and homogeneous town (i.e., without many ethnic minorities). The study was run at a time when wearing headscarves in public was only sometimes discussed in the German media. Data were not analysed until the entire sample had been obtained.

Results

Strategy of Data Analysis

The units of analyses were the overall frequencies of the different types of responses. In total there were 29 different response categories (shown in Table A1). Namely every statement coming from one of five sources (four target types and new statements) could either be correctly assigned to the target who made the statement or incorrectly assigned to

any of the five sources. Data were analysed by means of a multinomial model with the MultiTree software (Moshagen, 2010). The parameters of the multinomial model are mathematical representations of different cognitive processes contributing to the different response types, validated by previous research. When it comes to investigating categorization, the parameters of most interest in the WSW multinomial model are *category memory* (i.e., how well is the category of religion, accent, and the combination of both remembered for different types of targets?) and *person memory* (i.e., how well is each woman in a given category combination remembered?). Comparably, multinomial modelling also allows accounting for category guessing (i.e., guessing the category headscarf, or the category Arabic accent, or the combination of the two).

Briefly, the strategy of analysis can be described as follows: First, a baseline model is identified that fits the data. Typically, this requires imposing equality restrictions on some parameters that are of little theoretical interest so that the model has enough degrees of freedom to be identified. If this is not sufficient to obtain the necessary degrees of freedom, one can proceed by restricting other parameters as well to obtain a baseline model that fits the data. Then we can test hypotheses by introducing additional restrictions and testing the resulting model against the baseline model. For example, if setting equal the category memory parameters for different types of targets results in a misfit of the restricted model as compared the baseline model, the hypothesis is rejected that these category memory parameters are equal.

Baseline Model Identification

The model used for analyses consisted of 27 parameters representing different cognitive processes. Because this specific multinomial model has been validated for the WSW-paradigm (for further explanation and validation, see Klauer et al., 2003; Klauer &

Wegener, 1998; Schug et al., 2015) we will only introduce briefly the parameters crucial for the current analyses.

The distractor (DN) and old statement memory ($D_{nG}, D_{nA}, D_{hG}, D_{hA}$)³ parameters represent knowledge that a given statement is new or that it was presented previously, respectively. Further parameters are: A parameter for guessing the statement status as old (parameter b); *person memory* for a target person in a given category (parameters $c_{nG}, c_{nA}, c_{hG}, c_{hA}$); guessing a target within the correct subgroup (parameter $1/n$); category guessing headscarf (parameters x and y ; where $1-x$ and $1-y$ represent category guessing “no headscarf”) and category guessing Arabic accent (parameters f and p). Importantly, the *category memory* parameters are: category memory for the religious-cue (i.e., indicative of exclusive religion memory; parameters $d_{nG}, d_{nA}, d_{hG}, d_{hA}$); for methodological reasons, category memory for accent was subsequently split into: category memory for accent given no category memory for the religious-cue (i.e., indicative of exclusive accent memory; parameters $h_{nG}, h_{nA}, h_{hG}, h_{hA}$), and category memory for accent given category memory for the religious-cue (i.e., indicative of subgroup memory; parameters $e_{nG}, e_{nA}, e_{hG}, e_{hA}$).

Prior to hypothesis testing, we identified the baseline model by restricting some parameters. These restrictions followed theoretical assumptions (e.g., Klauer et al., 2003; Schug et al., 2015). Thus, we fixed the parameter for guessing the correct speaker in the correct subgroup (parameter $1/n$) at chance level, .5; further, because there was no reason for differentiation in guessing categories, guessing the category headscarf (parameters y and x) and guessing the category Arabic accent (parameters f and p) were set equal, respectively. After trying to set all memory parameters for old statements ($D_{nG}, D_{nA}, D_{hG}, D_{hA}$) equal to distractor memory (DN), a model resulted with unsatisfactory fit, indicating that there was a

