

SUBMITTED VERSION

Stereotype Content of Refugee Subgroups in Germany

Patrick F. Kotzur, ORCID: <https://orcid.org/0000-0002-5193-3359>

University of Osnabrück, Philipps-University of Marburg

Maria-Therese Friehs, ORCID: <https://orcid.org/0000-0002-5897-8226>

University of Osnabrück

Frank Asbrock

Chemnitz University of Technology

Maarten H. W. van Zalk

University of Osnabrück

First authorship: Patrick F. Kotzur and Maria-Therese Friehs share the first authorship.

Corresponding author: Patrick F. Kotzur, Department of Psychology, University of Osnabrück, Seminarstraße 20, 49074 Osnabrück, Germany (Phone: +49163 7201303; E-mail: pakotzur@uos.de)

Conflict of Interest: The authors declare no conflict of interest. Acknowledgments: We thank Peter Schmidt for his advice on conducting our analyses, as well as Lea Müller and Leonie Schmidt for their support preparing this manuscript.

NOTE: This paper was accepted 11 March 2019 at the European Journal of Social Psychology. Full citation: Kotzur, P. F., Friehs, M.-T., Asbrock, F., & van Zalk, M. H. W. (2019). Stereotype content of refugee subgroups in Germany. *European Journal of Social Psychology*. doi: 10.1002/ejsp.2585

Stereotype Content of Refugee Subgroups in Germany

Word Count:

Abstract: 136

Main text, Abstract and Footnotes: 9370

Graphics: 3 Tables, 1 Figure

Online Supplementary Material (OSF): 10 Tables

Abstract

Stereotypes of refugee subgroups are still understudied. We contribute to this body of research by investigating differences in stereotype content, meaning warmth and competence ratings, of refugee subgroups in Germany ($N = 264$). Most extant Stereotype Content Model research is based on observed warmth and competence means values. We applied latent variable modelling using the alignment optimisation to ensure meaningful and reliable mean value comparisons. Generic refugees were rated as lacking warmth and competence. Warmth assessments of refugee subgroups varied depending on flight motives, geographical origin, and religious affiliation, implying that perceptions of threat and competition differed between these subgroups. Less differences emerged in competence assessments, indicating that refugee groups are generally regarded as lacking status. Our results enhance knowledge of the stereotype content of refugee subgroups and make a methodological contribution to stereotype content research.

Keywords: stereotype content, refugees, subgroups, Germany, alignment optimisation

Stereotype Content of Refugee Subgroups in Germany

March 2016: A German news platform describes the situation of migrants in Athens, explaining that for some time, only *War refugees*¹ were allowed to continue on the so-called “Balkan route”, while *Economic refugees*² were forced to remain behind (n-tv, 2016). June 2016: When analysing the 2015 crime statistics of the federal criminal agency, the widely disseminated German newspaper DIE WELT explains why refugees of certain origins, especially from North Africa or the Balkan states, tend to become more delinquent than immigrants of other origins (Hackensberger, Kalnoky, & Smirnova, 2016). September 2016: The Christian-conservative Bavarian political party CSU demands that Christian migrants should be preferred over Muslim ones in the German migration system (“CSU will Zuwanderer aus dem christlichen Kulturkreis bevorzugen”, 2016). These instances exemplify various public discourses in Germany since the beginning of the so-called “refugee crisis” in 2015, when about 890,000 people sought refuge in Germany (Bundesministerium des Innern, für Bau und Heimat [BMI], 2016). Notably, these narratives suggest that refugees have not been represented as one homogenous social group in Germany, but that there are different subtypes of refugees to be distinguished along several dimensions, including the perceived flight motives, religious affiliation, and geographic origin (“CSU will christliche Zuwanderer bevorzugen”, 2016; Hackensberger et al., 2016; n-tv, 2016). Recently, empirical studies have begun to investigate whether this differentiation results in different social perceptions of flight- and migration-related subgroups (Bansak, Hainmueller, & Hangartner, 2016; Ditzmann, Koopmans, Michalowski, Rink, & Veit, 2016; Kotzur, Forsbach, & Wagner, 2017).

Further addressing this issue is of high scientific and social relevance: European countries are, and most likely will continue to be, important destinations for people seeking refuge (Eurostat, 2016, 2018), making refugees a relevant social group in these countries. This is especially true for Germany, which received more than a third of all first-time applications

within the European Union in 2015 (Eurostat, 2016). However, insights regarding factors shaping receiving society members' perception of the newcomers, such as perceived characteristics of refugees, are scarce. The social perception of refugees has profound consequences, as it is likely to govern refugee-receiving community relations, as well as the broader context of reception of immigrant groups in general (Kotzur, Tropp, & Wagner, 2018). Additionally, the social perception of groups influences whether they are supported or harmed (Cuddy, Fiske, & Glick, 2007). Identifying subgroups running elevated risks of becoming targets of hostility and aggression allows for target group-specific social interventions to improve receiving society-refugee relations. Therefore, understanding factors that impact the social perception of refugees is an important research goal.

While pursuing these goals, we acknowledge that vital statistical preconditions for conducting substantive group comparisons in the social perception literature have only scarcely been tested (i.e., measurement invariance; but see Janssens, Verkuyten, & Khan, 2015; Stanciu, Cohrs, Hanke, & Gavreliuk, 2017). We do so by analysing social perception on a latent level, which allows controlling for reliability differences and assuring valid comparison of constructs across social groups (Kline, 2010). Specifically, we apply the alignment optimisation method (Asparouhov & Muthén, 2014) to establish approximate measurement invariance in our research, a recent, researcher-friendly statistical technique that both tests for and achieves the necessary preconditions for latent mean comparisons.

Stereotype Content of Refugee Subgroups in Germany

An influential theoretical framework to study social perception of groups, such as subgroups of refugees, is the Stereotype Content Model (SCM; Fiske, Cuddy, Glick, & Xu, 2002). The SCM proposes that culturally shared stereotypes towards social groups in a given society are based on two fundamental dimensions of social perception: Warmth, the “potential harm or benefit of the target group’s goals” (Cuddy et al., 2007, p. 632) in relation to one’s ingroup’s goals; and competence, the “degree to which the group can effectively enact those

goals” (Cuddy et al., 2007, p. 632). Stereotype content research focusses on culturally shared stereotypes in a given society by asking participants to indicate what they assume most society members think about a specific social group (Fiske et al., 2002). While this procedure is assumed to reduce social desirability, it implies that stereotype content predominantly taps into the perceived majority’s perspective on stereotypes, which can be different from individually endorsed stereotypes (Ashmore & Del Boca, 1981). Perceptions of threat and competition serve as predictors of the social group’s perceived intentions (i.e., warmth perceptions), whereas perceived status engenders competence perceptions (Binggeli, Krings, & Sczesny, 2014a; Fiske et al., 2002; Kervyn, Fiske, & Yzerbyt, 2015). The SCM has been frequently applied to describe cultural stereotypes of social groups in different national contexts (e.g., Asbrock, 2010; Binggeli et al., 2014a; Burkley, Durante, Fiske, Burkley, & Andrade, 2017; Bye, Herrebrøden, Hietland, Røyset, & Westby, 2014; Clausell & Fiske, 2005; Cuddy et al., 2009; Durante et al., 2013, 2017; Eckes, 2002; Janssens et al., 2015; Sadler, Meagor, & Kaye, 2012; Stanciu et al., 2017). The combination of both dimensions is theorised to predict contemptuous (low warmth/low competence; e.g., homeless people), envious (low warmth/high competence; e.g., rich people), and paternalistic (high warmth/low competence; e.g., elderly people) outgroup perceptions, as well as positive perceptions of allied and ingroups (high warmth/high competence; e.g., one’s own national group; Cuddy et al., 2009; Fiske, 2018). These, in turn, result in facilitative (for high warmth and/or high competence groups) and harmful (for low warmth and/or low competence groups) action tendencies towards these groups (Cuddy et al., 2007).

Despite these implications, there are no comprehensive studies investigating the stereotype content of (subgroups of) refugees. Research has shown in many country contexts – including Germany – that immigrant groups are generally rated low on warmth and competence (e.g., Asbrock, 2010; Binggeli, Krings, & Sczesny, 2014b; Eckes, 2002; Lee & Fiske, 2006; “The Fiske lab”, n.d.). Intriguingly, studies suggested that stereotype content of

generic social groups, meaning social groups without any further describing characteristics, do not need to correspond to specific subgroups, for instance, when additional subgroup information along key dimensions are provided (Binggeli et al., 2014b; Burkley et al., 2017; Bye et al., 2014; Clausell & Fiske, 2005; Eckes, 2002; Lee & Fiske, 2006; Sadler et al., 2012). For instance, in one study, *Women* as a general group were rated as warm and incompetent (Asbrock, 2010). Whereas the stereotype content of the specific subgroup *Housewives* matched this profile, *Career women* were rated as cold and competent (Asbrock, 2010). Recent studies, which we review in the following, provided initial evidence that this mechanism may also apply to present-day subgroups of refugees (Bansak et al., 2016; Binggeli et al., 2014b; Ditlmann et al., 2016; Kotzur et al., 2017; Lee & Fiske, 2006).

