Journal of Personality and Social Psychology
2016, Vol. 110, No. 5, 675–709

0022-3514/16/$12.00 http://dx.doi.org/10.1037/pspa0000046

The ABC of Stereotypes About Groups: Agency/Socioeconomic Success, Conservative–Progressive Beliefs, and Communion

Alex Koch
University of Cologne

Roland Imhoff
University of Cologne and
Johannes Gutenberg University Mainz

Ron Dotsch
Utrecht University and Radboud University Nijmegen

Christian Unkelbach and Hans Alves
University of Cologne

Previous research argued that stereotypes differ primarily on the 2 dimensions of warmth/communion and competence/agency. We identify an empirical gap in support for this notion. The theoretical model constrains stereotypes a priori to these 2 dimensions; without this constraint, participants might spontaneously employ other relevant dimensions. We fill this gap by complementing the existing theory-driven approaches with a data-driven approach that allows an estimation of the spontaneously employed dimensions of stereotyping. Seven studies (total $N = 4{,}451$) show that people organize social groups primarily based on their **a**gency/socioeconomic success (A), and as a second dimension, based on their conservative–progressive **b**eliefs (B). Communion (C) is not found as a dimension by its own, but rather as an emergent quality in the two-dimensional space of A and B, resulting in a 2D ABC model of stereotype content about social groups.

Keywords: social groups, stereotype content, agency, beliefs, communion/warmth

Supplemental materials: http://dx.doi.org/10.1037/pspa0000046.supp

Stereotypes are everywhere. To navigate their social world, people quickly group individuals in meaningful social categories based on their age, gender, ethnic origin, occupation, or interest (Brewer, 1988; Fiske & Neuberg, 1990; Tajfel, 1969). Knowledge about these categories includes what typical members of this category are like, think, feel and do, and the schematic application of this knowledge provides an economical alternative to effortful individuation (Fiske & Pavelchak, 1986; Gilbert & Hixon, 1991; Macrae, Milne, & Bodenhausen, 1994). And stereotypes matter. They allow people to go beyond the information given (Bruner, 1957), make predictions about the future behavior of individuals based on their sheer category membership (Hamilton, Sherman, & Ruvolo, 1990), and they influence people's judgments, decisions, and behavior in a stereotype-consistent way (Wheeler & Petty, 2001), even without being aware of this (Bargh, Chen, & Burrows, 1996).

Each stereotype consists of a more or less unique set of attributes associated with the social group: White, Black, Latino, Middle Eastern, and Asian men are "rich," "athletic," "macho," "bearded," and "intelligent," respectively. White, Black, Latino, Middle Eastern, and Asian women are "arrogant," "have an attitude," are "feisty," "quiet," and "intelligent," respectively (Ghavami & Peplau, 2012, pp. 118–120). Librarians are shy, hairdressers are flamboyant, and stock-traders are greedy. Some attributes, however, may be of greater importance for effectively coordinating social behavior than others and thus are likely to serve as content of stereotypes about many, if not all, groups. That is, some attributes may serve as fundamental dimensions of stereotype content that stretch out people's social maps on which groups can be located as a function of scoring low or high on the respective dimensions.

Warmth and Competence are Meaningful Stereotype Content Dimensions

According to the stereotype content model (SCM; Fiske, Cuddy, Glick, & Xu, 2002), the most relevant criteria in intergroup interaction are the social groups members' intentions and their ability to carry out their plans. The central question is whether a group has goals compatible with the perceiver and is thus likely to help him or her, or whether it has antagonistic goals and thus might harm him or her (Fiske et al., 2002). Knowing this (i.e., a group's warmth, Fiske et al., 2002; communion, Abele & Wojciszke, 2007; morality, Wojciszke, 1994; other-profitableness, Peeters, 1983;

Alex Koch, Social Cognition Center Cologne, University of Cologne; Roland Imhoff, Social Cognition Center Cologne, University of Cologne, and Social and Legal Psychology, Johannes Gutenberg University Mainz; Ron Dotsch, Social and Organizational Psychology, Utrecht University, and Behavioral Science Institute, Radboud University Nijmegen; Christian Unkelbach and Hans Alves, Social Cognition Center Cologne, University of Cologne.

This research was supported by a University of Cologne Advanced Postdoc Grant to Roland Imhoff. The first two authors contributed equally to this article. Order of the first two authors was decided on a junior–senior principle.

Correspondence concerning this article should be addressed to Alex Koch and Roland Imhoff, Social Cognition Center Cologne, University of Cologne, Richard-Strauss-Str. 2, 50931 Köln, Germany. E-mail: alex.koch@uni-koeln.de and roland.imhoff@uni-mainz.de

trustworthiness, Oosterhof & Todorov, 2008), the second most relevant question has been theorized to be a group's ability to carry out their intentions (i.e., competence, Fiske et al., 2002; agency, Abele & Wojciszke, 2007; Wojciszke, 1994; self-profitableness, Peeters, 1983; instrumentality, Parsons & Bales, 1955).

More than a decade of research on these two dimensions of the SCM (Fiske et al., 2002) suggests that whether a group is perceived as warm and/or competent has implications for emotional reactions to the group (Cikara & Fiske, 2012; Cuddy, Fiske, & Glick, 2007), neurological responses to the group (Harris & Fiske, 2006), people's perception of what typical group members look like (Imhoff, Woelki, Hanke, & Dotsch, 2013), as well as behavioral intentions of harming and helping (Becker & Asbrock, 2012; Cuddy et al., 2007) like invitations to a job interview (Agerström, Björklund, Carlsson, & Rooth, 2012), or support for immigration politics (Reyna, Dobria, & Wetherell, 2013). Even beyond groups, the two SCM dimensions have been employed to assess people's perceptions of brands (Aaker, Garbinsky, & Vohs, 2012; Kervyn, Chan, Malone, Korpusik, & Ybarra, 2014), exonerees (Clow & Leach, 2015), and individuals in pain (Ashton-James, Richardson, Williams, Bianchi-Berthouze, & Dekker, 2014).

Some of these studies adopt what we would call a *relational* approach and aim to explore how individuals determine their concrete behavior toward an individual from a group based on their assumptions about this group's warmth and competence (e.g., Kervyn, Dolderer, Mahieu, & Yzerbyt, 2010). Others have adopted what might be framed as a *lay sociologist* perspective, that is: On which dimensions do people identify the most relevant differences between social groups (Cuddy et al., 2007; Fiske et al., 2002; Imhoff et al., 2013)? The present research addresses particularly the latter perspective.

Warmth and Competence may not be the Dimensions That People Spontaneously Use

Within the lay sociologist perspective we argue that although warmth and competence are meaningful dimensions of stereotype content, we currently lack empirical support for the notion that these are indeed the dimensions that individuals *spontaneously* employ when making sense of social groups. Spontaneously employed dimensions are the ones that come to people's mind without theoretical constraints made by the researchers. Most studies on stereotype content constrain participants to the two theoretically derived dimensions, because in most cases only these two dimensions are rated (e.g., "participants rated the 15 groups on scales of warmth, competence, status, and competition," Cuddy et al., 2009, p. 12; "participants rated the 53 categories on either competence or warmth," Durante, Volpato, & Fiske, 2010, p. 473; "participants rated the groups on scales reflecting warmth, competence, perceived status, and perceived competition," Fiske et al., 2002, p. 884; "participants rated the extent to which each group appeared warm (friendly, cold (reversed), likable [. . .]) and competent (capable, incompetent (reversed), smart [. . .])," Bergsieker, Leslie, Constantine, & Fiske, 2012, p. 1229). For participants, it is thus impossible to employ any other stereotype content dimensions.

Another source of constraints is the selection of groups to be rated. Although some studies sampled groups spontaneously named by participants, the instructions prompted race, gender, occupation, and so forth as criteria of what constitutes groups, thereby biasing the likelihood of certain categories to be named (e.g., "Blacks," "women," and "professionals," see "off the top of your head, what various types of people do you think today's society categorizes into groups (i.e., based on ethnicity, race, gender, occupation, ability, etc.)?," Fiske et al., 2002, p. 883; see also Kervyn, Fiske, & Yzerbyt, 2013, p. 676). It is conceivable and highly likely that a biased sample of certain social groups will make certain stereotype dimensions more salient than others (e.g., prompting race and gender will make dimensions associated with race and gender more salient).

To give another example, Imhoff, Woelki, Hanke, and Dotsch (2013) showed that visual facial representations of typical exemplars of two social groups pretested as differing on warmth and competence were judged by other raters as differing on both warmth and competence. While this finding supports that people are able to associate warmth and competence with facial features, it does not rule out that the space of group stereotypes also includes one, two, or more additional—and potentially more fundamental—dimensions that were not visually encoded in the faces because the two pretested groups (managers and kindergartners) were not different on them, and/or that were not decoded from the faces because the researchers never asked for ratings other than warmth and competence.