³ The abbreviations in the subscripts stand for: D_{nG} = women without a headscarf speaking standard German, D_{nA} = women without a headscarf speaking German with an Arabic accent, D_{hG} = women with a headscarf speaking standard German, D_{hA} = women with a headscarf speaking German with an Arabic accent.

difference between the parameters. Participants showed better item memory for new statements as well as for those made by prototypical targets ($DN, D_{hA}, D_{nG} = .68$) than for statements made by non-prototypical targets ($D_{nA}, D_{hG} = .59$). Imposing these restrictions resulted in a baseline model with acceptable but low fit, $\chi^2 (3, N = 3552) = 7.06, p = .07$. In order to increase model fit and in line with our hypotheses hinging on prototypicality, we set equal category memory for accent given religious-cue-memory for the prototypical targets (parameters e_{nG}, e_{hA}) and for the non-prototypical targets (parameters e_{nA}, e_{hG}), respectively. This resulted in a better baseline model fit, $\chi^2 (5, N = 3552) = 7.06, p = .22^4$. Table 1 gives an overview of parameters as well as the parameter estimates within the baseline model.

Before testing our hypotheses, we checked if there was any difference in bias for guessing headscarf (Muslim religion-cue) or Arabic accent (parameters x, y and f, p , respectively) by setting all four parameters equal. The model did not lose fit, $\Delta\chi^2 (1, N = 3552) = .13, p = .71$; the same was true when we restricted all of them to the value of .5, $\Delta\chi^2 (2, N = 3552) = 1.35, p = .51$, indicating that guessing in both cases was at the expected chance rate of 50%.

Finally, in order to determine the homogeneity of responses in our sample homogeneity analyses were conducted, details of which can be found in the SOM, Appendix B. These analyses show that result patterns do not differ between participants (for the individual parameter estimates and complete dataset, see Rakić, 2020). In other words, the same baseline model can be applied to all of them, and there are no participant clusters with different types of response patterns.

⁴ The overall pattern of results remains the same when using the baseline model with acceptable but low fit.

Table 1*Parameter Estimates and 95% Confidence Intervals (CIs) for the Baseline Model*

Process	Parameter	Estimate	CI
Distractor memory	D _N	.68 ^a	.65-.71
	D _{nG}	.68 ^a	.65-.71
	D _{nA}	.59 ^b	.56-.64
Item memory	D _{hG}	.59 ^b	.56-.64
	D _{hA}	.68 ^a	.65-.71
Item status guessing as old	b	.21	.17-.24
	c _{nG}	.25	.18-.32
	c _{nA}	.007	-.09-.09
Person memory	c _{hG}	.09	.011-.18
	c _{hA}	.16	.08-.24
Exclusive accent/ethnicity memory (no religious- cue memory)	h _{nG}	.18	-.08-.44
	h _{nA}	.22	-.05-.48
	h _{hG}	.04	-.24-.32
Exclusive religious-cue memory	h _{hA}	.44	.18-.69
	d _{nG}	.12	-.05-.29

Process	Parameter	Estimate	CI
	d_{nA}	.27	.14-.40
	d_{hG}	.14	-.002-.28
	d_{hA}	.32	.15-.48
	e_{nG}	.018 ^c	-.47-.51
Accent/ethnicity	e_{nA}	1.00 ^d	.57-1.43
memory given	e_{hG}	1.00 ^d	.57-1.43
religious-cue memory	e_{hA}	.018 ^c	-.47-.51
Category guessing	f	.51 ^e	.42-.59
Arabic	p	.51 ^e	.42-.59
Category guessing	x	.53 ^f	.48-.58
headscarf	y	.53 ^f	.48-.58
Person guessing	1/n	.50	constant

Note. nG = Women without headscarf who spoke standard German; nA = Women without headscarf who spoke with an Arabic accent; hG = Women with headscarf who spoke standard German; hA = Women with headscarf who spoke with an Arabic accent.

^a Parameters set equal.

^b Parameters set equal.

^c Parameters set equal.

^d Parameters set equal.

^e Parameters set equal.