Although they did not explicitly examine subgroups of *refugees*, two studies investigated the stereotype content of immigrant subgroups: One in the U.S. (Lee & Fiske, 2006) and one in Switzerland (Binggeli et al., 2014b). Since refugees are an immigrant group, the same stereotype content-organizing principles may apply to refugee subgroups. In both studies, the researchers found that the *region or country of origin* served as an important cue of subgroups' perceived competition and status in the respective society, predicting subgroups' stereotype content. In the U.S., African immigrants were rated as warmer and less competent than immigrants from the Middle East (Lee & Fiske, 2006). In Switzerland, African immigrants were rated as warmer, but less competent, than immigrants from the Balkans³ (Binggeli et al., 2014b). Recent applicants for asylum in Germany originated mainly from these just-mentioned regions (Middle East: Syria, Afghanistan, Iraq; Balkan: Albania, Kosovo, Serbia; Africa: Eritrea; Juran & Broer, 2017). Indeed, the country of origin mattered as to whether participants were willing to grant asylum to refugees, or send them back to their country of origin (Bansak et al., 2016), arguably a benevolent (helping) or hostile (harming) behavioural intention towards distinct subgroups of origin, that may be reflected in warmth and competence ratings of these groups.

Along with Bansak et al. (2016), recent studies have begun to explicitly focus on the social perception of refugee subgroups in Germany and other European countries (Ditlmann et al., 2016; Kotzur et al., 2017). The only published SCM study investigating subgroups of refugees examined the impact of the *flight motive* on warmth and competence (Kotzur et al., 2017). The authors found that refugees fleeing due to economic reasons (*Economic refugees*) were rated significantly less warm than those that fled due to war (*War refugees*). No significant differences emerged on the competence dimension. The migration motive was an important predictor of the willingness to grant asylum in two further studies (Bansak et al., 2016; Ditlmann et al., 2016). Participants were less willing to welcome refugees fleeing from economic hardship than from war (Bansak et al., 2016; Ditlmann et al., 2016), lending further evidence to the importance of flight motive as a determinant of refugee subgroups' social perception. Consequently, the motive of migration may be an important predictor for the warmth ratings of refugee subgroups.

We have identified further attributes affecting the willingness to grant asylum that arguably may serve as cues for warmth and/or competence ratings (Bansak et al., 2016): For competence, potential cues include age, previous occupation, and language skills (Bansak et al., 2016). For warmth, and its precursors threat and competition, possible cues are religion and vulnerability (Bansak et al., 2016). *Religious affiliation* may be particularly relevant, since present-day refugees often originate from dominantly Muslim countries (Juran & Broer, 2017). Prior SCM studies found that Muslims were rated less warm and competent than Christians in societies where Christians are the majority ("The Fiske lab", n.d.), such as Germany. Consequently, Muslim refugees might be perceived as more threatening and of lesser status than Christian refugees, and thus as less warm and less competent.

Taken together, these findings indicate that generic refugees are rated relatively cold and incompetent. Prior SCM research focusing on other social groups suggested, however, that the stereotype content of subgroups may diverge from this generic view. Researchers

have begun to examine characteristics that may lead to shifts in the social perception of subgroups of refugees, mostly by examining the willingness to accept refugees with certain characteristics to one's country (but see Kotzur et al., 2017). We aim to contribute to this body of research using the SCM, a comprehensive social psychological framework that makes a range of predictions regarding antecedents and consequences of such perceptions.

Lastly, we identified two further gaps in this literature that need addressing. Firstly, the refugee subgroups to be studied had been exclusively studied with a top-down approach in prior research on refugee subgroup perception, that is, subgroup characteristics were selected by the researchers (Ditlmann et al., 2016; Bansak et al., 2016; Kotzur et al., 2017). Although there may be a wide range of characteristics of a social group that may be used to derive warmth and competence assessments, the question is still underexplored which of them are typically used by participants to meaningfully distinguish between refugee subgroups. Consequently, using a bottom-up approach might grant valuable insight into the characteristics that participants assume relevant to organise refugee subgroups. Secondly, prior studies have examined the social perception of refugees independently from other social groups. However, including reference groups, whose stereotype content within the SCM model has been reliably depicted within a specific society, is important for a meaningful contextualization of the relative location of refugee subgroups within the two-dimensional warmth by competence space (see, e.g., Lee & Fiske, 2006). Examining how the novel subgroups are rated relative to established social groups scoring particularly high or low on warmth or competence shows insight into the new groups' standing in society (i.e., in relation to important societal benchmarks).

Methodological Advances of SCM Research

In the spirit of the "crisis of confidence" (Kruglanski, Chernikova, & Jasko, 2017, p. 1) that has led researchers to question the appropriateness of the methods used and robustness of social psychological findings, we raise methodological concerns related to the extant SCM

research. One major criticism is that most SCM research is based on the analysis of observed values (i.e., computed scale means), confounding the true scale score with measurement error (Kline, 2010). Latent variable modelling is thus more appropriate in most cases, since it accounts for measurement error in the model (for a general readable introduction into the topic, see Cai, 2012). Another criticism is that most extant SCM research refrained from establishing measurement invariance (MI), that is, to test whether an “instrument measures the same concept in the same way across various subgroups of respondents” (Davidov, Meuleman, Cieciuch, Schmidt, & Billiet, 2014, p. 58; but see Janssens et al., 2015; Stanciu, 2015; for a general introduction into the topic, see Vandenberg & Lance, 2000). SCM research usually aims at comparing (aggregated) mean values of social groups’ assessments of warmth and competence in its analyses. This is often done using cluster analysis (e.g., Fiske et al., 2002), or observed mean value comparison (e.g., Kotzur et al., 2017), interpreting results for example as groups x_1, \dots, x_n are rated as warmer/less warm and/or more/less competent than groups y_1, \dots, y_n . Thus, measurement invariance is a key prerequisite to meaningful and valid mean value comparisons. The measures of warmth and competence should show at least (partial) scalar measurement invariance across target groups to avoid the proverbial comparison of “apples and oranges” (Chen, 2008) and to ensure that no systematic bias leads to over- or underestimation of any dimension between social groups (Vandenberg & Lance, 2000). Scalar MI is obtained when factor loadings and intercepts are constrained to be equal across groups, most commonly tested using multiple group confirmatory factor analysis (MGCFAs; van de Schoot, Schmidt, De Beuckelaer, Lek, & Zoder-van-Zwijnenburg, 2015). Partial measurement invariance refers to releasing the equality constraints, and thus the strict assumptions of the exact measurement invariance model, of highly deviating parameters while establishing equality for at least two other parameters (Byrne, Shavelson, & Muthén, 1989). We are aware of very few studies that have explicitly tested for measurement invariance of SCM constructs, thus ensuring meaningful and valid mean value comparisons of

observed and of latent variables between social groups (Janssens et al., 2015; Stanciu, 2015), all using the MGCFA approach.

A recent alternative is the multiple-group factor analysis alignment, or *alignment optimisation* method (Asparouhov & Muthén, 2014). This procedure allows to conduct latent mean value comparisons across groups while establishing a mathematically optimised (partial) scalar measurement invariance pattern (Asparouhov & Muthén, 2014). This approach is recommended for comparisons of large numbers of groups, and combines two advantages compared to other MI approaches: In the alignment optimisation procedure, (a) MI is tested less strictly and arguably enables a researcher to more realistically examine a broader range of group comparisons, without rejecting a group comparison early in the procedure because of MI violation (thus accounting for recent criticism of the MGCFA approach; see Van de Schoot, Kluytmans, Tummers, Lugtig, Hox, & Muthén, 2013); and (b) (partial) MI is established in an automated, easy-to-interpret manner using an algorithm that discovers “a solution where there are a few large noninvariant measurement parameters and many approximately invariant measurement parameters” (Asparouhov & Muthén, 2014, p. 3). This avoids cumbersome manual model improvements using modification indices that are required when establishing partial measurement invariance in the MGCFA framework, which might result in (a) potential errors made by the researcher at each manual modification index assessment, and (b) difficulties in replication of findings (Asparouhov & Muthén, 2014). Applied to the research at hand, alignment optimisation allows us to strengthen the internal and construct validity as well as the interpretability of our findings by generating (partial) measurement invariant latent mean values of warmth and competence assessments across a variety of target groups, while relaxing the oftentimes unrealistic exact measurement invariance assumption (Muthén & Asparouhov, 2013).

The Present Study’s Contributions and Predictions

With the present study, we intend to contribute to the literature by investigating stereotype content of subgroups of refugees. While researchers of extant studies investigating the social perception of refugee subgroups have selected the subgroup attributes themselves (Bansak et al., 2016; Ditlmann et al., 2016; Kotzur et al., 2017), we asked participants to nominate meaningful subgroup categories in a pilot study. This allowed us to investigate the stereotype content of a range of subgroup dimensions that are most likely meaningful to our participants (see, e.g., Binggeli et al., 2014a; Lee & Fiske, 2006). Moreover, we included reference groups that have been identified previously to score particularly high or low on either or both stereotype content dimensions in the present country context (Asbrock, 2010; Eckes, 2002), enabling us to map groups within the SCM space more comprehensively than prior studies (such as Kotzur et al., 2017).

Prior research indicates that generic refugees are rated relatively cold and incompetent (“The Fiske lab”, n.d.). We expected that subgroups of refugees diverge from this generic view: Specifically, we expected differences in warmth and competence ratings between subgroups, whereby country or region of origin, flight motives, and religious affiliations may serve as meaningful subgroup organisers (Bansak et al., 2016; Binggeli et al., 2014b; Ditlmann et al., 2016; Kotzur et al., 2017; Lee & Fiske, 2006). Based on previously reported findings, we expected that, as for flight motive, subgroups fleeing from war should be rated warmer than subgroups that fled for economic reasons (Bansak et al., 2016; Ditlmann et al., 2016; Kotzur et al., 2017). As for country or region of origin, refugees from African countries and regions are expected to be rated warmer and less competent than refugees from Balkan or Middle Eastern countries (Binggeli et al., 2014b; Lee & Fiske, 2006). As for religious affiliation, we expected Muslim refugees to be rated less warm and less competent than Christian refugees, based on findings in other SCM studies on religious groups (e.g., “The Fiske lab”, n.d.).