Thus, the above-mentioned studies lack representative design (Brunswik, 1955, 1956). To illustrate this important aspect, imagine one wants to find out the fundamental dimensions people spontaneously use to compare cars. A nonrepresentative sample of cars of the same price, size, and fuel efficiency, but in different colors will probably prompt the result that the most fundamental dimension on which people spontaneously distinguish cars is their color. While that might very well be the case, the biased sampling prevented other dimensions from being detected because there was no meaningful variance on these other dimensions. Likewise, even if there is a representative sample of cars, but participants rate them only on the number of airbags and the maximum speed, this will give us a two-dimensional space on which all cars can be positioned, with one dimension being number of airbags, and the other being maximum speed. Crucially, though, we have no empirical base to judge whether these two dimensions are indeed the fundamental dimensions that individuals spontaneously employ when comparing cars even if we replicate the rating multiple times in many different environments. Without a more representative sampling approach, one cannot rule out that empirical findings are influenced by sampling biases (Fiedler, 2011). As much as we ideally draw representative participant samples from the population we aim to generalize to, a representative design also calls for an unbiased sampling of stimuli (to be able to generalize to the universe of stimuli) as well as dimensions (to generalize to the universe of attributes; see Wells & Windschitl, 1999; Westfall, Kenny, & Judd, 2014 for a more elaborate discussion of the problems of stimulus sampling and generalization).

In summary, we believe the available evidence for the nature of stereotype content dimensions about social groups suffers from (a) a nonrepresentative sampling of social groups, which prevents generalization to the population of groups; and (b) a nonrepresentative sampling of rated attributes, which prevents generalization to the population of all conceivable attributes.

How to Explore the Nature of Spontaneous Stereotype Content About Groups

To gain insights into the fundamental, spontaneously employed dimensions of stereotypes about groups, one thus needs a different approach that more closely follows the ideal of a representative design (Brunswik, 1955, 1956). In such a design, a sample of participants organizes a random (i.e., without any theoretical constraints) sample of stimuli on dimensions without being constrained in what these dimensions are. Sampling of groups can be achieved by asking people to name groups and selecting the most frequently named ones. In doing so, we avoid theory-driven a priori assumptions about the most relevant criteria for segmenting society into groups, such as age, sex, race, occupation, ability, and so forth.

Assessing fundamental dimensions on which people align social groups without influencing participants by naming theoretically derived candidate dimensions requires more effort. Here, we rely on a data-driven strategy; such data-driven methods have proven to be extremely successful tools to identify fundamental dimensions of social perception with as little bias as possible in areas like face and gender perception (Broverman, Vogel, Broverman, Clarkson, & Rosenkrantz, 1972; Deaux & Lewis, 1984; Ghavami & Peplau, 2012; Todorov, Dotsch, Wigboldus, & Said, 2011; Williams & Best, 1990).

One well-established data-driven method is multidimensional scaling based on global dissimilarity estimates (Nosofsky, 1992; Schiffman, Reynolds, & Young, 1981). In this approach, participants merely provide estimates of the similarity/dissimilarity between social groups. Importantly, they are free to rely on any dimension that spontaneously comes to their mind and seems most diagnostic to them for that decision. When judging for instance the similarity between lawyers, nurses, and maids, individuals could resort to relatively consensual impressions of warmth and thus see lawyers and maids as similar (cold), but both different from nurses who are seen as warm (see Fiske & Dupree, 2014). If competence, however, is the most salient and subjectively diagnostic dimension, participants should see lawyers and nurses as similar compared to the dissimilar (relatively incompetent) maids. Finally, it is conceivable that people make use of completely different characteristics and see assumed gender as more central, with nurses and maids as occupations typically perceived to be female-dominated compared with lawyers evoking associations with men.

Exploring the dimensionality of stimulus spaces in this way is well established in the social psychology of personality impressions (Good–Bad × Hard–Soft; Rosenberg, Nelson, & Vivekananthan, 1968), emotions (Valence × Intensity; Russell, 1980; Shaver, Schwartz, Kirson, & O'Connor, 1987), animals (Size × Ferocity; Henley, 1969), power strategies (Rationality × Directness; Falbo, 1977), and responses to dissatisfaction in the job and one's relationship (Active–Passive × Constructive–Destructive; Farrell, 1983; Rusbult & Zembrodt, 1983). More relevant to the focus of the present article, Pattyn, Rosseel, and van Hiel (2013) recently asked participants to complete a hierarchical sorting task to estimate dissimilarities between individuals who belonged to predefined social groups. Across three studies they reported converging support for five to six meaningful dimensions of the social group space (conventional vs. alternative, old vs. young, male vs. female, cognitive vs. physical, deviant vs. nondeviant, and to a lesser extent: cold vs. warm).

Although these results are thought-provoking as they suggest very different dimensions than the well-received stereotype content model, a closer look at their stimulus sampling procedure indicates that, as in all previous work, biased sampling might have again played a major role in producing these findings (for another example of the large impact of stimulus sampling on results see Frable, 1993; Jones & Ashmore, 1973). Specifically, the researchers searched for pictures of (male and female) individuals who belonged to a predefined set of social groups, among them "punk," "hippie," "yuppie," "typical woman," and "senior citizen." Accordingly, two of the central dimensions turned out to be conventional (typical woman) versus alternative (punks, hippies) as well as old versus young and a similar argument can be made for the other dimensions. This study thus illustrates how stimulus sampling may influence the inferred underlying dimensions.

The Present Research

This article aims to investigate the fundamental, spontaneously employed dimensions of stereotype content about social groups. To achieve this aim, we followed the proposed data-driven research strategy. We asked participants in two cultural contexts (U.S.-based MTurkers and German students) to name examples of what constitutes groups without biasing the selection by any examples or criteria. The groups that were most frequently named and appeared most often in contemporary mass media were then judged on dissimilarity to one another in order to compute stereotype maps of groups with multidimensional scaling. The dimensions of the emerging scaling solutions were then interpreted via property fitting analyses (Chang & Carroll, 1969) with a variety of candidate stereotype content dimensions on which the groups had been judged by independent raters. As these candidate dimensions may constitute an experimenter influence, we finally asked participants to label all rotated content dimensions that run through the origin of the groups' stereotype maps. Other independent raters confirmed that these labels did not reflect a dimension that was not included in our selection of candidate stereotype content dimensions. We believe this strategy avoided biases due to selective sampling of stimuli and/or dimensions and allowed participants to spontaneously employ any dimension they saw as important to distinguish between the groups that they saw as important to distinguish. In a total of seven studies with 4,451 participants, we found, confirmed, and generalized what we refer to as the 2D ABC model of spontaneous stereotypes about groups. According to the data, people distinguish groups based on differences in agency/socioeconomic success (A: "powerless–powerful," "poor–wealthy," "low status–high status," "dominated–dominating," "unconfident–confident," and "unassertive–competitive") and conservative–progressive beliefs (B: "traditional–modern," "religious–science–oriented," "conventional–alternative," and "conservative–liberal"). Further, the groups' communion/warmth (C; "cold–warm," "untrustworthy–trustworthy," "dishonest–sincere," "repellent–likable," "threatening–benevolent," and "egoistic–altruistic") emerges as a function of centrality in the stereotype map spanned by A and B. That is, groups that appear average on both dimensions appear to be warm, trustworthy, sincere, likable, benevolent, and altruistic. Just like the stereotype content model by Fiske, Cuddy, Glick, and Xu (2002; see also Cuddy et al., 2007), the 2D

ABC model addresses consensual rather than idiosyncratic group stereotypes.

We conducted five more studies within this project that we do not report for reasons of brevity. All studies consistently supported the pattern of results reported in this article. These 12 studies represent the full set of all studies we have conducted up to this point to explore the number and nature of the stereotype content dimensions that people spontaneously employ to distinguish large sets of social groups sampled without bias in favor of a specific stereotype content model.

Study 1

We first generated a large sample of social groups by asking people to name groups and then selected the most frequently named ones (consensus >10%). Then, new participants judged the dissimilarity between each group and each other group, allowing participants to spontaneously choose dimensions on which they base their judgment (Forgas, 1976; Rosenberg et al., 1968). Dissimilarity per se is unspecific and open to idiosyncratic interpretation—that is, it needs to be construed in one or another respect (Medin, Goldstone, & Gentner, 1993). The chosen dimensions might be different for each participant, but highly idiosyncratic approaches will be filtered out by aggregation across individuals so that the average pairwise estimates of the dissimilarities between the groups will reflect a consensual view. The dimensions might be different for each pairwise comparison and each dissimilarity rating might be a judgment based on the integration of many dimensions. However, as long as all participants employ more or less identical dimensions in making the dissimilarity judgments, the multidimensional scaling (MDS; for a review, see Borg & Groenen, 2005) algorithm will compute a multidimensional social space in which the groups' coordinates retain almost all the variance contained in the original dissimilarity judgments.

If there are fundamental stereotype content dimensions, then the next question is their nature, which can be addressed with a property fitting analysis (ProFit, (Chang & Carroll, 1969; e.g., Pattyn, Rosseel, & van Hiel, 2013) during which rating dimensions are sought that can be best predicted by the social groups' MDS coordinates. This approach is ideal to "help systematize data in areas where organizing concepts and underlying dimensions are not well-developed" (Schiffman, Reynolds, & Young, 1981, p. 3, see also Giguère, 2006). The properties to be fitted were 24 trait dimensions ("unfriendly–friendly," "incompetent–competent," etc.) that were identified as possible candidates of being fundamental to stereotype content, both in light of the data as well as established theories. While we diverge here from a purely data-driven approach, 24 dimensions present a much larger sample of possible candidates than in previous studies on the dimensionality and nature of spontaneous stereotype content about groups (e.g., Fiske et al., 2002). Studies 5 and 6 will solve this deviation from a purely data-driven approach and show that our selection of candidates included all stereotype content dimensions that participants employed to distinguish between groups. At this point we refrained from making predictions regarding the existence, number, and nature of the fundamental stereotype content dimensions for groups.