^f Parameters set equal.

Hypothesis Testing

Hypothesis testing within the multinomial model is obtained by constraining parameters of interest and comparing the new model to the baseline model (e.g., Klauer et al., 2003; Rakić et al., 2011); if the new model does not lose fit, the constrained parameters can

be set equal, alternatively the loss of fit indicates that given parameters are of different size, corroborating hypotheses. Parameter estimates are probabilities (i.e., range: 0-1).

Person Memory. We first tested for differences in *person memory* (i.e., memory for individual targets). Following the hypothesis and the baseline model indication for a possible difference between prototypical and non-prototypical targets we first checked the person memory parameters in terms of prototypicality. Namely, person memory for non-prototypical targets was set equal ($c_{hG}, c_{nA} = .06$), $\Delta\chi^2(1, N = 3552) = 2.35, p = .13$, and the model maintained a good fit, indicating that these parameters are not significantly different. Similarly, the model maintained a good fit when the person memory parameters for prototypical targets ($c_{hA}, c_{nG} = .21$) were set equal, $\Delta\chi^2(1, N = 3552) = 2.35, p = .13$. These two were significantly different, $\Delta\chi^2(1, N = 3552) = 14.48, p < .001$, indicating that participants remembered prototypical targets individually better than non-prototypical targets.

For further analyses following previous work on ethnic categorization showing that participants showed an outgroup homogeneity effect for nonnative accent rather than foreign appearance (Rakić et al., 2011), we set parameters equal within the same accent. However, setting person memory equal within Arabic and within German accents both resulted in a loss of fit, $\Delta\chi^2(1, N = 3552) = 6.67, p = .009$ and $\Delta\chi^2(1, N = 3552) = 8.03, p = .005$, respectively. The same happened when parameters were set equal for all targets without headscarf, $\Delta\chi^2(1, N = 3552) = 18.69, p < .001$. Only for targets with a headscarf, $\Delta\chi^2(1, N = 3552) = 1.18, p = .28$, the model retained good fit. Taking together these analyses, participants showed best individual memory for targets without headscarf speaking standard German (.25), followed by targets wearing a headscarf, irrespective of their accent (.13), and effectively no individual memory for targets without headscarf speaking with an Arabic accent (.007). When person memory for targets without headscarf speaking with Arabic accent was set equal to zero, the model did not lose fit, $\Delta\chi^2(1, N = 3552) = 0.03, p = .86$, indicating practically no person

memory for those targets. This indicates that high person memory for prototypical targets is mostly driven by ingroup targets (i.e., without headscarf speaking standard German), and person memory for non-prototypical targets is restricted to targets wearing headscarf speaking standard German.

Category Memory. To first test for exclusive *category memory* by cue type, we set the four individual target parameters equal for category memory for *religious-cue* (parameters d) and for category memory *accent given no religious-cue* (parameters h), respectively, and the model maintained acceptable fit, $\Delta\chi^2(6, N = 3552) = 11.16, p = .08$. The same was true once these two were additionally set equal ($d, h = .19$), $\Delta\chi^2(1, N = 3552) = .03, p = .86$, indicating that participants remembered equally well the category based on the religious cue and accent given no religious-cue-memory.

Furthermore, these did not differ from category memory for accent given religious-cue-memory for prototypical targets (i.e., women without headscarf speaking standard German and women with a headscarf speaking with Arabic accent; parameters $e_{nG}, e_{hA} = .18$), $\Delta\chi^2(1, N = 3552) = .04, p = .84^5$. However, category memory for accent given religious-cue-memory for non-prototypical targets (i.e., women without headscarf speaking with Arabic accent and women with a headscarf speaking standard German, parameters $e_{nA}, e_{hG} = 1.00$), was significantly higher than the other category memory parameters, $\Delta\chi^2(1, N = 3552) = 17.75, p < .001$. In other words, while non-prototypical targets were individually not remembered well, participants did remember very well their non-prototypical combination of categories (i.e., they subtyped them).