With regard to the societal reference groups, we selected groups that in prior research reliably scored in the four quadrants of the two-dimensional warmth-competence space (Asbrock, 2010; Fiske, 2018). In accordance with our expectation that refugees, and thus refugee subgroups, are rated relatively cold and incompetent, we expected that *Germans*, a high warmth-high competence reference group, receive higher values than refugee subgroups on both dimensions. *Elderly people*, a high warmth-low competence reference group, should receive higher warmth assessments and *Rich people*, a low warmth-high competence group, higher competence ratings compared to refugee subgroups. Lastly, we included *Homeless people* as a low warmth-low competence group. Based on Asbrock (2010), who found “foreigners” to be rated similar in warmth and higher in competence compared to “the homeless”, we assume that *Homeless people* are rated less competent than refugee subgroups⁴.

The last intended contribution was to move beyond observed mean analysis and cumbersome MGCFA and apply state-of-the-art methods appropriate to compare latent means of warmth and competence across large numbers of target groups. In order to allow for meaningful and internally as well as construct valid comparisons of the SCM dimensions, we employed the innovative alignment optimisation procedure that automatically establishes the best fitting (partial) scalar invariant solution and estimates as well as compares latent mean values of different target groups accordingly (Asparouhov & Muthén, 2014).

Methods

All data were collected at two German universities in the context of a larger research endeavour on the social perception of social groups. In Germany, the conduction of studies based on anonymous and confidential questionnaires that are not expected to entail any lasting harms or risks for the participants requires no additional permission by an internal review board. Thus, formally obtaining an internal review board approval was not necessary. All procedures were performed in full accordance with the ethical guidelines of the Deutsche

Gesellschaft für Psychologie (German Society for Psychology), and adhered to the low-risk study requirements of the universities where the studies have been carried out. Supplementary material, including the surveys, pilot study results, syntaxes, outputs, and results of additional analyses, has been made accessible on the Open Science Framework's website (see <https://osf.io/5j7t6/>). Raw data is available upon request from the corresponding author; it cannot be made publicly available since participants were informed that data management would be controlled by the study authors at all times.

Pilot Study and Subgroup Generation

We conducted an online pilot study to explore meaningful subcategory dimensions of refugees following the procedures described in Asbrock (2010). Participants ($N = 80$) were simultaneously recruited online through students' mailing lists from two mid-sized German universities ($n_1 = 40$ and $n_2 = 40$, respectively). Participants were not incentivised for their participation. All participants gave their informed consent prior to their inclusion in the study. To offer participants some guidance for this task (see Asbrock, 2010), they read the following instruction: "In the text box below, please list all migration-related groups in Germany that come to your mind. Please specifically consider groups with flight experiences and their backgrounds (e.g., relating to geographical, religious, flight cause characteristics). There are no right or wrong answers." The subsequent text box was unlimited. To increase participants' focus on listing groups, the online survey's continue-button was suppressed for 90 seconds (a full list of all identified groups and further procedural details how the open answers were coded can be found in <https://osf.io/rqk96/>).

In total, 83 and 108 groups were identified by the two samples of universities 1 and 2, respectively. We deemed groups that were mentioned by more than 20% of one of the samples to be meaningful and commonly used subgroup labels ($n > 8$; Asbrock, 2010), and consequently included them into the main study. As expected, subgroups with reference to flight motive, region or country of origin, and religion emerged: *Syrians, Afghans, Iraqis,*

Eritreans, Turks, North Africans, People from the Balkans, War refugees, Economic refugees, Muslims, Christians.

We amended all group descriptions with the label *refugees* (e.g., *Syrian refugees*), except for *Turks*. Historically, *Turks* have been economic migrants to Germany since the 1960s (Bade, 1992), so we labelled this group *Turkish migrants*. We kept the overall generic group *Refugees* in the list to receive reference information. Finally, we included four social groups that have shown to be located in the four extreme quadrants of the SCM as additional reference groups (*Germans* for high warmth/high competence, *Rich people* for low warmth/high competence, *Elderly people* for high warmth/low competence, and *Homeless people* for low warmth/low competence; see also Asbrock, 2010; Binggeli et al., 2014b; Fiske, 2018; Lee & Fiske, 2006).

Main Study

Participants and procedure. Using parallel online surveys, data for the main study were simultaneously collected between February and March 2017 using university-wide mailing lists from the same mid-sized German universities where we conducted our pilot studies. Bentler and Chou (1987) suggested 5 to 10 observations per estimated parameter for latent variable modelling. Asparouhov and Muthén (2014, p. 10) suggested “good recovery for all parameters except the factor variances is found already for $N_g = 100$ ” when applying alignment optimisation. We thus recruited $N = 264$ ($n_1 = 79$; $n_2 = 185$)⁵ German adults (72.3% female, 1.1 % other; $M_{age} = 24.21$, $SD_{age} = 4.65$; 95.5% university students, 4.6% other; 84.8% without migration background, 0.8% missing⁶) to fulfil these prerequisites. The two subsamples from the two universities did not differ significantly on any demographic variables (all $ps > .05$). Thus, we collapsed both subsamples to one joint sample on which we based all subsequent analyses. All participants gave their informed consent prior to their

inclusion in the study and were compensated with course credit and the opportunity to donate one Euro to a non-governmental aid organization of their choice.

The survey contained items concerning demographic information, stereotype content, participants' membership to one of the surveyed outgroups, and other variables not relevant to the study at hand (for a complete list of constructs we assessed, see Questionnaire: <https://osf.io/e3mqg/>). Following the procedure of previous SCM studies (e.g., Cuddy et al., 2009, study 1; Eckes, 2002), we presented stereotype content items one indicator per page, alternating warmth and competence indicators. For each indicator, we asked participants to evaluate the stereotype content of *Refugees* first, followed by a random order of refugee subgroups, *Turkish migrants*, and a random order of the non-migrant anchor groups. To prevent participant fatigue, we implemented a Three-Form-Design by inducing completely random, planned missingness on a subset of refugee groups (Graham, 2009; for an overview over the randomization, please see Online Supplementary Material Table SM1: <https://osf.io/y76kq/>).

SCM measure. We measured the stereotype content, and thus warmth and competence ratings, with the German-language SCM scales used in Asbrock (2010): "From the perspective of most Germans, how [ITEM] are the following social groups?". A social groups' warmth was assessed using the items "good-natured", "warm", and "likeable", while competence was assessed with "competent", "independent", and "competitive". Answers were given on a scale from 1 (*not at all*) to 5 (*very much*).

Analysis strategy. Following Asparouhov and Muthén (2014), we first assessed the general baseline measurement model fit for all 16 social groups under investigation using confirmatory factor analyses. We determined model fit to be adequate if all criteria formulated by Schermelleh-Engel, Moosbrugger, and Müller (2003) were met: $\chi^2/df < 3$; root mean standard error of approximation (RMSEA) $< .08$; standardised root mean square residual (SRMR) $< .10$; comparative fit index (CFI) $> .95$. Second, as a prerequisite for

alignment optimisation, we established a configural model across all groups with acceptable model fit; for these groups, “the number of subscales (i.e., factors), the location of the items (i.e., pattern by which items load onto each factor), and postulated correlations among the subscales (i.e., existence of covariances)” (Byrne, 2008, p. 873) were specified to be equal across groups. In a final step, we used *alignment optimisation*. This analysis strategy allowed us to estimate trustworthy latent means and comparing them for significant group mean differences while at the same time generating an optimised approximate measurement invariance pattern (Cieciuch, Davidov, & Schmidt, 2018).

Results

We conducted all analyses in Mplus Version 8, using robust maximum likelihood estimator (MLR) to account for multivariate non-normality and non-independence of observations (Muthén & Muthén, 1998-2017). Descriptive information on the stereotype content of the surveyed social groups in terms of warmth and competence are outlined in the Online Supplementary Material (see Table SM2: <https://osf.io/y76kq/>). Correlation tables of all indicators within one social group (Table SM4) and of warmth and competence scales within and across social groups (Table SM5) are also provided in the Online Supplementary Material.

Baseline Model Fit

Following the procedure described above, we first ran 16 single-group confirmatory factor analyses (one per target group). For each group, we specified one warmth factor with the indicators “good-natured”, “warm”, and “likeable” and one competence factor with the indicators “competent”, “independent”, and “competitive”. Warmth and competence factors were correlated, no cross-loadings or indicator residual covariations were allowed (for the syntax, see the folder Analysis material, 1 – Baseline models, e.g., <https://osf.io/y3hwq/>). Results are presented in Table 1. Ten out of 16 groups achieved an acceptable model fit: *Refugees, Syrian refugees, Muslim refugees, Afghan refugees, War refugees, Economic*

refugees, Refugees from Eritrea, Refugees from North Africa, Elderly people and Homeless people. The six remaining social groups *Christian refugees, Refugees from the Balkans, Iraqi refugees, Turkish migrants, Germans and Rich people* were discarded from further analysis due to non-acceptable model fit.

Please insert Table 1 about here

Alignment Optimisation Model

Configural measurement invariance. We entered the ten social groups showing adequate model fit into a simultaneous analysis for the configural measurement model. The model showed good fit, $\chi^2(80) = 114.890$, $p = .006$, $\chi^2/df = 1.436$, RMSEA = .046, SRMR = .034, CFI = .988, allowing us to focus on our research questions using the subsequent alignment optimisation procedure.