Method and Results

To avoid having an overly homogenous sample of undergraduate students, we recruited a more diverse sample in terms of educational and professional background as well as age, via Amazon's crowdsourcing platform Mechanical Turk.

Study 1a: Naming social groups. We paid 213 people (101 women, 112 men; $M = 34.41$ years, $SD = 11.02$) \$1.5 to "name 40 social groups." Importantly, we refrained from recommending sampling strategies to get at people's naive understanding of groups (for a different approach, see Fiske et al., 2002). In the upper half of the screen, people read "Dear participant, each society is not only made up by the individuals that live in the society, but these individuals also constitute what we call 'social groups.' People belong to social groups either because they have a specific characteristic that is seen as typical for a social group or because they have chosen to become part of a social group. Thus, some social groups are based on how people are, while others are based on how people behave or see the world. (These groups do not have to be mutually exclusive in the sense that being part of one social group means one cannot also be part of another social group.) Although this definition may sound very abstract to you, you probably have examples of social groups in your mind. We ask you to name 40 social groups that spontaneously come to your mind. Just think for a moment of the groups that structure society and name 40 of them." In the bottom half, people entered 40 groups into 40 text boxes.

Table 1 shows all 80 social groups named by more than 10% of people in Study 1a. Apparently people selected groups based on race or ethnicity (Whites, Blacks, Asians), social class (poor, middle-class, rich), and political or religious beliefs (Democrats, Atheists, Republicans, Christians). The combination of these 80 groups results in 3,160 possible pairs for which we collected dissimilarity judgments in Study 1b.

Study 1b: Multidimensional scaling of 80 groups. We paid 843 other people (420 women, 423 men; $M = 36.33$ years, $SD = 12.65$) \$0.6 to "rate the similarity-dissimilarity of 80 pairs of social groups." Multidimensional scaling operates on the stimulus level (here: groups); as it was not feasible to do all pairwise dissimilarity comparisons, we presented each participant with 80 randomly selected pairs of stimuli out of the full 3,160 pairs of stimuli and averaged the ratings on the stimulus level. On the first screen slide, they read "Dear participant, please rate the similarity–dissimilarity of these two social groups." Below, they used a 9-point very *similar–very dissimilar* scale to rate the two randomly selected groups. On the next screens, people rated 79 other randomly selected pairs of groups.

On average, each of the 3,160 dissimilarities was judged by $M = 20.94$ participants, $SD = 4.77$. We subjected the full matrix of 3,160 mean pairwise dissimilarities to multidimensional scaling (MDS; for a review, see Borg & Groenen, 2005). We used the ALSCAL procedure (Young, Takane, & Lewyckyj, 1978); assuming an interval scale, we estimated coordinates for the 80 social groups in dissimilarity spaces in which euclidean distances can be interpreted as dissimilarity. The further apart two groups are in these spaces, the more dissimilar people judged them to be. We estimated coordinates for six MDS solutions, varying from a one-dimensional to a six-dimensional dissimilarity space.

Table 1
Most Frequently Named Social Groups in the U.S. (Consensus > 10%) in Study 1a

1st–20th most frequent	21st–40th most frequent	41st–60th most frequent	61st–80th most frequent
Whites (66%)	Teenagers (28%)	Buddhists (19%)	Upper-class (14%)
Democrats (51%)	Muslims (27%)	Working class (19%)	Military (14%)
Blacks (48%)	Politicians (27%)	Young (19%)	Religious (14%)
Poor (47%)	Catholics (26%)	Elderly (18%)	Techies (14%)
Middle class (45%)	Gays (26%)	Hipsters (18%)	Sports fans (13%)
Asians (45%)	Men (25%)	Actors (18%)	Heterosexuals (13%)
Rich (44%)	Teachers (25%)	Homeless (17%)	Lower class (13%)
Atheists (42%)	Children (25%)	Libertarians (17%)	Drug users (12%)
Republicans (41%)	Goths (24%)	Independents (17%)	Employed (12%)
Christians (37%)	Jocks (22%)	Mexicans (17%)	Hindu (12%)
Liberals (36%)	Parents (22%)	Businesspeople (16%)	Lawyers (12%)
Conservatives (35%)	Hippies (22%)	Educated (16%)	Straight (12%)
Nerds (34%)	Doctors (21%)	White collar (16%)	Families (12%)
Students (33%)	Adults (21%)	Indians (16%)	Lesbians (12%)
Athletes (31%)	Blue collar (21%)	Old (16%)	Skaters (12%)
Jews (30%)	Geeks (21%)	Bisexuals (14%)	Stoners (12%)
Hispanics (30%)	Preps (21%)	Criminals (14%)	Agnostics (11%)
Women (30%)	Scientists (20%)	Homosexuals (14%)	Latinos (11%)
Artists (29%)	Americans (19%)	Immigrants (14%)	Rednecks (11%)
Musicians (29%)	Gamers (19%)	Unemployed (14%)	Tea Party (11%)

Note. Percentage in parentheses is proportion of participants who spontaneously named this group as a social group that is representative of the structure of U.S. society.

There are two indicators of goodness of scaling fit: scaling stress (*S*; should be preferably low) and the proportion of original dissimilarity variance accounted for by the scaling solution (R^2; should be preferably high). Table 2 shows *S* and R^2 for the six scaling solutions (1D, 2D, 3D, 4D, 5D, and 6D). Balancing goodness of scaling fit and ease of interpretation (Jaworska & Chupetlovska-Anastasova, 2009), we proceeded with the social groups' 1D, 2D, and 3D dissimilarity spaces. The scree plots of *S* and 1-R^2 showed that extracting a fourth, fifth, and sixth dimension only slightly improved *S* and R^2. Next, we inspected the corresponding scatter plots, searching for and selecting a number of candidate stereotype content dimensions deemed suitable to interpret the 1D, 2D, and 3D space. These data-driven candidates were augmented with candidate dimensions derived from the main theories of stereotype content (e.g., Fiske et al., 2002). Table osm.1 in the online supplementary material shows all 24 candidate dimensions.

Study 1c: Disambiguating the dissimilarity ratings. Finally, 620 people (275 women, 336 men and nine unassigned; $M = 34.94$ years, $SD = 12.17$) were paid $1 to "rate 80 social groups on a

Table 2
Goodness of 1D, 2D, 3D, 4D, 5D, and 6D Scaling Fit in Studies 1–6

	Method	People	Groups	1D	2D	3D	4D	5D	6D
Study 1	sequential dissimilarity judgment	U.S.	80	.23	.19	.16	**.15**	**.14**	**.13**
				.57	.73	.81	.84	.86	.88
Study 2	simultaneous dissimilarity arrangement	U.S.	80	.16	**.13**	**.11**	**.10**	**.09**	**.08**
				.62	.78	.85	.89	.91	.92
Study 3	simultaneous dissimilarity arrangement	German	76	.16	**.12**	**.10**	**.09**	**.08**	**.07**
				.75	.87	.91	.93	.95	.96
Study 4	simultaneous rating on prespecified scales	U.S.	80	**.14**	**.13**	**.12**	**.11**	**.10**	**.10**
				.72	.80	.84	.86	.88	.89
Study 5	simultaneous dissimilarity arrangement	U.S.	42 minimal.	**.12**	**.09**	**.07**	**.06**	**.04**	**.04**
				.75	.87	.93	.95	.97	.98
Study 5	simultaneous dissimilarity arrangement	U.S.	61 natural.	**.13**	**.09**	**.08**	**.06**	**.06**	**.05**
				.70	.86	.91	.94	.95	.96
Study 6	simultaneous dissimilarity arrangement	U.S.	42 minimal.	**.14**	**.10**	**.08**	**.07**	**.06**	**.05**
				.85	.93	.95	.97	.98	.98
Study 6	simultaneous dissimilarity arrangement	U.S.	61 natural.	.16	**.13**	**.10**	**.09**	**.09**	**.08**
				.71	.82	.88	.90	.92	.93

Note. Upper values indicate scaling stress (for a review, see Borg & Groenen, 2005). Lower values indicate percent of original variance retained in the scaling solution. According to Kruskal and Wish (1978), stress $<= .20$, $<= .15$, $<= .10$, $<= .05$, and $<= .025$ may be interpreted as poor, sufficient, satisfactory, good, and excellent, respectively. Bold values are stress $=<.15$, which are sufficient. In all studies except Study 1, the 2D scaling solution achieved a sufficient low stress.

stereotype dimension" (e.g., "unfriendly friendly"). On the first screen slide, they read "Dear participant, some kind of people in our society are [friendly], while other kind of people in our society are [the opposite stereotype; unfriendly]. Please rate the following 80 social groups according to how [friendly] or [unfriendly] they are." People then used 0–100 slider scales to rate the groups in a random order, one below the other on the same screen slide. There were between 22 and 27 raters per candidate stereotype content dimension. Raters' agreement about the groups was very high, $ICC(2,k) > = .84$, for all 24 candidate stereotype content dimensions (McGraw & Wong, 1996; Shrout & Fleiss, 1979).