⁵ These parameters though small (.18) significantly differ from zero, $\Delta\chi^2(1, N = 3552) = 34.95, p < .001$.

Discussion

Our data show a very interesting pattern: On the one hand participants demonstrate an implicit bias in remembering individual targets who are prototypical of intersecting categories (headscarf and Arabic accent, or no headscarf and standard German accent) better than non-prototypical targets (headscarf and standard German accent, or no headscarf and Arabic accent). This was also reflected in their memory for targets' statements (as part of the baseline model); statements from prototypical targets (together with the new statements) were better remembered compared to statements from non-prototypical targets. However, it is worth noting that better memory of prototypical individuals seems to be a function of excellent memory for ingroup targets (women without headscarf speaking standard German) and practically inexistent memory for individual targets without headscarf speaking with Arabic accent. Targets wearing a headscarf were in-between, independently of their accent. This pattern possibly indicates that low prototypicality and high expectancy violations disable individuation (individuated memory) of women without headscarf but speaking with Arabic accent, for whom person memory was lowest (see Berthold et al., 2019; Hansen et al., 2017; Macrae et al., 1993; Sesko & Biernat, 2010).

Interestingly, there was no difference between overall religious-cue memory and memory for ethnicity/accent (given no religious-cue-memory), indicating that ethnicity and religion are perceived as comparably informative, possibly because the categories are perceived to be linked. On the contrary, for accent-memory given religious-cue-memory (i.e., memory for both categories simultaneously) the hypothesized difference emerged: the combined categories of non-prototypical women were well remembered, indicating subtyping, compared to category memory for prototypical women. Taken together, both individual memory and categorization were guided by the intersectionality of the two

categories in question: ethnicity/accent and religion, and not by an individual dominance of either category.

These findings are relevant for several reasons. First, they further support the relevance of accents in categorization; relatedly religion is well represented with a headscarf and in both cases the category cues are highly salient and ecologically valid. Rakić and colleagues (2011) had found that facial appearance ‘typical’ of ethnicity (i.e., German vs. Italian) did not influence categorization when in the presence of accents. In contrast to typical facial appearance, the present research suggests that headscarves symbolizing religion determine categorization as much as accents do. Whereas previous research indicated that people are blinded by accents, here the intersectionality of the two categories seems to contribute to memory for individuals (i.e., prototypical targets were better remembered individually than non-prototypical ones). Indeed, these findings fit well with previous research on intersectionality (e.g., Goff et al., 2008; Schug et al., 2015; Sesko & Biernat, 2010, 2018).

Furthermore, in this experiment we did not explicitly name the categories used or mention any category-related stereotypes. This further reflects real life where people “extract” or account for available information. Additionally, on the empirical and theoretical front this study serves as a sort of bridge between the traditional impression formation and stereotyping literatures. The former usually tests how multiple cues are interpreted and combined and the latter investigates how different cues trigger individual social categorization (Carlston & Schneid, 2015). Our study used multiple cues in multiple categories. We were able to test how participants processed this information and spontaneously categorized different targets.

Our data support the idea that there is perceived intersectionality between ethnicity/accent and religion, and at least for our German participants this caused the targets

to be perceived based on both pieces of information. The negligible memory of individual non-prototypical targets that we found could contribute to stereotype maintenance by subtyping. Interestingly the outgroup homogeneity effect was not “reserved” for outgroup members on both dimensions but was present also for those with a non-prototypical match between religion and ethnicity/accent. At the same time, category memory for the latter was excellent. Because categorization is often considered the first step in the process of stereotyping and discrimination (Fiske, 1998), being perceived as an *in-* or an *outgroup* member can lead to possible discrimination. One could speculate that these findings indicate that sometimes it is better to be an outgroup member on both dimensions than just on one. However, our experiment did not assess evaluation but only spontaneous categorization. Likewise, it would be interesting if future research related these categorization findings to stereotyping: Among these different groups of targets, who will be stereotyped as a typical “warm, but incompetent” outgroup (see Asbrock, 2010)? Recent findings suggest that in some cases stereotype-inconsistency (e.g., Hansen et al., 2017; Niedlich et al., 2016) and non-prototypicality (Biernat & Sesko, 2013) could suffice to avoid negative impressions, but as that research used other social categories, the findings are as of yet inconclusive.