Measurement non-invariance. The fixed alignment optimisation solution we obtained with *Refugees* as a reference group showed two out of 120 parameters (two indicator intercepts; less than 2% of all parameters) to be non-invariant⁷. This finding indicated that a trustworthy estimation and comparison of latent warmth and competence means was possible, as the share of non-invariant parameters did not exceed 25% (Asparouhov & Muthén, 2014). The result was a latent mean value comparison based on a metric and partial scalar approximate measurement invariant model.

Latent mean values of stereotype content and significance testing. The ranking of the social groups, their latent mean values, and the significant differences to other social groups are outlined in Table 2 for warmth, and in Table 3 for competence. The findings are depicted in Figure 1. Further information on the alignment optimisation model (including factor loadings, indicator intercepts, factor means and variances, and the factor covariation of warmth and competence) are provided in Table SM3 in the Online Supplementary Materials (for more information, see <https://osf.io/y76kq/>).

Please insert Table 2 about here

Please insert Table 3 about here

Please insert Figure 1 about here

We examined differences in warmth and competence assessments between all refugee subgroups that had achieved acceptable baseline measurement model fit. As expected, both stereotype content dimensions showed statistically significant differences between refugee subgroups. In the following, we focus on selected group differences according to our predictions. For a complete list of significant differences that emerged on both warmth and competence between all groups, see Tables 2 and 3.

Differences regarding implied flight motive. Regarding the implied flight motive, we expected that *War refugees* should be rated warmer than *Economic refugees* (Kotzur et al., 2017). In accordance with this expectation, *War refugees* (latent factor mean $\alpha_W = 0.548$, rank 2) scored significantly higher on warmth than *Economic refugees* ($\alpha_W = -1.105$, rank 10). In fact, *War refugees* received the highest warmth ratings of all refugee subgroups. In contrast, *Economic refugees* showed the lowest warmth assessments, significantly lower than any other subgroup included.

We had not formulated any expectations regarding differences in terms of competence based on implied flight motive. In fact, prior research had found none (Kotzur et al., 2017). Unexpectedly, *Economic refugees* received the highest competence ratings of all refugee subgroups ($\alpha_C = 0.559$, rank 2); significantly higher than *War refugees* ($\alpha_C = 0.015$, rank 5). In sum, from all refugee subgroups, participants attributed the highest levels of warmth to *War refugees*. In accordance with our prediction, *War refugees* were rated substantially warmer than *Economic refugees*. These received the highest competence ratings of all refugee subgroups, differing significantly from *War refugees*.

Differences regarding origin. Due to non-acceptable baseline model fit, we could not include *Refugees from the Balkans*, *Iraqi Refugees* and *Turkish migrants* into the analysis at hand. For refugees from African countries, acceptable baseline model fit was found for

Refugees from Eritrea and *Refugees from North Africa*. For refugees from Middle Eastern countries, acceptable baseline model fit was found for *Syrian refugees* and *Afghan refugees*. Therefore, we compared these groups. Based on previous research on the stereotype content of immigrant subgroups of different countries and regions of origins conducted in Switzerland and the U.S. (Binggeli et al., 2014b; Lee & Fiske, 2006), we expected that refugees from African countries should be rated warmer than refugees from Middle Eastern countries.

Our expectations for warmth were not confirmed: *Syrian refugees* (Middle Eastern country) received the highest warmth ratings ($\alpha_w = 0.000$, rank 4) from all subgroups that indicated a region or country of origin; significantly higher than *Afghan refugees* ($\alpha_w = -0.375$, rank 7, Middle Eastern country), *Refugees from Eritrea* ($\alpha_w = -0.335$, rank 6; African country), and *Refugees from North Africa* ($\alpha_w = -0.854$, rank 9; African region). *Refugees from Eritrea* (African country) and *Afghan refugees* (Middle Eastern region) were rated significantly warmer than *Refugees from North Africa* (African region). Thus, contrary to our assumption, refugees from Middle Eastern countries were rated warmer than or non-significantly different in warmth from refugees of African origin.

In terms of competence, we expected that refugees from Middle Eastern countries should be rated as more competent than refugees from African countries (Binggeli et al., 2014b; Lee & Fiske, 2006). Our expectations were confirmed: *Syrian refugees* (Middle Eastern country) received highest competence ratings ($\alpha_c = 0.216$, rank 3), non-significantly different to *Afghan refugees* ($\alpha_c = -0.018$, rank 7; Middle Eastern country), but significantly higher than *Refugees from Eritrea* ($\alpha_c = -0.214$, rank 8; African country) and *Refugees from North Africa* ($\alpha_c = -0.348$, rank 9; African region). *Afghan refugees* (Middle Eastern country) were rated significantly more competent than *Refugees from North Africa* (African region), but non-significantly different from *Refugees from Eritrea* (African region). In sum, regarding origin, our predictions for warmth were contradicted, for competence partially confirmed. Refugees from Middle Eastern countries were rated warmer or not significantly different in

warmth compared to refugees from African countries, and partially more competent than refugee subgroups of African origin.

Differences regarding religious affiliation. We anticipated that *Muslim refugees* should be rated less benevolently than *Christian refugees*, i.e., receive lower warmth and competence ratings. Unfortunately, we were unable to test this expectation, since we had to exclude *Christian refugees* from our analyses due to non-acceptable baseline model fit. Nonetheless, the results did reveal that *Muslim refugees* were among the groups that received comparatively low warmth ratings ($\alpha_w = -0.639$, rank 8), providing indirect evidence for a depreciation of *Muslim refugees* relative to other refugee subgroups. *Muslim refugees* were however rated comparatively high in competence ($\alpha_c = 0.043$, rank 4). Thus, we found evidence that *Muslim refugees* were generally depreciated – at least on the warmth dimension – although we were unable to contrast *Muslim refugees* with *Christian refugees*.

Differences to reference groups. We were also interested in identifying the refugee subgroups' locations within the warmth by competence space in relation to societal reference groups. *Elderly people*, *Homeless people*, and generic *Refugees* were eligible for analysis, while *Germans* and *Rich people* had to be discarded due to non-acceptable model fit. Regarding warmth, we assumed *Elderly people*, a group particularly high on warmth (Fiske, 2018), to score highest of all groups – an assumption that was supported empirically. All groups received significantly lower warmth ratings than *Elderly people* ($\alpha_w = 1.931$, rank 1). Asbrock (2010) found “foreigners” to be rated similar in warmth to “the homeless”. Similarly, *Homeless people* ($\alpha_w = -0.108$, rank 5), showed non-significantly different warmth ratings from generic *Refugees* ($\alpha_w = 0.000$, rank 3). Yet, we found significant differences in warmth ratings of *Homeless people* to particular refugee subgroups in both directions: *Homeless people* were rated significantly less warm than *War refugees*, but significantly warmer than *Afghan refugees*, *Muslim refugees*, *Refugees from North Africa* and *Economic refugees*. Regarding generic *Refugees*, subgroups that were rated significantly less warm were

Refugees from Eritrea, Afghan refugees, Muslim refugees, Refugees from North Africa, and Economic refugees.

For competence, we assumed *Homeless people* to show lowest competence assessments. This expectation was empirically supported, as *Homeless people* ($\alpha_C = -0.733$, rank 10) indicated significantly lower competence ratings than any other social group. Surprisingly, all refugee subgroups received significantly lower competence ratings than *Elderly people* ($\alpha_C = 1.301$, rank 1), a group that has also been associated with low competence (Fiske, 2018). Generic *Refugees* ($\alpha_C = 0.000$, rank 6) were rated significantly less competent than *Economic refugees* and significantly higher in competence than *Refugees from North Africa* ($\alpha_C = -0.348$, rank 9).⁸

In sum, participants rated refugee subgroups overall less warm than *Elderly people*, a reference group that had been shown to score particularly high on warmth. Most subgroups were rated less warm than generic *Refugees*; only *War Refugees* were rated warmer. *Refugees* were rated non-significantly different in warmth compared to *Homeless people*, the reference group scoring particularly low on warmth. Whereas all subgroups were rated less competent than *Elderly people*, all groups were rated more competent than *Homeless people*, both low competence reference groups.

Discussion

This paper provided multiple insights in the contemporary stereotype content of refugees in Germany. We contributed by investigating the stereotype content of refugee subgroups in the SCM framework and applying state-of-the-art methods appropriate for comparing latent means of warmth and competence. Our results indicated that the stereotype content depended on the refugee subgroup in question. Our research thereby contributed to a growing body of literature that shows that stronger nominal differentiations of groups can lead to distinct stereotype content (Binggeli et al., 2014a, 2014b; Bye et al., 2014; Clausell & Fiske, 2005; Eckes, 2002; Lee & Fiske, 2006).

Stereotype Content of Refugee Subgroups in Germany

Prior research identified many dimensions along which refugees may be categorised into subgroups (Bansak et al., 2016; Ditlmann et al., 2016; Kotzur et al., 2017). Of those, flight motives, country or region of origin, and religious affiliation emerged as meaningful organisers of the social perception of refugee subgroups in our research. By asking participants to freely nominate subgroups of refugees within a society, we complemented previous research that has focused on researcher-generated subgroup dimensions (Bansak et al., 2016; Ditlmann et al., 2016; Kotzur et al., 2017).