To facilitate the interpretation of the social groups' 1D, 2D, and 3D dissimilarity space, we ran principal component analyses (PCA; Jolliffe, 2002) on the 24 candidate stereotype content dimensions, using varimax rotation. First, we determined the number of components to be extracted. The first, second, third, fourth, fifth, and sixth component explained 35%, 28%, 15%, 5%, 4%, and 3% of the total variance, respectively. Based on the scree plot, we proceeded with the extraction of three components. Aiming for simple structure, we omitted all eight candidate stereotype content dimensions that had no primary factor loading of $\geq .75$ and/or a cross-loading of $\geq .45$ on any of the three components. The eight omitted dimensions were: "incompetent–competent," "unintelligent–smart," "masculine–feminine," "communal–individualistic," "typical (in the U.S.)–unusual (in the U.S.)," "unfriendly–friendly," "intolerant–tolerant," and "unable–skillful." The third step validated the simple structure and no more omissions of candidate stereotype content dimensions were necessary.

Table 3 shows the varimax rotated component loadings of the 16 candidate stereotype content dimensions retained in this solution. Based on these component loadings, we composed the three combined candidate stereotype content dimensions *agency/socioeconomic success* (A: "powerless–powerful," "poor–wealthy," "low status–high status," "dominated–dominating," "unconfident–confident," and "unassertive–competitive;" $\alpha = .955$), *conservative–progressive beliefs* (B: "traditional–modern," "religious–science-oriented," "conventional–alternative," and "conservative–liberal;" $\alpha = .900$), and *communion* (C: "cold–warm," "untrustworthy–trustworthy," "dishonest–sincere," "repellent–likable," "threatening–benevolent," and "egoistic–altruistic;" $\alpha = .953$).[1] For short, the analysis yielded the dimensions A, B, and C. A and B were almost but not entirely orthogonal, $r = -.29$, $p < .01$; A and C were orthogonal, $r = .07$, $p = .55$; B and C were orthogonal, $r = -.01$, $p = .90$.

To compare the suitability of A, B, and C for interpreting the social groups' 1D, 2D, and 3D dissimilarity space, we carried out a series of nine multiple linear regressions with the groups' mean A, B, and C as criterion and their x-, x-/y-, and x-/y-/z-coordinates in the 1D, 2D, and 3D space as predictors, respectively (Forgas, 1976; Rosenberg et al., 1968; Shaver et al., 1987). Figuratively speaking, each of these nine property fitting analyses (ProFit; Chang & Carroll, 1969; e.g., Pattyn et al., 2013) finds out how much the location of groups in the 1D, 2D, or 3D dissimilarity spaces can be mapped onto either A, B, or C by means of rotating the dissimilarity spaces. Ideally, in the 1D group space, consisting of one axis, the groups' coordinates (consisting of scores on the single axis) correlate as high as $R(1D\ axis) = 1$ with only A, B, or C. For example, if $R(1D\ axis) = 1$ for A, but $R(1D\ axis) = 0$ for B and C, then A is maximally suitable as an axial interpretation of the groups' 1D space, accounting for 100% of the dissimilarity variance in this space. To account for 100% of the dissimilarity variance in the groups' 2D space, consisting of two axes, two *orthogonal* candidate stereotype content dimensions with $R(2D\ axis) = 1$ need to be fitted, because each axis should map onto one of the two stereotype content dimension. If so, these two can be interpreted as the two primary independent stereotype content dimensions on which people spontaneously judged the dissimilarities between the 80 social groups. This reasoning can be generalized to higher dimensions (e.g., three axes).

Table 4 shows the results.[2] The higher a multiple correlation $R(1D\ axis)$, $R(2D\ axis)$, and $R(3D\ axis)$, the more suitable was the corresponding candidate stereotype content dimension as an *axial* interpretation of the 1D, 2D, and 3D social group space, respectively. In Study 1, A, $R(2D\ axis) = .72$, $p < .001$, and B, $R(2D\ axis) = .91$, $p < .001$, were almost maximally suitable axial interpretations of the 2D group space (see also Figure 1), whereas C was not suitable as an axial interpretation of the 2D space, $R(2D\ axis) = .23$, $p = .13$.

Table 4 also shows the correlations r(1D pole), r(2D pole), and r(3D pole), the extent to which the 80 social groups' proximity to the origin of their 1D, 2D, and 3D space related to their score on A, B, and C, respectively. The closer to the origin a group is positioned, the higher that group scores on the respective dimension, resulting in a positive correlation. In the 2D space, to some extent this was the case for A and C, but not for B. Especially C was interesting, because it was not a maximally suitable axial interpretation of the 2D space. So, to some extent C was suitable as a polar interpretation of the 2D space, r(2D pole) $= .41$, $p < .001$, emerging from the two axes that represent A and B. That is, the more average a group on A and B, the more communal it was stereotyped to be; the more extreme a group on A and B, the less communal it was stereotyped to be.

Discussion

We refrained from preselecting candidate stereotype content dimensions as well as social groups because we wanted to identify the dimensions that people spontaneously use to distinguish between groups sampled without bias. The results showed that the first two of these spontaneously used stereotype content dimensions can be interpreted as *agency/socioeconomic success* (A) and *conservative–progressive beliefs* (B), because

[1] Each subdimension of our three *combined* candidate stereotype content dimensions was rated by different people. It might be argued that valid estimates of the social groups' A, B, and C requires judging the groups on all subcomponents of A, B, and C at once, as the whole is more than the sum of its parts. To address this possibility, 79 MTurkers (36 women, 43 men; $M = 32.46$ years, $SD = 10.04$) were paid $1 to rate the 80 groups on *compound* items of A, $n = 25$, $ICC(2,k) = .96$; B, $n = 27$, $ICC(2,k) = .95$; or C, $n = 27$, $ICC(2,k) = .92$. Each of the corresponding 0–100 slider scale items was anchored with a meaning cloud of all subdimensions that are included in the combined items. All compound items showed very high convergence with the combined items, $rs \geq .97$, $ps < .001$, and all analyses reported below led to identical conclusions if compound rather than combined items were used.

[2] The results of ($24 * 3 = $) 72 multiple linear regressions with the social groups' means on each of the 24 candidate stereotype content dimensions as the criterion and the groups' coordinates in their 1D, 2D, and 3D dissimilarity spaces as predictors are shown in the online supplementary material, Table osm.1, and are consistent with Table 4.

Table 3
Factor Loadings and Interpretation of the 16 Retained Dimensions in Study 1c

Candidate stereotype content dimension	1st component: Agency (A)	2nd component: Beliefs (B)	3rd component: Communion (C)
Powerless–powerful	**.940**	−.097	.102
Dominated–dominating	**.928**	−.205	−.150
Low status–high status	**.924**	−.097	.284
Poor–wealthy	**.905**	−.015	.019
Unconfident–confident	**.873**	−.174	.034
Unassertive–competitive	**.808**	.032	−.247
Traditional–modern	−.124	**.964**	.143
Religious–science-oriented	.313	**.855**	.044
Conventional–alternative	−.417	**.819**	−.234
Conservative–liberal	−.445	**.815**	.141
Untrustworthy–trustworthy	.081	−.014	**.953**
Dishonest–sincere	−.022	.025	**.936**
Repellent–likable	.226	.033	**.913**
Threatening–benevolent	.167	.103	**.910**
Cold–warm	−.178	.068	**.909**
Egoistic–altruistic	−.417	−.067	**.790**

Note. Primary factor loadings are printed in bold.

the statistical fit of A and B modeled as axes of the 2D group space was almost maximal and far better than the statistical fit of C as an axis, and A, B and C as poles at the origin of the 2D space. Unexpectedly, this data-driven A and B space (see Figure 1) was different from the warmth and competence space (Fiske et al., 2002).

Although one of the identified principal components, which we labeled agency/socioeconomic success, seemed to align somewhat with the competence dimension in the SCM (Fiske, Cuddy, & Glick, 2007), we decided against labeling it that way. Recent research suggested that stereotypic competence and stereotypic agency are distinct, and that agency is more related to socioeconomic success than to competence in the sense of ability (Carrier, Louvet, Chauvin, & Rohmer, 2014). Indeed, the items that loaded on the principal component in question (i.e., power, dominance, status, wealth, confidence, and competiveness) seemed to reflect agency better than competence. The items that reflected competence in the sense of ability (i.e., smartness, skill, and competence) either did not load strongly on this component or showed substantial cross-loadings on other components and were thus excluded. In other words, a janitor might be very smart and highly skilled, but would lack status and wealth. Conversely, a manager has high status and wealth, but might not be smart and skilled.