A first limitation is that the present study is a single study with a relatively small sample. Our participants were all autochthonous Germans from a relatively uniform small town. Though we did not explicitly ask for participants’ religion, none of them displayed any obvious religious affiliation (e.g., by wearing a headscarf). The overall homogeneity of our sample was also confirmed with our hierarchical homogeneity analyses. In other words, it is possible that a relatively low social identity complexity (SIC; Roccas & Brewer, 2016) of our participants compared to our targets has contributed to the pattern of results. Participants with either higher SIC or from a more multicultural place could potentially perceive differently targets based on their ethnicity/accent and religion and not demonstrate the same prototype

bias as our sample did. Indeed, it would be very interesting to see how women identifying as German Muslims would categorize these targets. Experimentally, future studies could test categorization before and after a habituation phase, where participants are confronted with different stereotype-inconsistent exemplars over different periods of time and contexts.

Additionally, one might wonder whether the present findings could be due to a possible confound in the similarity of the voices and faces used. However, targets were counterbalanced across the two conditions, with prototypical targets in one condition being non-prototypical in the other. Therefore, individual stimulus features cannot account for our findings. With regard to guessing, participants showed no bias in guessing any of the categories. This supports the conclusion that both categories (ethnicity/accent and religious-cue) were equally important in processing target information. Furthermore, the choice to use “neutral” statements was made to not prompt guessing biases. This could also be seen as a potential limitation, and it would be an interesting extension to see how the stereotypicality of statements affects individual and category memory for targets.

In order to increase experimental control, we used the same neutral faces that appeared neither highly prototypically Arabic or German and thus could be used both in the headscarf and in the no-headscarf condition without adding further to the complexity of the experiment. Future research could test whether the present findings generalize to targets who look more typical of the respective social category. One could speculate that as previous evidence suggests that accent trumps (facial) appearance when it comes to ethnic categorization (Rakić et al., 2011), using more typical ethnic appearance would be redundant. Still, because this type of experiment refrains from either explicitly naming or asking participants to name the given categories it is speculative whether they associate accent to a particular ethnicity (e.g., Birney et al., 2020; Dragojevic et al., 2018). Participants are

however very good in determining that a given accent is foreign, suggesting that our participants used accent and headscarf as proxies for foreignness.

Conclusion

We found that when asked to form an impression of a group, participants tended to rely on prototypicality of targets based on their ethnicity/accent and religion. In other words, the intersectionality between ethnicity/accent and religion was shown both in terms of how targets were individually remembered as well as categorized. Though this study further confirms the importance of accents and language in social categorization, the challenge of fully understanding how exactly language is being used is still open. This is a challenge for studies on language in the XXI century: further continue to use and explore the impact of language and accents in everyday encounters. How people perceive appears slightly different from how they explicitly explain it (use of accents for categorization vs. low ability to correctly identify accents, see Birney et al., 2020). Also, the language used to refer to certain groups (e.g., immigration background) and consequently the tendency to perceive someone as ingroup or outgroup matters. Though we may never find an answer to why certain stereotypes started to exist, comprehending how people tend to perceive certain categories and related stereotypes, as well as which categories are processed independently or concurrently can have important implications in understanding how to promote stereotype change. This is a challenge that we are hoping will be addressed in future research and theory development.