As for flight motives, our results regarding warmth were in line with previous research and our corresponding expectations (Kotzur et al., 2017). *War refugees* were rated highly on warmth; in fact, higher than any other refugee group, indicating high levels of benevolence (Cuddy et al., 2007). In contrast, *Economic refugees* were rated less warm (the least warm of all refugee subgroups), indicating perceptions of threat and competition, and thus elevated risk to become recipients of outright rejection (Cuddy et al., 2007). Inconsistently with prior research that found no significant difference in competence ratings between both subgroups (Kotzur et al., 2017), *Economic refugees* received higher competence ratings than *War refugees*. Indeed, *Economic refugees* received the highest competence ratings of all refugee subgroups. These findings indicate that *Economic refugees* are seen as particularly capable of enacting their intentions (Fiske et al., 2002). Thus, in combination with the finding that this subgroup also rated the lowest warmth ratings suggests that *Economic refugees* are perceived as relatively skilled to enact their relatively harmful goals towards German society.

As for the country of origin, we expected that refugees from African countries and regions were rated warmer and less competent than refugees from Balkan or Middle Eastern countries (Binggeli et al., 2014b; Lee & Fiske, 2006). Whereas our expectations for competence were partially supported by the data (*Refugees from Eritrea* and *Refugees from North Africa* were rated less competent than *Syrian refugees*; *Refugees from North Africa* (yet

not *Refugees from Eritrea*) were rated less competent than *Afghan refugees*), our expectations for warmth were not. Warmth assessments did either not differ significantly between these refugee groups (*Afghan refugees* and *Refugees from Eritrea*), or the differences were in the opposite direction (*Syrian refugees* were rated more highly on warmth than *Refugees from Eritrea* and *Refugees from North Africa*). These unexpected findings may be related to the general observation that outgroup perceptions, particularly for racial, ethnic, and religious groups, can be context specific (Fiske, 2017). Thus, findings from other country contexts on immigrant groups do not necessarily need to be applicable to Germany. For instance, some Middle Eastern countries are regions of armed conflicts, such as wars and civil wars (Sørli, Gleditsch, & Strand, 2005). When people think about refugees fleeing from war, many Germans first think about refugees from Middle Eastern countries, especially Syria (Kotzur et al., 2017). Since our results showed that refugees fleeing from wars and armed conflicts were rated warmer than refugees that flee for other reasons, it may thus not be surprising that *Syrian refugees* and *Afghan refugees*, fleeing from war-ridden zones, were rated non-significantly different in warmth or even warmer than refugees from African regions. Moreover, the research we based our expectations on used labels referring to larger geographical units (e.g., Africa, Middle East; Binggeli et al., 2014b; Lee & Fiske, 2006), whereas we referred to specific subregions (e.g., North Africa) and countries (e.g., Syria, Afghanistan) within these geographical units. Just like overall perceptions of generic social groups do not necessarily correspond to subgroup perceptions, perceptions of groups from larger geographical units may not correspond to perceptions of specific subgroups within these regions. Two findings supported this conclusion: The finding that two African groups, namely *Refugees from North Africa* and *Refugees from Eritrea*, were rated significantly less warm than the groups from the Middle Eastern countries, and the finding that subgroups stemming from the same geographical region also differed in their stereotype content.

Moreover, note that *Homeless people* were usually perceived as among “the lowest of the low” (Fiske, 2018, p. 68; see also Asbrock, 2010, for a German sample), receiving the lowest warmth and competence scores. Our results showed, however, that *Refugees from North Africa* were rated lower on warmth, while also being rated low on competence. Thus, refugees from this region appeared to be among the most despised subgroups we have investigated in our study. This depreciation may potentially relate to perceptions of particularly high levels of threat emanating from *Refugees from North Africa* after nationwide media reports associating this subgroup with serious criminal offences (e.g., Drüeke, 2016; Hackensberger et al., 2016).

As for religious affiliation, we expected that *Muslim refugees* would be rated less benevolently than *Christian refugees* (“The Fiske lab”, n.d.). Although we had to exclude *Christian refugees* from our analyses, we found that *Muslim refugees* had an overall rather low rank in the warmth rating. *Muslim refugees* were also rated significantly less warm than generic *Refugees*. Both findings provide indirect evidence for a depreciation of refugees based on their belief – in line with findings that *Muslim refugees* would be granted asylum less often if participants were to decide (Bansak et al., 2016). Overall, then, our findings are highly compatible with societal discourses on refugees and provide an explanation of differential treatment of subgroups with different flight motives (n-tv, 2016), region or country of origin (Hackensberger et al., 2016), and religion (“CSU will christliche Zuwanderer bevorzugen”, 2016).

We found slightly more observable differences between refugee subgroups on the warmth compared to the competence dimension. An explanation may be that whereas different subgroups of refugees were associated with different levels of threat and competition (predictors of warmth; Fiske et al., 2002; Kervyn et al., 2015), refugees may have been perceived as a low status, and thus low competence immigrant group (Fiske et al., 2002); a perception that additionally specified characteristics could hardly change. Indeed,

institutionalised barriers and restrictions to access to the labour market (“Access to the labour market”, 2018), education, (“Access to education”, 2018) and other sources of status may limit the extent in which subgroups might be perceived differently on this dimension. Low status, then, may be a common and defining feature of all refugee groups we investigated.

Indeed, the analysis of the relative location to societal reference groups in the warmth by competence space corroborates this interpretation. We expected that refugee groups were rated relatively low on both warmth and competence. Our results confirmed the assumption that refugee subgroups are rated less warm than *Elderly people*, a high warmth-low competence reference group. Moreover, our results confirmed the assumption that refugee subgroups are rated more competent than *Homeless people*, a low warmth-low competence group. Thus, overall, the cultural stereotypes of refugee subgroups ranged from the low warmth-low competence area (that is, around the low warmth-low competence group *Homeless people*; see also Asbrock, 2010) up to the high warmth-low competence quadrant (that is, close to *Elderly people*, see also Asbrock, 2010).

The relative location within the warmth by competence space does not only hint at differential levels of threat, competition, and status associated with refugee subgroups, but also differential emotional and behavioural intentional consequences (Cuddy et al., 2007). That is, refugee subgroups are likely targets of either elevated contempt and harming intentions, or targets of pity and facilitative intentions, depending on their warmth and competence perceptions (high warmth-low competence or low warmth-low competence; Cuddy et al., 2007). As such, our analyses provide a first step towards identifying subgroups running elevated risks of becoming targets of hostility and aggression that allows for target group-specific social interventions to improve receiving society-refugee relations.

Methodological Advances of SCM Research

A further contribution of our work was that we applied recent and sophisticated methods to the study of stereotype content and social perception in general: By using latent

variable modelling instead of analyses based on observed means, which is the dominant approach in previously published SCM literature, we computed warmth and competence scores corrected for measurement error (Cai, 2012), thus increasing reliability of our findings. Moreover, by using the alignment optimisation procedure, we also greatly strengthened the validity of our findings (Asparouhov & Muthén, 2014). Surprisingly, the basic factor structure could not be established in six out of 16 cases, indicating that the two stereotype content dimensions proposed by the SCM could not be replicated empirically in all cases. If we had relied on observed variable analysis, this fact would have remained unnoticed, which might have resulted in comparisons of scale values that would not have validly represented equal warmth and competence constructs in some cases – the well-known “comparing apples with oranges”-problem. Thus, our procedure safeguarded us from erroneously including measures in our final analysis that did not fulfil the basic criteria for mean comparisons, and ultimately protected us from drawing inappropriate conclusions regarding mean differences in warmth and competence ratings across social groups.

Using confirmatory baseline modelling resulted in substantial data reduction, which in turn decreased the information and deductions we were able to present. Thus, our and others’ MGCFA findings suggest that measurement invariance is no naturally occurring scale characteristic (Janssen et al., 2015), and, consequently, should be examined carefully in all instances in SCM research. Given the novelty of applying confirmatory latent modelling procedures in SCM research, we can only speculate why some groups did not fit the baseline model. One implication would be that the global claim of the “generality across place, levels, and time” (Fiske, 2018, p. 67) of stereotype content dimensions may not be upheld without (to date unknown) boundary conditions. Therefore, we strongly recommend the increased usage of latent modelling approaches in future SCM research to cast more light on these questions and thus strengthen the empirical foundation of the theoretical framework of the SCM. We want to emphasize that we do not wish to devalue the vast body of prior SCM literature that

has contributed to the knowledge about social perception in important ways. To the contrary – we hope that our contribution helps the field to move forward; towards more robust, valid and authentic research demanded for in the light of the “crisis of confidence” in psychology (Kruglanski et al., 2017, p. 1).

Applying the alignment optimisation procedure to generate latent mean values apt for meaningful cross-group comparison appears to be a promising approach: Compared to manually establishing scalar measurement invariant models in an MGCFA, the alignment optimisation is less cumbersome and produces partial measurement invariant results that are not based on the individual decision of a researcher. In line with recent criticism towards the unduly strictness of the MGCFA approach (Muthén & Asparouhov, 2013), the alignment optimisation has less stringent prerequisites, thus producing an optimised approximate measurement invariance pattern when approaches like MGCFA failed (Asparouhov & Muthén, 2014; Cieciuch et al., 2018). Alignment optimisation is especially advantageous when large numbers of groups are to be compared – given that this is usually the case in SCM research, we feel that this approach is very promising for following SCM studies.

Limitations and Future Directions

We were not able to establish a baseline measurement model for more than 35% of groups. This is astounding given that we used warmth and competence scales that had been used in SCM studies in the German context before (e.g., Asbrock, 2010; Eckes, 2002; Kotzur et al., 2017) and that were developed directly from the original SCM scales (see Eckes, 2002). We attribute this loss of data to the more rigorous statistical analyses we applied compared to previously published research (but see Janssens et al., 2015; Stanciu et al., 2017), thus increasing our findings’ validity. The baseline model fit of the social groups might be improved through adapted warmth and competence scales: Firstly, including more items that tap into warmth and competence may enhance the reliability of the scales, and may allow ad-hoc adjustments (e.g., excluding certain items that underperform) to include more social

groups in our analyses. We are aware that this strategy would put a strain on the overall number of groups that can be studied within one survey. However, since a higher number of groups does not help to produce more insights when their scores cannot be compared, we would like to encourage future research to increase the numbers of items when measuring warmth and competence.