This alone, of course, can be an artifact of the item list we started from. However, a property fitting analysis with the 24 single items also suggested that wealth, power, dominance, and status are statistically better fitting single candidate stereotype content dimensions than smartness, skill, and competence, and that confidence and competitiveness are statistically better fitting candidate stereotype content dimensions than skill and competence (see online supplementary material, Table osm.1). Based on these results we propose agency/socioeconomic success and not competence as one of the fundamental stereotype content dimensions on which people distinguish social groups.

We labeled the second dimension conservative–progressive beliefs. Judgments of how traditional versus modern, how conven-

Table 4
Property Fitting Results for Studies 1–4

	Group sample	Stereotype content	*R(1D axis)*	*R(2D axis)*	*R(3D axis)*	r(1D pole)	r(2D pole)	r(3D pole)
Study 1	80 U.S.	Agency (A)	**.615**	**.715**	**.842**	**.419**	**.456**	**.426**
		Beliefs (B)	**.677**	**.904**	**.935**	−.116	−.060	−.037
		Communion (C)	.162	.229	**.468**	**.358**	**.406**	**.415**
Study 2	80 U.S.	Agency (A)	**.766**	**.812**	**.898**	.069	.183	.234
		Beliefs (B)	**.720**	**.812**	**.954**	−.035	.076	.017
		Communion (C)	.071	.175	.236	**.562**	**.580**	**.581**
Study 3	76 German	Agency (A)	**.821**	**.903**	**.909**	.283	.272	.334
		Beliefs (B)	.257	**.857**	**.831**	**−.372**	−.187	−.174
		Communion (C)	**.502**	**.464**	**.479**	**.783**	**.745**	**.727**
Study 4	80 U.S.	Agency (A)	**.856**	**.893**	**.893**	.221	.208	.222
		Beliefs (B)	.339	**.848**	**.890**	.015	.009	.010
		Communion (C)	.205	.375	**.622**	**.482**	**.506**	**.477**

Note. *R(1D–3D axis)* indicate the maximal correlations between the 80/76 U.S./German social groups' agency/socioeconomic success, conservative–progressive beliefs, and communion ratings and their projections on an axis rotated around the origin of their 1D, 2D, and 3D dissimilarity space in Studies 1–4; *r*(1D–3D pole) indicate correlations between the groups' A, B, and C ratings and their proximity to the origin of these spaces. Bold correlations are significant at $p = < .001$.

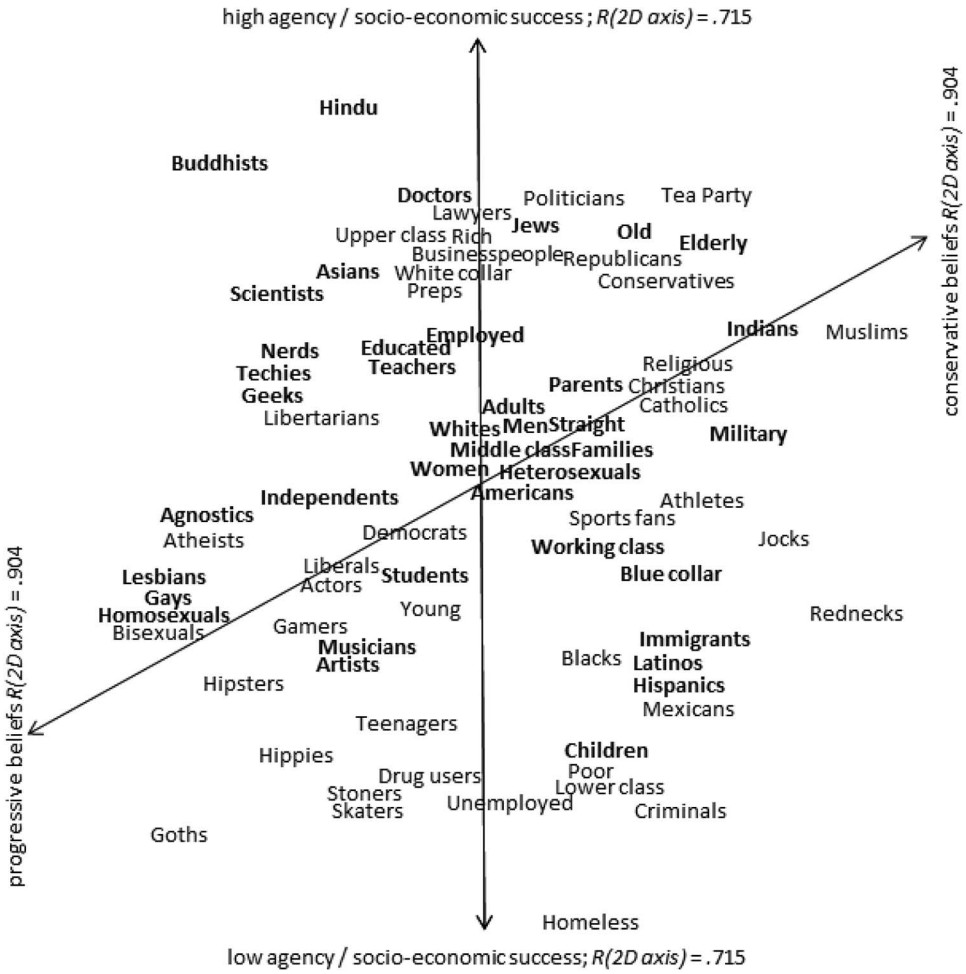

Figure 1. Study 1 (U.S. participants and target social groups): This 2D space of 80 representatively sampled social groups was computed based on pairwise dissimilarity ratings, and can be interpreted by agency/socioeconomic success and conservative–progressive beliefs. Communion emerges within these two dimensions. Groups that are average on A and B are perceived as more communal (the 40 most communal social groups are bold), whereas groups that are extreme on A and B are perceived as less communal (the 40 least communal social groups are not bold).

tional versus alternative, how conservative versus liberal, and how religious versus science-oriented the groups are loaded high on this dimension; we thus concluded that it captures socially shared convictions about groups' conservative–progressive beliefs. The discovery of this dimension underlines the usefulness and necessity of data-driven approaches because few theories have previously addressed what a group believes in as relevant for stereotyping. Participants seem to systematically differentiate groups on the basis of them either striving to keep up traditions/preserving the status quo (e.g., conservatives, religious, Republicans) or striving to overcome traditions/altering the status quo (e.g., gays, atheists, liberals). In a sense, much like warmth in the stereotype content model is conceptualized as informative of mainstream society's views about a group's intention to help/care versus harm/neglect, conservative–progressive beliefs are informative of mainstream society's views about a group's intention to preserve versus change the status quo.

This finding is in line with Jones and Ashmore (1973) and Pattyn et al. (2013), who found a similar dimensions (modern–backward and alternative–conventional, respectively) using a theory-driven selection of stimuli (e.g., an image of a punk and an elderly person). Our findings are thus the first to establish the centrality of this dimension for distinguishing between representatively sampled social groups.

The dimension of conservative–progressive beliefs is also compatible with fundamental dimensions from other areas of psychology. On the level of personality traits, the Big 5 factor openness to experience (McCrae & Costa, 1987; McCrae & John, 1992) taps into a similar construct, and this personality trait has been identified as one of the central predictors of political conservatism (Jost, Glaser, Kruglanski, & Sulloway, 2003). Moreover, one of the two central dimensions on which human values can be positioned is openness to change (self-direction, stimulation) versus conservation (security, conformity, tradition; Schwartz, 1994; Schwartz &

Bilsky, 1987, 1990). In U.S. society this dimension has received increasing attention over the recent years, leading some scholars to speak of a divide or even polarization between liberal and conservative camps (Brewer, 2005; Haidt, 2012).

Of the three combined candidate stereotype content dimensions that we composed, communion (Abele & Wojciszke, 2014; highly akin to the dimension warmth in the SCM; Fiske et al., 2007) did not appear to be one of the two stereotype content dimensions that participants most often used to judge the dissimilarities between the 80 social groups. Importantly, this was not due to the fact that the groups' C (trustworthiness, sincerity, warmth, benevolence, likability, and altruism) ratings were unreliable. In fact, the reliability was very high for all subdimensions of stereotypic C, $ICC(2,k) > = .84$ (McGraw & Wong, 1996; Shrout & Fleiss, 1979). This contrasts with the pivotal role of communion/warmth in existing theories of stereotype content (Cuddy et al., 2007; Fiske et al., 2002), and also in theories of social perception in general (Abele & Wojciszke, 2014).

Despite this lack of support for C as one of the first two spontaneously employed stereotype content dimensions, we found support for C as emerging from the first two spontaneously employed stereotype content dimensions. Specifically, groups positioned closer to the origin of the 2D A and B space were judged as relatively more communal. Therefore, highly communal groups were seen as neither too rich, nor too poor, as well as neither too conservative, nor too progressive. Less communal groups were peripheral (see groups marked in blue in Figure 1), whereas more communal groups were central (see groups marked in red in Figure 1). This finding reconciles our 2D solution with existing models that consider communal attributes to be fundamental to stereotype content: Study 1 suggests that communion is encoded by the two stereotype content dimensions that we refer to as A and B in a nonlinear way.