References

- Asbrock, F. (2010). Stereotypes of social groups in Germany in terms of warmth and competence. *Social Psychology, 41*(2), 76–81. <https://doi.org/10.1027/1864-9335/a000011>
- Berthold, A., Steffens, M. C., & Mummendey, A. (2019). What did they say? How subgroup stereotypes influence memory for superordinate groups. *Journal of Experimental Social Psychology, 83*, 23-36. <https://doi.org/10.1016/j.jesp.2019.02.007>
- Biernat, M., & Sesko, A. (2013). Evaluating the contributions of members of mixed-sex work teams: Race and gender matter. *Journal of Experimental Social Psychology, 49*, 471–476. <http://dx.doi.org/10.1016/j.jesp.2013.01.008>
- Birney, M. E., Rabinovich, A., & Morton, T. A. (2020). Where are you from? An investigation into the intersectionality of accent strength and nationality status on perceptions of non-native speakers in Britain. *Journal of Language and Social Psychology, 39*(4). [THIS ISSUE - Please add doi]
- Bruner, J. S. (1957). On perceptual readiness. *Psychological Review, 64*(2), 123-152. <https://doi.org/10.1037/h0043805>
- Carlston, D. E., & Schneid, E. D. (2015). When literatures collide: Synergies between stereotyping and impression formation. In S. J. Stroessner & J. W. Sherman (Eds.), *Social perception: From individuals to groups* (pp. 93-125). Psychology Press. <https://doi.org/10.4324/9781315734798>
- Corneille, O., Huart, J., Becquart, E., & Brédart, S. (2004). When memory shifts toward more typical category exemplars: Accentuation effects in the recollection of ethnically ambiguous faces. *Journal of Personality and Social Psychology, 86*(2), 236-250. <https://doi.org/10.1037/0022-3514.86.2.236>

- Crisp, R. J., & Hewstone, M. (2007). Multiple social categorization. In M. P. Zanna (Ed.), *Advances in experimental social psychology*, (Vol. 39, pp. 163-254). Elsevier Academic Press. [https://doi.org/10.1016/s0065-2601\(06\)39004-1](https://doi.org/10.1016/s0065-2601(06)39004-1)
- Dolderer, M., Mummendey, A., & Rothermund, K. (2009). And yet they move: The impact of direction of deviance on stereotype change. *Personality and Social Psychology Bulletin*, 35(10), 1368-1381. <https://doi.org/10.1177/0146167209339783>
- Dragojevic, M., Berglund, C., & Blauvelt, T. K. (2018). Figuring out who's who: The role of social categorization in the language attitudes process. *Journal of Language and Social Psychology*, 37(1), 28-50. <https://doi.org/10.1177/0261927x17706942>
- Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In D. T. Gilbert & G. Lindzey (Eds.), *The handbook of social psychology* (4th ed., Vols. 1 and 2, pp. 357-411). McGraw-Hill. <https://doi.org/10.1002/9780470561119>
- Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category based to individuating processes: Influence of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 1-74). Academic Press. [https://doi.org/10.1016/s0065-2601\(08\)60317-2](https://doi.org/10.1016/s0065-2601(08)60317-2)
- Giles, H., Bourhis, R. Y., & Taylor, D. M. (1977). Towards a theory of language in ethnic group relations. In H. Giles (Ed.), *Language, ethnicity and intergroup relations* (pp. 307-348). Academic Press.
- Giles, H., & Johnson, P. (1981). The role of language in ethnic group relations. In J. C. Turner & H. Giles (Eds.), *Intergroup behavior* (pp. 199-243). Blackwell.
- Giles, H., & Johnson, P. (1987). Ethnolinguistic identity theory: A social psychological approach to language maintenance. *International Journal of the Sociology of Language*, 68, 69-99. <https://doi.org/10.1515/ijsl.1987.68.69>