What is more, the established competence items tap mainly into economic and professional competence (“competitive”, “competent”, “independent”). However, other competence areas may be more associated with competencies of refugee groups, including withstanding threats to their survival, such as very adverse living conditions in their regions of origins, or threats while migrating. Thus, future research could explore whether our findings replicate with items tapping into other competence areas.

From its beginnings, stereotype content has been measured as a rating of how much participants assume a social group is viewed by most society members (Fiske et al., 2002). This strategy aims both at assessing cultural stereotypes shared by most members of a society, and at reducing social desirability bias in the data (Fiske et al., 2002). Nonetheless, this approach calls for cautious interpretation, as the results do not need to translate directly into participants’ personal perceptions of social groups, but rather display what participants believe how most Germans perceive groups. Thus, our results must not be understood as individual expressions of stereotypes. Although cultural and individual stereotypes may differ from each other (Ashmore & Del Boca, 1981), these potential deviations have not yet been systematically evaluated in the SCM framework. Especially in the important context of the social perception of (subgroups of) refugees, such a direct comparison appears a worthy goal of future research.

Further limitations relate to the sample composition: Like in numerous SCM studies and comparative research before us (Asbrock, 2010), we based our research on (typically young and liberal) student samples of two specific universities. Previous research found that student

samples, as well as representative samples, support the SCM structure (Fiske, 2015): Both student and representative samples lead to similar conclusions regarding the shared perception of social groups within a given society. Nonetheless, we would welcome studies that base their research on large representative samples to test whether this is also the case stereotype content of refugee subgroups.

Similar to prior research in this domain, we used a cross-sectional design. Although such designs are efficient to investigate the social perceptions of groups at a given time, they do not provide any indication regarding the stability of findings. The stereotype content of groups might change over time for several reasons; one of them being that intergroup contact with refugee groups can help to enhance warmth and competence assessments (Kotzur, Schäfer, & Wagner, 2018) - that is ever more likely to occur in refugee-receiving countries like Germany (Bundesamt für Migration und Flüchtlinge, 2019). Thus, we recommend the repeated assessment of social perception of refugee subgroups in Germany, ideally on a longitudinal basis.

As previously stated, SCM findings from one country context do not necessarily need to generalize to another ("The Fiske lab", n.d.). Moreover, some researchers found within-country regional differences in the endorsement of warmth and competence of social groups (Binggeli et al., 2014a; Stanciu et al., 2017). Although our data stemmed two different research sites within Germany, we found no such differences between samples (see additional analyses), although power for such comparative analyses was admittedly limited ($n_g < 100$ for some of the samples; Asparouhov & Muthén, 2014). Therefore, we encourage future studies investigating the social perception of refugee subgroups in other countries that are both destinations and origins of refugees, as well as potential within-country differences of the social perception of refugees. Such studies would further the understanding of the particularities of the social perception of this social group of high social relevance within and across cultures and nations.

Conclusion

In this study, we investigated the stereotype content of subgroups of refugees in Germany. Generic refugees were rated as lacking warmth and competence. Subgroup assessments differed significantly, depending on the insinuated flight motive, region or country of origin, or religious affiliation. Overall, the subgroups' warmth and competence ratings ranged from low warmth/low competence to high warmth/low competence. We produced these insights using alignment optimisation, an appropriate state-of-the-art method to compare multiple latent means. Given its relative user-friendliness and more realistic approach to measurement invariance compared to conventional methods, we hope will be adopted by others in this research field.

References

- “Access to education”. (2018). Retrieved from
<http://www.asylumineurope.org/reports/country/germany/reception-conditions/employment-education/access-education>
- “Access to the labour market”. (2018). Retrieved from:
<http://www.asylumineurope.org/reports/country/germany/reception-conditions/employment-education/access-labour-market>
- Asbrock, F. (2010). Stereotypes of social groups in Germany in terms of warmth and competence. *Social Psychology, 41*(2), 76-81. doi:10-1027/1864-9335/a000011
- Ashmore, R. D., & Del Boca, F. K. (1981). Conceptual approaches to stereotypes and stereotyping. In D. L. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior* (pp. 1-36). Hove, UK: Psychology Press.
- Asparouhov, T., & Muthén, B. O. (2014). Multiple-group factor analysis alignment. *Structural Equation Modeling, 21*, 1-14. doi: 10.1080/10705511.2014.919210

- Bade, K. J. (1992). *Deutsche im Ausland – Fremde in Deutschland [Germans abroad – foreigners in Germany]*. München: Beck.
- Bansak, K., Hainmueller, J., & Hangartner, D. (2016). How economic, humanitarian, and religious concerns shape European attitudes toward asylum seekers. *Science*, *354*(6309), 217-222. doi:10.1126/science.aag2147
- Bentler, P. M., & Chou, C. H. (1987). Practical issues in structural modeling. *Sociological Methods & Research*, *16*, 78-117.
- Binggeli, S., Krings, F., & Sczesny, S. (2014a). Perceived competition explains regional differences in the stereotype content of immigrant groups. *Social Psychology*, *45*(1), 62-70. doi:10.1027/1864-9335/a000160
- Binggeli, S., Krings, F., & Sczesny, S. (2014b). Stereotype content associated with immigrant groups in Switzerland. *Swiss Journal of Psychology*, *73*(3), 123-133. doi:10.1024.1421-0815/a000133
- Bundesamt für Migration und Flüchtlinge (2019). *Aktuelle Zahlen zu Asyl*. Retrieved from <http://www.bamf.de/SharedDocs/Anlagen/DE/Downloads/Infothek/Statistik/Asyl/aktuelle-zahlen-zu-asyl-dezember-2018.html?nn=7952222>
- Bundesministerium des Innern, für Bau und Heimat (2016, September 30). *Pressemitteilung – 890.000 Asylsuchenden im Jahr 2015*. Retrieved from <https://www.bmi.bund.de/SharedDocs/pressemitteilungen/DE/2016/09/asylsuchende-2015.html>
- Burkley, E., Durante, F., Fiske, S. T., Burkley, M., & Andrade, A. (2017). Structure and content of Native American stereotypic subgroups: Not just (ig)noble. *Cultural Diversity and Ethnic Minority Psychology*, *23*(2), 209-219. doi:10.1037/cdp0000100
- Bye, H. H., Herrebrøden, H., Hjetland, G. J., Røyset, G. Ø., & Westby, L. L. (2014). Stereotypes of Norwegian social groups. *Scandinavian Journal of Psychology*, *55*(5), 469-476. doi:10.1111/sjop.12141

- Byrne, B. M. (2008). Testing for multigroup equivalence of a measuring instrument: A walk through the process. *Psicothema, 20*(4), 872-882.
- Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and mean structures: The issue of partial measurement invariance. *Psychological Bulletin, 105*(3), 456-466.
- Cai, L. (2012). Latent variable modeling. *Shanghai Archives of Psychiatry, 24*(2), 118-120. doi:10.3969/j.issn.1002-0829.2012.02.010
- Chen, F. F. (2008). What happens if we compare chopsticks with forks? The impact of making inappropriate comparisons in cross-cultural research. *Journal of Personality and Social Psychology, 95*(5), 1005-1018. doi:10.1037/a0013193
- Cieciuch, J., Davidov, E., & Schmidt, P. (2018). Alignment optimization: Estimation of the most trustworthy means in cross-cultural studies even in the presence of noninvariance. In E. Davidov, P. Schmidt, & J. Billiet J. (Eds.), *Cross Cultural Analysis: Methods and Applications* (2nd ed., pp. 571-592). NY: Routledge.
- Clausell, E., & Fiske, S. T. (2005). When do subgroup parts add up to the stereotypic whole? Mixed stereotype content for gay male subgroups explains overall ratings. *Social Cognition, 23*(2), 161-181. doi:10.1521/soco.23.2.161.65626
- Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2007). The BIAS map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology, 92*(4), 631-648. doi:10.1037/0022-3514.92.4.631
- Cuddy, A. J. C., Fiske, S. T., Kwan, V. S. Y., Glick, P., Demoulin, S., Lexens, J.-P., ... Ziegler, R. (2009). Stereotype content model across cultures: Towards universal similarities and some differences. *British Journal of Social Psychology, 48*(1), 1-33. doi:10.1348/014466608X314935

- “CSU will Zuwanderer aus dem christlichen Kulturkreis bevorzugen” [Christian social union wants to privilege immigrants from Christian cultural milieus]. (2016, October 23). Die Welt. Retrieved August 28, 2018 from <https://www.sueddeutsche.de/politik/bayern-csu-will-vorrang-fuerzuwanderer-aus-christlich-abendlaendischem-kulturkreis-1.3153027>
- Davidov, E., Meuleman, B., Cieciuch, J., Schmidt, P., & Billiet, J. (2014). Measurement equivalence in cross-national research. *Annual Review of Sociology*, *40*, 55-75. doi:10.1146/annurev-soc-071913-043137
- Ditlmann, R., Koopmans, R., Michalowski, I., Rink, A., & Veit, S. (2016). Verfolgung vor Armut: Ausschlaggebend für die Offenheit der Deutschen ist der Fluchtgrund [Persecution for poverty: Decisive for the openness of Germans is the reason for flight]. *WZB Mitteilungen*, *151*, 1–27. Retrieved from https://www.wzb.eu/sites/default/files/publikationen/wzb_mitteilungen/veits24-27151-webpdf-2.pdf
- Drüeke, R. (2016). *Die TV-Berichterstattung in ARD und ZDF über die Silvesternacht 2015/2016 in Köln [Media analysis: Public television coverage of the Cologne New Year's Eve 2015/16]*. Gunda-Werner-Institut für Feminismus und Geschlechterdemokratie der Heinrich Böll Stiftung. Retrieved from www.gwi-boell.de/sites/default/files/web_161122_e-paper_gwi_medienanalysekoeln_v100.pdf
- Durante, F., Fiske, S. T., Kervyn, N., Cuddy, A. J., Akande, A. D., Adetoun, B. E., ... Barlow, F. K. (2013). Nations' income inequality predicts ambivalence in stereotype content: How societies mind the gap. *British Journal of Social Psychology*, *52*(4), 726-746. doi: 10.1111/bjso.12005
- Durante, F., Fiske, S. T., Gelfand, M. J., Crippa, F., Suttora, C., Stillwell, A., ... Teymoori, A. (2017). Ambivalent stereotypes link to peace, conflict, and inequality across 38 nations.