In sum, Study 1 suggests that fundamental stereotype content about social groups can be described by a 2D space spanned by A and B from which communion emerges as a function of centrality within that space (see Figure 1). At this point, we cannot be certain that this 2D ABC (agency/socioeconomic success, conservative–progressive beliefs, and communion) stereotype content model provides a full description of the dimensions that people spontaneously used for judging the dissimilarities between the representatively sampled groups. Most problematically, the goodness of scaling fit of the groups' 3D coordinates was more than slightly better than the scaling fit of the groups' 2D coordinates, suggesting that the 2D ABC model misses a third spontaneously used stereotype content dimension. Based on that C was a moderately suitable axial interpretation of the social groups' 3D space, this third dimension could be C (note that C can at the same time be a third independent dimension and emerge from centrality on the first two). Further, according to Kruskal and Wish (1978), a scaling fit of $S < = .20$, $< = .15$, $< = .10$, $< = .05$, and $< = .025$ is poor, sufficient, satisfactory, good, and excellent, respectively. Using these criteria we have to concede that neither the 2D nor the 3D solution showed a sufficient fit. Although the 4D space met this standard ($S = .15$), the improvement was only marginal compared with the 3D solution. These findings suggest two aspects. First, people based their group dissimilarity judgments primarily but not solely on A and B. Second, the dimensions they employed additionally are not consensually shared to the degree that they form more than one orthogonal dimension that explains a noteworthy increase in explained dissimilarity variance.

To a certain degree, this was a consequence—and an advantage—of our design. Each participant judged the dissimilarity of only ~ 2.5% of all unique pairs of social groups, and thus each participant judged dissimilarity in a highly different context. This might have added additional noise to the data because dissimilarity (i.e., the way it is construed) varies as a function of context of judgment (Goldstone, Medin, & Halberstadt, 1997; Krumhansl, 1978; Tversky, 1977). The sequential, pairwise mode of dissimilarity judgment also could have encouraged people to switch between many circumstantial stereotype content dimensions rather than to stick with the essential ones. The advantage lies in that the context of judgment is so variable across participants that it could not have constrained the outcome of judgment to any dimension. The fact that we nevertheless obtained at least two meaningful dimensions speaks to the centrality of these.

Another factor that might have contributed to the nonoptimal scaling fit might lie in the repetitive nature of making 80 pairwise dissimilarity judgments sequentially. It is conceivable that the repetitive nature of the task tempted our online participants to pay increasingly less attention and thus added noise to the data. In our next study, we aimed to ameliorate these problems by employing a more stimulating research design in which participants judged the dissimilarities between large arrays of social groups simultaneously. Such an alternative to the classic pairwise method has recently been proposed as the *spatial arrangement method* (SpAM; Hout et al., 2013).

Study 2

Consistent with the geometric model of similarity (Carroll & Wish, 1974; Nosofsky, 1992; Torgerson, 1965), the spatial arrangement method (SpAM; Hout et al., 2013; see also Goldstone, 1994; Koch, Alves, Krüger, & Unkelbach, in press; Kriegeskorte & Mur, 2012) rests on the assumption that people can reliably and validly sort attitude objects in a way that more dissimilar attitude objects are located further apart. To illustrate, Goldstone (1994) presented participants with multiple variants of the letter A (in different font styles) all at once and in random locations on the computer screen. Their task was to use the computer mouse to "move the letters around so that letters that are similar to each other are close. The more similar two letters are, the closer they should be" (Goldstone, 1994, p. 382). The distances between the spatially arranged letters correlated strongly with sequential, pairwise dissimilarity judgments collected from a different sample of people (see also Hout et al., 2013; Koch et al., 2016). Thus, sequential, pairwise judgment and SpAM seem to be equally effective ways to measure interstimulus dissimilarity. The advantage of SpAM is that it is a lot more efficient, because the dragging-and-dropping of a single attitude object simultaneously adjusts the distances between that attitude object and all other attitude objects on the dissimilarity map. In fact, with the help of SpAM, people are able to assess the entire pattern of dissimilarities between dozens of attitude objects in a quick, easy, and readjustable way, because all attitude objects can be moved to a different location on the dissimilarity map at all times during the task. Thus, SpAM is ideal to improve on the design of Study 1. Based on the results of Study 1, we hypothesized that people would spontane-

ously use the stereotype content dimensions A and B to spatially arrange the dissimilarities between the 80 social groups. We expected C to—again—emerge as a function of centrality in the 2D space spanned by the other two dimensions. Together, these findings would further support our 2D ABC model of stereotype content.

Method

Participants and stimuli. We paid 131 MTurkers (67 women, 64 men; $M = 34.74$ years, $SD = 11.84$) \$1 to "sort 40 social groups on the computer screen." They received a random sample of 40 out of the 80 groups chosen as representative of U.S. society in Study 1.

Procedure. On the first screen slide, participants read "Dear participant, your task is to sort 40 social groups based on how similar / dissimilar they are. The social groups will appear in the middle of the screen one at a time, and you can drag-and-drop them at any time to change their location on the screen. Please sort the social groups in such a way that more similar social groups are more close to each other, while more dissimilar social groups are more distant to each other. That is, please use the social groups to draw a map in which greater proximity means greater similarity, and in which greater distance means greater dissimilarity". After clicking on an "I understand" button, the button disappeared, and a randomly selected group appeared in the middle of the screen. Once that group was dragged to another location on the screen, the button reappeared as a "Next social group" button in the center of the screen bottom, and with a click on the button the next randomly selected group appeared in the middle of the screen, and the button disappeared again. This procedure was repeated until all 40 groups were positioned on the screen. After the 40th group was arranged on the dissimilarity map, the button changed to "I finished" (see online supplementary material, Figure osm.1). Upon clicking this button, the dissimilarity distances between the groups were recorded as proportions of the greatest possible distance—the screen diagonal.

Results

First, we computed the mean distance—that is, the mean dissimilarity—for each of the 3,160 unique pairs of social groups across all people who had dragged-and-dropped that pair ($M = 32.19$, $SD = 4.94$). Next, we subjected the mean dissimilarities to MDS (with the same parameter values as in Study 1. The goodness of fit of the 1D, 2D, 3D, 4D, 5D, and 6D scaling solutions are shown in Table 2. Balancing goodness of scaling and ease of interpretation, as in Study 1, we proceeded with the 1D, 2D, and 3D group spaces.

To compare the suitability of the social groups' stereotypic A, B, and C for interpreting the groups' 1D, 2D, and 3D space, as in Study 1, we again carried out nine multiple linear regressions with the groups' means on A, B, and C from Study 1 as the criterion and the groups' coordinates in the 1D, 2D, and 3D space as predictors, respectively. The results are shown in Table 4.[3] As in Study 1, A, $R(2D\ axis) = .81$, $p < .001$, and B, $R(2D\ axis) = .81$, $p < .001$, were the most and almost maximally suitable axial interpretations of the 2D space (see Figure 2), whereas C was not a suitable axial interpretation of the 2D space, $R(2D\ axis) = .18$, $p = .30$. Table 4 also shows the linear relations between the groups' A, B, and C and the groups' proximity to the origin of their 1D, 2D, and 3D space. As in Study 1, C was a suitable polar interpretation of the 2D space, $r(2D\ pole) = .58$, $p < .001$, whereas A and B were not. The same pattern of results was found for the groups' 3D space.

Discussion

In Study 2, participants spatially arranged sequentially appearing social groups on the 2D screen (more dissimilar groups had to be positioned further apart). Our results show that given two dimensions to distinguish between the groups, people used A and B, but not C; as in Study 1, C again emerged as a function of centrality within the stereotype content space spanned by A and B (see Figure 2). Thus, our results provided further support for the 2D ABC model identified in Study 1. First, the suitability of A and B as axial interpretations of the 2D group space was almost maximal, while C was not a suitable axial interpretation of the 2D space. Second, the suitability of C as a polar interpretation of the 2D space was substantial and higher than in Study 1, while A and B were not suitable as polar interpretations of the 2D space. Third, the suitability of A and B as axial interpretations of the 2D space was higher than the suitability of C as a polar interpretation of the 2D space. And fourth, the scaling fit of the groups' 2D coordinates was sufficient ($S <= .15$; see Table 2).

As in Study 1, the scaling fit of the social groups' 3D coordinates was better than the scaling fit of the social groups' 2D coordinates, suggesting that the 2D ABC model missed a third independent stereotype content dimension (in the 2D ABC model, C is not an independent dimension, because it emerges as a function of centrality within the stereotype content space spanned by A and B). However, in contrast to Study 1, Study 2 showed no evidence that this third independent dimension might be described as C. Based on our data, we could not adequately interpret the third independent dimension (if there is any). Therefore, the more parsimonious 2D ABC model was the best available interpretation of the stereotype content dimensions that people spontaneously used to spatially arrange the dissimilarities between the 80 representative U.S. groups. In other words, Study 2 confirmed that the two most fundamental dimensions of stereotype content can be described as A and B, and that C can be described to emerge from A and B, and not vice versa. Specifically, the more average a social group on A and B, the more communal it was stereotyped to be; in contrast, the more extreme a social group on A and B, the less communal it was stereotyped to be.