- Giles, H., & Rakić, T. (2014). Language attitudes: The social determinants and consequences of language variation. In T. Holtgraves (Ed.), *Oxford handbook of language and social psychology* (pp. 11-26). Oxford University Press.
<https://doi.org/10.1093/oxfordhb/9780199838639.013.030>
- Goff, P. A., Thomas, M. A., & Jackson, M. C. (2008). 'Ain't I a woman?': Towards an intersectional approach to person perception and group-based harms. *Sex Roles, 59*(5-6), 392-403. <https://doi.org/10.1007/s11199-008-9505-4>
- Hansen, K., Rakić, T., & Steffens, M. C. (2017). Competent and warm? How mismatching appearance and accent influence first impressions. *Experimental Psychology, 64*(1), 27-36. <https://doi.org/10.1027/1618-3169/a000348>
- Hugenberg, K., & Wilson, J. P. (2013). Faces are central to social cognition. In D. E. Carlston (Ed.), *Oxford handbook of social cognition* (pp. 167-193). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199730018.013.0009>
- Klauer, K. C., Ehrenberg, K., & Wegener, I. (2003). Crossed categorization and stereotyping: Structural analyses, effects patterns, and dissociative effects of context relevance. *Journal of Experimental Social Psychology, 39*(4), 332-354.
[https://doi.org/10.1016/s0022-1031\(03\)00017-9](https://doi.org/10.1016/s0022-1031(03)00017-9)
- Klauer, K. C., & Wegener, I. (1998). Unraveling social categorization in the 'who said what?' paradigm. *Journal of Personality and Social Psychology, 75*(5), 1155-1178.
<https://doi.org/10.1037//0022-3514.75.5.1155>
- Lambert, W. E., Hodgson, R. C., Gardner, R. C., & Fillenbaum, S. (1960). Evaluational reactions to spoken languages. *The Journal of Abnormal and Social Psychology, 60*(1), 44-51. <https://doi.org/10.1037/h0044430>

- Macrae, C. N., & Bodenhausen, G. V. (2000). Social cognition: Thinking categorically about others. *Annual Review of Psychology, 51*, 93-120.
<https://doi.org/10.1146/annurev.psych.51.1.93>
- Macrae, C. N., Hewstone, M., & Griffiths, R. J. (1993). Processing load and memory for stereotype-based information. *European Journal of Social Psychology, 23*(1), 77-87.
<https://doi.org/10.1002/ejsp.2420230107>
- Moshagen, M. (2010). Multitree: A computer program for the analysis of multinomial processing tree models. *Behavior Research Methods, 42*, 42-53.
<https://doi.org/10.3758/brm.42.1.42>
- Niedlich, C., Steffens, M. C., Cattarius, M., & Michel, C. (2016, January 28-30). *Benefiting from stigmatized identities? – discrimination on the basis of interactions regarding sexual orientation, gender and ethnicity in hiring decisions* [Conference presentation]. SPSP 2016 Convention, San Diego, CA, United States.
- Oakes, P. (1987). The salience of social categories. In J. C. Turner, M. A. Hogg, P. Oakes, S. Reicher, & M. S. Wetherell (Eds.), *Rediscovering the social group: A self-categorization theory* (pp. 117-141). Blackwell. <https://doi.org/10.2307/2073157>
- Rakić, T. (2020). *Dataset for study on ethnicity-religion cross-categorization*. Lancaster University. [10.17635/lancaster/researchdata/103](https://doi.org/10.17635/lancaster/researchdata/103)
- Rakić, T., Steffens, M. C., & Mummendey, A. (2011). Blinded by the accent! The minor role of looks in ethnic categorization. *Journal of Personality and Social Psychology, 100*(1), 16-29. <https://doi.org/10.1037/a0021522>
- Richards, Z., & Hewstone, M. (2001). Subtyping and subgrouping: Processes for the prevention and promotion of stereotype change. *Personality & Social Psychology Review, 5*(1), 52-73. https://doi.org/10.1207/s15327957pspr0501_4