Proceedings of the National Academy of Sciences of the United States of America,
114(4), 669-674. doi: 10.1073/pnas.1611874114

- Eckes, T. (2002). Paternalistic and envious gender stereotypes: Testing predictions from the stereotype content model. *Sex Roles*, 47(3-4), 99-114. doi:10.1023/A:1021020920715
- Eurostat. (2016). *Asyl in den EU-Mitgliedsstaaten – Rekordzahl von über 1,2 Millionen registrierten erstmaligen Asylbewerbern im Jahr 2015* [Asylum in EU member states – record number of more than 1.2 million registered first-time asylum seekers in 2015]. Retrieved from <http://ec.europa.eu/eurostat/documents/2995521/7203837/3-04032016-AP-DE.pdf/9fcd72ad-c249-4f85-8c6d-e9fc2614af1b>
- Eurostat. (2018). – *Asylum report quarterly – First time asylum applicants and first instance decisions on asylum decisions: Fourth quarter 2017*. Retrieved from http://ec.europa.eu/eurostat/statistics-explained/index.php/Asylum_quarterly_report#cite_note-3
- Fiske, S. T. (2015). Intergroup biases: A focus on stereotype content. *Current Opinion in Behavioral Sciences*, 3, 45-50. doi:10.1016/j.cobeha.2015.01.010
- Fiske, S. T. (2017). Prejudice in cultural contexts: Shared stereotypes (gender, age) versus variable stereotypes (race, ethnicity, religion). *Perspectives on Psychological Science*, 12(5), 791-799. doi:10.1177/1745691617708204
- Fiske, S. T. (2018). Stereotype content: Warmth and competence endure. *Current Directions in Psychological Science*, 27(2), 67-73. doi:10.1177/0963721417738825
- Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82(6), 878-902. doi:10.1037/0022-3514.82.6.878
- Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, 60, 549-576. doi:10.1146/annrev.psych.58.110405.085530

- Hackensberger, A., Kalnoky, B., & Smirnova, J. (2016, June 09). *Warum Flüchtlinge aus diesen Ländern oft kriminell werden [Why refugees from these countries often become delinquent]*. *Die Welt*. Retrieved from <https://www.welt.de/politik/ausland/article156077320/Warum-Fluechtlinge-aus-diesen-Laendern-oft-kriminell-werden.html>
- Janssens, H., Verkuyten, M., & Khan, A. (2015). Perceived social structural relations and group stereotypes: A test of the Stereotype Content Model in Malaysia. *Asian Journal of Social Psychology*, *18*(1), 52-61. doi:10.1111/ajsp.12077
- Juran, S., & Broer, P. N. (2017). A profile of Germany's refugee populations. *Population and Development Review*, *43*(1), 149-157. doi:10.1111/padr.12042
- Kervyn, N., Fiske, S., & Yzerbyt, V. (2015). Forecasting the primary dimension of social perception. *Social Psychology*, *46*(1), 36-45. doi:10.1027/1864-9335/a000219
- Kline, R. B. (2010). *Principles and practise of structural equation modelling* (3rd ed.). New York: Guilford.
- Kotzur, P. F., Forsbach, N., & Wagner, U. (2017). Choose your words wisely: Stereotypes, emotions, and action tendencies toward fled people as a function of the group label. *Social Psychology*, *48*(4), 226-241. doi:10.1027/1864-9335/a000312
- Kotzur, P. F., Schäfer, S. J., & Wagner, U. (2018). Meeting a nice asylum seeker: Intergroup contact changes stereotype content perceptions and associated emotional prejudice, and encourages solidarity-based collective action intentions. *British Journal of Social Psychology*. Advance online publication. doi: 10.1111/bjso.12304.
- Kotzur, P. F., Tropp, L. R., & Wagner, U. (2018). Welcoming the unwelcome: How contact shapes contexts of reception for new immigrants in Germany and the United States. *Journal of Social Issues*, *74*(4), 812-832. doi:10.1111/josi.12300

- Kruglanski, A. W., Chernikova, M., & Jasko, K. (2017). Social psychology circa 2016: A field on steroids. *European Journal of Social Psychology, 47*(1), 1-10.
doi:10.1002/ejsp.2285
- Lee, T. F., & Fiske, S. T. (2006). Not an outgroup, not yet an ingroup: Immigrants in the stereotype content model. *International Journal of Intercultural Relations, 30*(6), 751-768. doi:10.1016/j.ijintrel.2006.06.005
- Muthén, B. O., & Asparouhov, T. (2013, January 11). BSEM measurement invariance analysis. *Mplus Web Notes* (Number 17). Retrieved from <http://www.statmodel.com/examples/webnotes/webnote17.pdf>
- Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus User's Guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.
- n-tv (2016). *Athen: Kaum noch Kriegsflüchtlinge*. [Athens: Hardly any war refugees anymore] Retrieved from <http://www.n-tv.de/politik/Athen-Kaum-noch-Kriegsfluechtlinge-article17144711.html>
- Sadler, M. S., Meagor, E. L., & Kaye, K. E. (2012). Stereotypes of mental disorders differ in competence and warmth. *Social Science & Medicine, 74*(6), 915-922.
doi:10.1016/j.socscimed.2011.12.019
- Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online, 8*(2), 23-74.
doi:10.1016/j.ssresearch.2003.11.003
- Sørli, M. E., Gleditsch, N. P., & Strand, H. (2005). Why is there so much conflict in the Middle East?. *Journal of Conflict Resolution, 49*(1), 141-165.
doi:10.1177/0022002704270824
- Stanciu, A. (2015). Four sub-dimension of stereotype content: Explanatory evidence from Romania. *International Psychology Bulletin, 19*(4), 14-20.

- Stanciu, A., Cohrs, J. C., Hanke, K., & Gavreliuk, A. (2017). Within-culture variation in the content of stereotypes: Application and development of the stereotype content model in an Eastern European culture. *The Journal of Social Psychology, 157*(5), 611-628.
doi:10.1080/00224545.2016.1262812
- “The Fiske lab: Intergroup relations, social cognition, and social neuroscience”. (n.d.)
Retrieved from www.fiskelab.org/cross-cultural-wc-maps
- van de Schoot, R., Kluytmans, A., Tummers, L., Lugtig, P., Hox, J., & Muthén, B. O. (2013). Facing off with Scylla and Charybdis: A comparison of scalar, partial, and the novel possibility of approximate measurement invariance. *Frontiers in Psychology, 4*, 770.
doi:10.3389/fpsyg.2013.00770
- van de Schoot, R., Schmidt, P., De Beuckelaer, A., Lek, K., & Zondervan-Zwijnenburg, M. (2015). Editorial: Measurement invariance. *Frontiers in Psychology, 6*, 1064.
doi:10.3389/fpsyg.2015.01064
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*(1), 4-70. doi:10.1177/109442810031002

Footnotes

¹ In German: “Kriegsflüchtling”, a term commonly used in Germany to refer to people that seek refuge due to war and civil war.

² In German: “Wirtschaftsflüchtling”, a term commonly used in Germany to refer to people that seek refuge due to economic hardship.

³ We calculated differences for these social groups based on the information Binggeli et al., (2014b) provided in Table 1, p. 128.

⁴ Although the authors are fully supportive of the aims and strategies for study preregistration of the Center for Open Science and other initiatives, we refrained from pre-registering these predictions as (I) they mostly stem from published research that was not based on the Stereotype Content Model framework (except for Kotzur et al., 2017) or that partially relied on the more general target groups (immigrants; Binggeli et al., 2014b, Lee & Fiske, 2006); and (II) due to the descriptive (i.e., with reference to the relative location of refugee subgroups compared to reference groups in the two-dimensional SCM framework) and exploratory (i.e., with reference to what societally relevant refugee subgroups would be generated) character of the study at hand.

⁵ In both subsamples, we excluded participants that provided answers for 50% or less of the variables ($n_1 = 6$; $n_2 = 46$), that were multivariate outliers as identified by Mahalanobi's distance ($n_2 = 13$), identified with at least one of the surveyed outgroups ($n_1 = 3$; $n_2 = 11$), had non-German nationality ($n_1 = 1$; $n_2 = 6$), or did not reside mainly in the respective area where the university was located ($n_1 = 4$; $n_2 = 1$).