Both Study 1 and 2 recruited Amazon.com Mechanical Turk workers as participants, because their demographics have repeatedly been shown to be relatively more population-representative than the demographics of other convenience samples such as university students (Berinsky, Huber, & Lenz, 2012; Buhrmester, Kwang, & Gosling, 2011; Casler, Bickel, & Hackett, 2013; Mason & Suri, 2012; Paolacci, Chandler, & Ipeirotis, 2010). Nevertheless, one could argue that the results might be specific to our participant sample. Despite their greater representativeness in terms of age, education, and income, MTurkers might constitute a biased sample

[3] The results of separate analyses for the 24 candidate stereotype dimensions are shown in the online supplementary material, Table osm.2, and are consistent with Table 4.

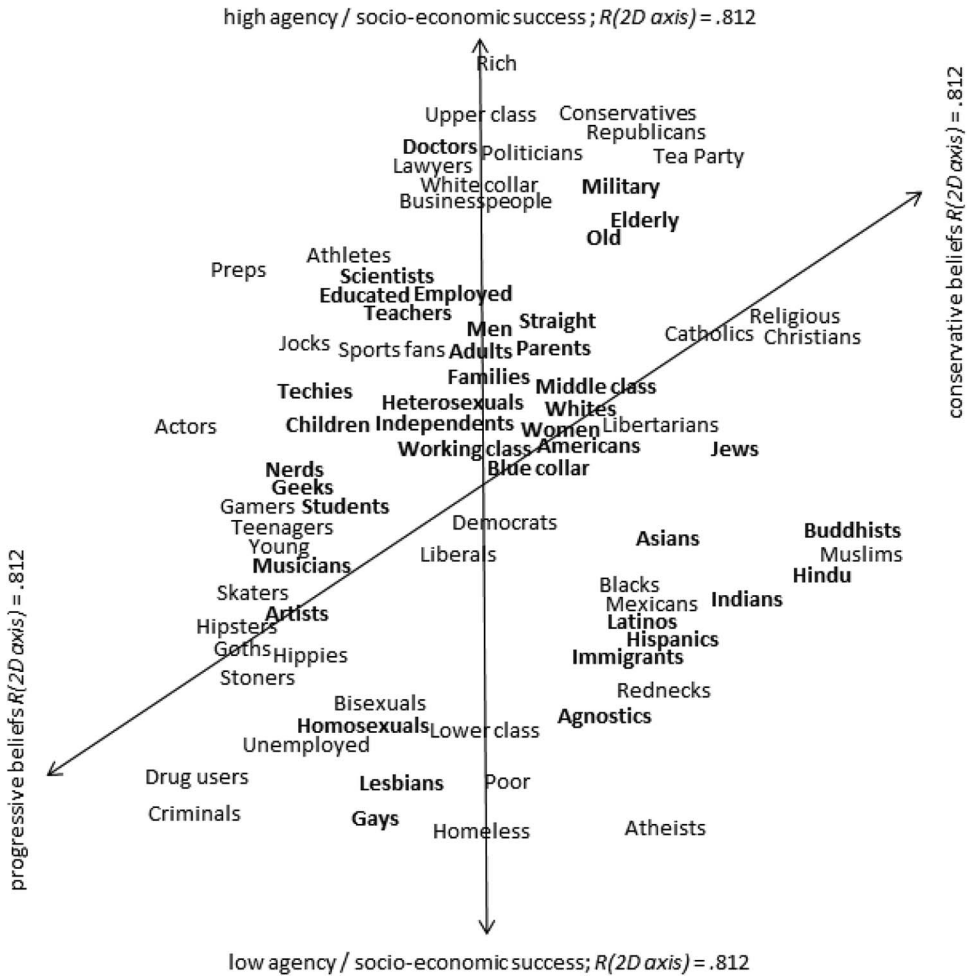

Figure 2. Study 2 (U.S. participants and target social groups): This 2D space of 80 representatively sampled social groups was computed based on spatially arranged dissimilarity distances, and can be interpreted by agency/socioeconomic success and conservative–progressive beliefs. Communion emerges within these two dimensions. Groups that are average on A and B are perceived as more communal (the 40 most communal social groups are bold), whereas groups that are extreme on A and B are perceived as less communal (the 40 least communal social groups are not bold).

in terms of other variables like affinity with computers. Even more relevant, the population of the U.S. is not representative of other nations. The strong topicality of the divide between two political camps in the U.S. (Brewer, 2005; Layman & Carsey, 2002; McCarty, Poole, & Rosenthal, 2008) might have increased the salience and accessibility of the conservative–progressive beliefs dimension. We thus sought to bolster the generalizability of our 2D ABC model by replicating it in a different culture.

Study 3

Study 3 replicated Study 2 with German rather than U.S. American participants.

Method and Results

Study 3a: Naming social groups in Germany. We collected data from 178 online participants contacted through an e-mail list of individuals interested in participating in studies at the University of Cologne (119 women, 53 men; $M = 26.35$ years, $SD = 6.11$). They were offered a chance to win one of five vouchers (€20) for a large online retailer. In the top half of the first screen slide, they read the same instructions (in German) as the people who named social groups in Study 1. In the bottom half, they entered 40 social groups into 40 text boxes. Table 5 shows all 76 social groups named by more than 10% of all people.

Study 3b: Dissimilarity arrangement of 76 groups. Another 69 students were recruited on the campus of the University of Cologne (47 women, 22 men; $M = 23.37$ years, $SD = 4.53$) to participate in a lab study for a small monetary compensation (€2). Their instructions were the same as in Study 2 (in German), namely to spatially arrange a random sample of 50 of the 76 social groups. More similar groups had to be placed more close to one another, and more dissimilar social groups had to be placed further apart from one another.

Table 5
Most Frequently Named Social Groups (Consensus ≥ 10%) in Germany in Study 3a

1st–20th Most frequent	21st–40th Most frequent	41st–60th Most frequent	61st–76th Most frequent
Students (70%)	Christians (27%)	Adults (20%)	Germans (14%)
Children (58%)	Foreigners (27%)	Drug addicts (20%)	Goths (13%)
Employed (56%)	Religious (27%)	Catholics (19%)	Alcoholics (13%)
Unemployed (56%)	Academics (26%)	Conservatives (19%)	Single parents (13%)
Young (47%)	Homosexuals (26%)	Self-employed (18%)	Rightists (12%)
Pupils (46%)	Musicians (26%)	Welfare recipients (18%) (17%)	Sick (12%)
Pensioners (44%)	Jews (24%)	Criminals (18%)	Nazis (12%)
Muslims (38%)	Trainees (24%)	Lower class (16%)	Blue collar (11%)
Officials (37%)	Parents (23%)	Upper class (16%)	Hip-Hopper (11%)
Workers (36%)	Vegans (22%)	Leftists (16%)	Emos (11%)
Athletes (34%)	Hipsters (22%)	Rural (16%)	Scientists (11%)
Politicians (33%)	Singles (22%)	Economic-liberals (16%)	Right-wing extremists (11%)
Migrants (33%)	Teachers (21%)	Employers (16%)	Rockers (11%)
Artists (31%)	Atheists (21%)	Car drivers (15%)	Managers (11%)
Middle class (31%)	Vegetarians (20%)	Nerds (15%)	Bicycle drivers (10%)
Punks (30%)	Poor (20%)	Educated (15%)	Soccer players (10%)
Elderly (30%)	Urban (20%)	Buddhists (15%)	
Disabled (29%)	Doctors (20%)	Hippies(15%)	
Rich (29%)	Heterosexuals (20%)	Environmentalists (15%)	
Homeless (28%)	Families (20%)	Celebrities (14%)	

Note. Percentage in parentheses is proportion of participants who spontaneously named this group as a social group that is representative of the structure of German society.

As in the previous spatial arrangement study we subjected the mean distances for each of the 2,850 unique pairs of social groups (average number of raters per pair $M = 29.23$, $SD = 5.15$) to MDS (same settings as in Studies 1 and 2). The goodness of fit of the 1D, 2D, 3D, 4D, 5D, and 6D scaling solutions are shown in Table 2. Balancing goodness of scaling fit and ease of interpretation, we extracted, analyzed, and interpreted only the 1D and 2D social group spaces, because the scree plots of S and 1-R^2 showed that extracting a third and higher dimensions did only to a slight degree improve S and R^2. Nevertheless, to better compare Study 3 with Studies 1 and 2, we extracted and analyzed a third dimension, but refrained from interpreting the respective 3D group space.

Study 3c: Disambiguating the dissimilarity arrangement. Finally, 60 other participants recruited on the campus of the University of Cologne (41 women, 19 men; $M = 22.55$ years, $SD = 4.55$) received a piece of candy to participate in a lab-study to rate the 76 social groups on compound A, B, or C. Twenty participants rated all groups on a slider scale ranging from 1 ("low power/low status/low dominance/low confidence") to 100 ("high power/high status/high dominance/high confidence"), measuring stereotypic A, $ICC(2, k) = .965$, 20 different participants rated all social groups' B on an identical scale with the anchors "traditional/religious/conservative/conventional—modern/faithless/liberal/alternative;" $ICC(2, k) = .912$, and a third group rated all social groups' C ("low trustworthiness/low sincerity/low benevolence/low likability—high trustworthiness/high sincerity/high benevolence/high likability"), $ICC(2, k) = .952$. A and C were correlated, $r = .32$, $p < .01$; A and B were orthogonal, $r = -.10$, $p = .41$; and B and C were orthogonal, $r = .08$, $p = .49$.