- Roccas, S., & Brewer, M. B. (2016). Social identity complexity. *Personality and Social Psychology Review, 6*(2), 88-106. https://doi.org/10.1207/s15327957pspr0602_01
- Schug, J., Alt, N. P., & Klauer, K. C. (2015). Gendered race prototypes: Evidence for the non-prototypicality of Asian men and Black women. *Journal of Experimental Social Psychology, 56*, 121-125. <https://doi.org/10.1016/j.jesp.2014.09.012>
- Sesko, A., & Biernat, M. (2010). Prototypes of race and gender: The invisibility of Black women. *Journal of Experimental Social Psychology, 46*(2), 356-360. <https://doi.org/10.1016/j.jesp.2009.10.016>
- Sesko, A., & Biernat, M. (2018). Invisibility of Black women: Drawing attention to individuality. *Group Processes & Intergroup Relations, 21*(1), 141–158. <https://doi.org/10.1177/1368430216663017>
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*(11), 1359-1366. <https://doi.org/10.1177/0956797611417632>
- Stangor, C., Lynch, L., Duan, C., & Glas, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology, 62*(2), 207-218. <https://doi.org/10.1037//0022-3514.62.2.207>
- Taylor, S. E., Fiske, S. T., Etoff, N. L., & Ruderman, A. J. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology, 36*(7), 778-793. <https://doi.org/10.1037//0022-3514.36.7.778>
- Verkuyten, M. (2007). Religious group identification and inter-religious relations: A study among Turkish-Dutch Muslims. *Group Processes & Intergroup Relations, 10*(3), 341-357. <https://doi.org/10.1177/1368430207078695>

- Viruell-Fuentes, E. A., Miranda, P. Y., & Abdulrahim, S. (2012). More than culture: Structural racism, intersectionality theory, and immigrant health. *Social science & medicine*, 75(12), 2099-2106. <https://doi.org/10.1016/j.socscimed.2011.12.037>
- Weber, B. M. (2015). Gender, race, religion, faith? Rethinking intersectionality in German feminisms. *European Journal of Women's Studies*, 22(1), 22-36. <https://doi.org/10.1177/1350506814552084>
- Wenzel, M., Mummendey, A., & Waldzus, S. (2007). Superordinate identities and intergroup conflict: The ingroup projection model. *European Review of Social Psychology*, 18, 331-372. <https://doi.org/10.1080/10463280701728302>
- Zàrate, M. A., & Smith, E. R. (1990). Person categorization and stereotyping. *Social Cognition*, 8(2), 161-185. <https://doi.org/10.1521/soco.1990.8.2.161>

Appendix A

Table A1

Frequencies of Statements from Different Sources Assigned to Targets or New Items in the Matching Task

Assignment to:		An incorrect speaker from set of				
Source	The correct speaker	Headscarf	No Headscarf	Headscarf	No Headscarf	New items
		Arabic accent	Arabic accent	German accent	German accent	
Headscarf	1 (N=109)	2 (N=59)	3 (N=68)	4 (N=73)	5 (N=30)	6 (N=105)
Arabic accent						
No Headscarf	7 (N=69)	8 (N=72)	9 (N=67)	10 (N=57)	11 (N=34)	12 (N=145)
Arabic accent						
Headscarf	13 (N=46)	14 (N=59)	15 (N=63)	16 (N=49)	17 (N=59)	18 (N=138)
German accent						
No Headscarf	19 (N=111)	20 (N=50)	21 (N=59)	22 (N=67)	23 (N=35)	24 (N=122)
German accent						
New	n.a.	25 (N=31)	26 (N=29)	27 (N=28)	28 (N=30)	29 (N=1658)

Note. n.a., not applicable. The response categories are numbered from 1 to 29 with respective obtained response frequencies in brackets.

Author Biographies

Tamara Rakić is a Lecturer in Social Psychology at Lancaster University, UK. Her research interests include the influence of accents on person perception, categorization, and accommodation; and different language and cross-cultural aspects of social psychology.

Melanie C. Steffens is a full professor of social psychology at the University of Koblenz–Landau and the head of the Department of Social, Environmental, and Economic Psychology. Her research interests are centered on gender and diversity, attitudes toward lesbians and gay men, implicit processes, and memory phenomena.

Atena Sazegar is a researcher at Department of Medical Psychology, Research Group on Psychosocial Migration Research, University Medical Center Hamburg-Eppendorf, Germany.