⁶ The subsample 1 was 81.0% female, 1.3% other; ; $M_{age} = 23.49$ years, $SD_{age} = 4.76$, $Min_{age} = 19$, $Max_{age} = 50$; 98.7% University students, 1.3% other; 88.6% without migration background. The subsample 2 was 68.6% female, 1.6% other; $M_{age} = 24.52$ years, $SD_{age} = 4.57$, $Min_{age} = 19$, $Max_{age} = 50$; 94.1% University students, 6.0% other; 83.2% without migration background, 1.1% missing; 0.5% missing nationality.

⁷ For *Homeless people*, the intercepts of the competence item “independent” as well as the warmth item “likeable” were non-invariant.

⁸ Although the sample size is very low for such analyses (below $n = 100$ for subsample 1; Asparouhov & Muthén, 2014), we conducted all presented analyses also in each of the subsamples separately to rule out that the social perception of investigated groups differed systematically between sampling sites. The findings can be found in Tables SM6 and SM7 (for the baseline measurement model fit of all social groups in subsample 1 and 2, respectively), Tables SM8 and SM9 (for the results of the alignment optimisation approach for warmth and competence, respectively), Table SM10 (for further information on the alignment optimisation model) in the online supplementary material: <https://osf.io/y76kq/>. The analysis output can be found in the folder “Additional analysis: Separated for research location”, e.g., <https://osf.io/4h75y/>. No substantial differences emerged when comparing warmth and competence ratings of the same subgroups across subsamples. Thus, these additional analyses supported the findings presented above, lending further support for the robustness of our findings.

Tables

Table 1

Single Group Confirmatory Factor Analysis Model Fit

#	Group	<i>N</i>	χ^2	<i>df</i>	<i>p</i>	χ^2/df	<i>RMSEA</i>	<i>SRMR</i>	<i>CFI</i>
1	Refugees	264	18.008	8	.021	2.251	.069	.037	0.975
2	Christian Refugees	188	18.859	8	.016	2.357	.085	.048	0.955
3	Syrian Refugees	188	11.104	8	.196	1.388	.045	.033	0.991
4	Germans	264	27.566	8	<.001	3.446	.096	.043	0.940
5	Turkish Migrants	264	24.084	8	.002	3.011	.087	.035	0.967
6	Muslim Refugees	264	11.808	8	.160	1.476	.042	.030	0.991
7	Afghan Refugees	188	15.376	8	.054	1.922	.070	.039	0.977
8	Rich People	264	38.872	8	<.001	4.859	.121	.068	0.881
9	War Refugees	174	7.219	8	.513	0.902	.000	.029	1.000
10	Refugees from the Balkans	174	23.923	8	.002	2.990	.107	.044	0.953
11	Iraqi Refugees	174	30.485	8	<.001	3.811	.127	.057	0.929
12	Elderly People	264	15.995	8	.043	1.999	.062	.040	0.969
13	Economic Refugees	166	11.044	8	.199	1.381	.048	.035	0.988
14	Refugees from Eritrea	166	3.145	8	.925	0.393	.000	.019	1.000
15	Refugees from North Africa	166	3.226	8	.917	0.403	.000	.017	1.000
16	Homeless People	264	20.019	8	.010	2.502	.075	.038	0.962

Note. *N* = Number of participants; *df* = degrees of freedom; *p* = probability value; *RMSEA* = root mean square error of approximation; *SRMR* = standardized root mean square residual; *CFI* = comparative fit index. Acceptable model fit is indicated if all following requirements were fulfilled: $\chi^2/df < 3$; *RMSEA* < .08; *SRMR* < .10; *CFI* > .95 (Schermelleh-Engel et al., 2003). *Christian refugees*, *Germans*, *Turkish migrants*, *Rich people*, *Refugees from the Balkans*, and *Iraqi refugees* indicated poor model fit.

Table 2

Rank Order of Latent Mean Values for Warmth Assessment Across Social Groups

Rank	#	Group	Latent Mean Value	Groups with Significantly Smaller Factor Means
1	12	Elderly People	1.931	9, 1, 3, 16, 14, 7, 6, 15, 13
2	9	War Refugees	0.548	1, 3, 16, 14, 7, 6, 15, 13
3	1	Refugees†	0.000	14, 7, 6, 15, 13
4	3	Syrian Refugees	0.000	14, 7, 6, 15, 13
5	16	Homeless People	-0.108	7, 6, 15, 13
6	14	Refugees from Eritrea	-0.335	6, 15, 13
7	7	Afghan Refugees	-0.375	6, 15, 13
8	6	Muslim Refugees	-0.639	13
9	15	Refugees from North Africa	-0.854	
10	13	Economic Refugees	-1.105	

Note. Significance testing was conducted at a 5% significance level (two-sided). †Due to the fixed alignment optimisation model, this mean value was constrained to be zero.

Table 3

Rank Order of Latent Mean Values for Competence Assessment Across Social Groups

Rank	#	Group	Latent Mean Value	Groups with Significantly Smaller Factor Means
1	12	Elderly People	1.301	13, 3, 6, 9, 1, 7, 14, 15, 16
2	13	Economic Refugees	0.559	3, 6, 9, 1, 7, 14, 15, 16
3	3	Syrian Refugees	0.216	14, 15, 16
4	6	Muslim Refugees	0.043	14, 15, 16
5	9	War Refugees	0.015	15, 16
6	1	Refugees [†]	0.000	15, 16
7	7	Afghan Refugees	-0.018	15, 16
8	14	Refugees from Eritrea	-0.214	16
9	15	Refugees from North Africa	-0.348	16
10	16	Homeless People	-0.733	

Note. Significance testing was conducted at a 5% significance level (two-sided). [†]Due to the fixed alignment optimisation model, this mean value was constrained to be zero.

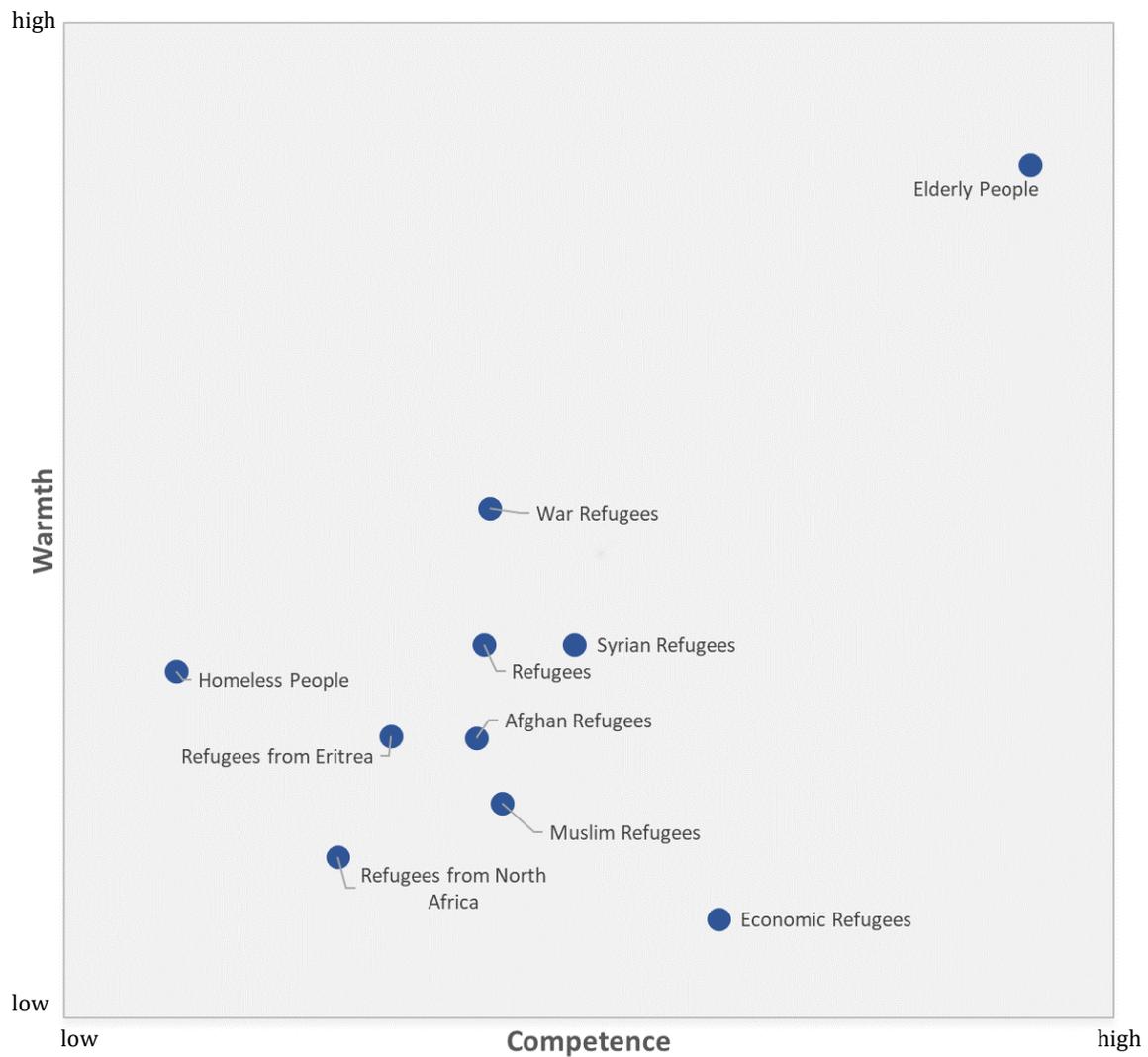
Figures

Figure 1. Latent warmth (Y-axis) and competence (X-axis) mean values for social groups. Scaling was achieved by constraining the latent mean values of warmth and competence of the group *Refugees* to zero.