Employing the same property fitting strategy as in the previous studies suggested that A, $R(2D\ axis) = .90$, $p < .001$, and B, $R(2D\ axis) = .86$, $p < .001$, were far better axial interpretations of the 2D social group space than C, $R(2D\ axis) = .38$, $p < .001$ (see Table 4 and Figure 3). As in Study 2, C was a suitable polar interpretation of the 2D space, $r(2D\ pole) = .75$, $p < .001$, whereas A and B were not. The same pattern of results was found for the groups' 3D space.

Discussion

Study 3 was set in another national context (Germany), and we used another type of sample (mostly university students) and another research setting (Studies 3b and 3c were conducted in the lab rather than online). The scaling fit of the social groups' 2D coordinates was again sufficient ($S <= .15$) and did not markedly improve with the addition of another dissimilarity dimension. Thus, modeling a third independent stereotype content dimension was nonessential. Overall, Study 3 supported the 2D ABC stereotype content model identified in Studies 1 and 2 (see Figure 3), and also refined it, potentially due to reduced noise in the data. Specifically, A was found to be more fundamental than B, as in the groups' 1D space (already accounting for 75% of the original spatially arranged dissimilarity variance) A was a suitable axial interpretation, whereas B was not. B only accompanied A as a suitable axial interpretation in the 2D space. Because A and B were more suitable axial interpretations of the 2D space compared to the suitability of C as a polar interpretation of the 2D space, the primary and secondary fundamental stereotype content dimensions should be interpreted as A and B, respectively. Further, the suitability of C as a polar interpretation was already maximal in the 1D space, which, in contrast to Studies 1 and 2, suggests that C is primarily a function of not being too high or too low on A (i.e., C can be inferred from A, but not from B; see blue and red social groups in Figure 3).

Study 3 thus refined the conclusions drawn from Studies 1 and 2 and at the same time strengthened the empirical base of the

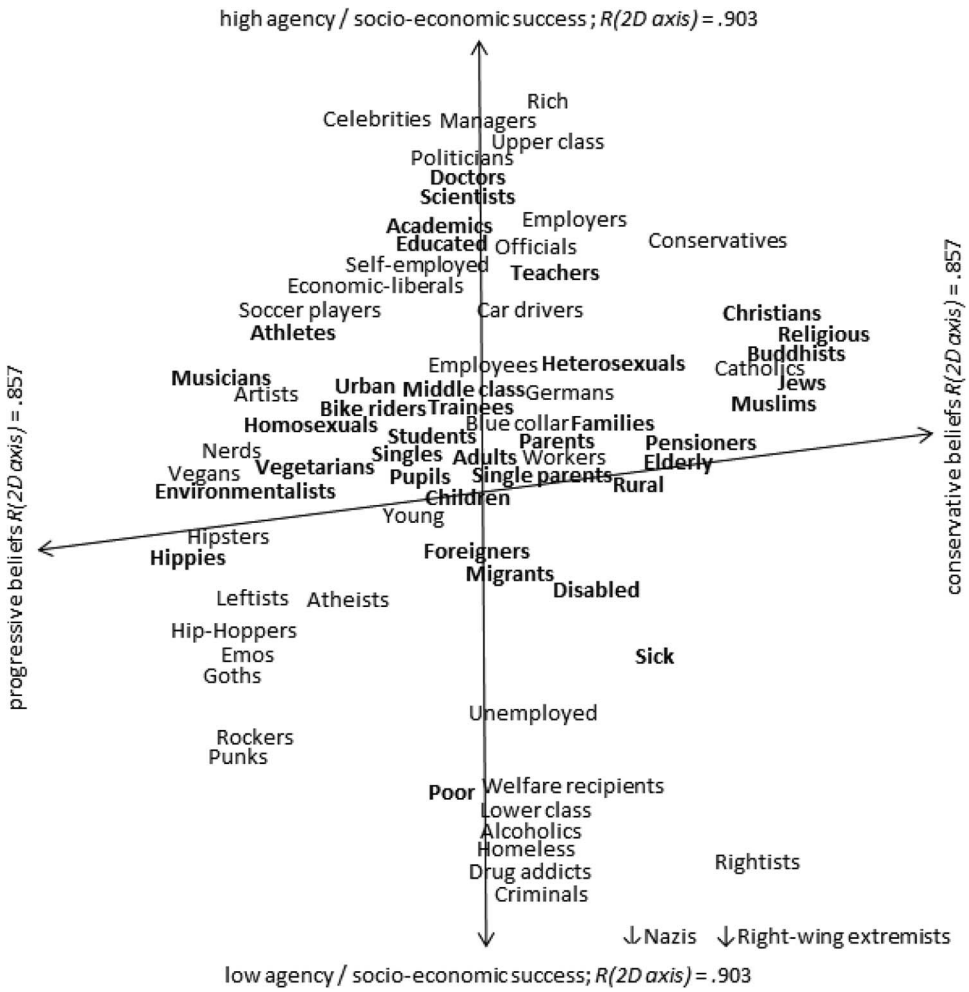

Figure 3. Study 3 (German participants and target social groups): This 2D space of 76 representatively sampled social groups was computed based on spatially arranged dissimilarity distances, and can be interpreted by agency/socioeconomic success and conservative–progressive beliefs. Communion emerges within these two dimensions. Groups that are average on A and B are perceived as more communal (the 40 most communal social groups are bold), whereas social groups that are extreme on A and B are perceived as less communal (the 40 least communal social groups are not bold).

proposed 2D ABC model of stereotype content. Our results became increasingly clear from Study 1 to Studies 2 and 3, speaking to the usefulness of employing the spatial arrangement method. In particular, if participants are confined to a 2D space, they use A and B to organize the social groups, while C emerges as a nonlinear function of A but not B.

Verticality, Horizontality, and Centrality Metaphors as Possible Alternative Explanations

Despite this, there may also be pitfalls in using this method. It is conceivable that semantic concepts are intrinsically associated with spatial locations. One example is that the concept of power is commonly found to be intuitively represented vertically with high power on the top and low power on the bottom (Schubert, 2005; see also Meier & Robinson, 2004; Slepian, Masicampo, & Ambady, 2015). Specifically, words representing high versus low power (e.g., employer vs. employee) can be more rapidly identified as connoting high versus low power when they are presented in the metaphorically corresponding area of a screen (top for high power, bottom for low power; Schubert, 2005). Likewise, people presented at the top of the screen are perceived to be more powerful (Giessner & Schubert, 2007; Meier, Hauser, Robinson, Friesen, & Schjeldahl, 2007). Thus, the vertical nature of the spatial arrangement board could have primed people to construe the dissimilarities between the social groups with respect to their stereotypic agency/socioeconomic success.

A similar argument could be made for the belief dimension. Going back to the seating arrangement in legislative bodies during the French revolution era, progressives are often referred to as left-wing, whereas conservative beliefs are referred to as right-wing. Importantly, this is not only an abstract reference but hori-

zontal positions on the left or the right are intuitively connected with the corresponding political attitudes (Farias, Garrido, & Semin, 2013; Oppenheimer & Trail, 2010; van Elk, van Schie, & Bekkering, 2010). The horizontal nature of the spatial arrangement board could thus have primed people to construe the dissimilarities between the social groups with respect to their conservative–progressive beliefs and sort the groups accordingly.

Finally, even the position of high communal groups in the center and low communal groups near the margin of the screen could be construed as merely reflecting an instance of embodied semantics or metaphors. The center often stands for relevant ingroups, whereas outgroups are labeled as (in this case literally) peripheral. Moreover, motivated explanations are conceivable. Pushing unpleasant (i.e., unfriendly and cold) groups to the margin of the screen and keeping the friendly and warm groups in the center of frequent attention could be experienced as more pleasant than the opposite. Alternatively, being motivated to keep the most frequent, typical, and familiar groups in the center of frequent attention could explain why high and low communal groups were positioned in the center and at the margins, respectively, because "what is typical is good" (Sofer, Dotsch, Wigboldus, & Todorov, 2015).

To test whether these confounds are, at least in principal, a problem, we correlated the social groups' spatially arranged average verticality (higher values: more upward), horizontality (higher values: more rightward), and centrality (higher values: more inward) with the social groups' A, B, and C ratings, respectively. In line with spatial metaphors, more agentic social groups were positioned further upward, $r_{Study\ 2} = .74$, $p < .001$, $r_{Study\ 3} = .60$, $p < .001$; and, more communal social groups were positioned further inward, $r_{Study\ 2} = .52$, $p < .001$, $r_{Study\ 3} = .69$, $p < .001$. Contradicting a spatial metaphor, more liberal (conservative) social groups were not positioned further leftward (rightward), $r_{Study\ 2} = .03$, $p = .78$, $r_{Study\ 3} = -.23$, $p = .04$.

Although the social groups' spatially arranged coordinates cannot explain our empirical support for the B dimension, and although the spatial A, B, and C metaphors account cannot explain the highly convergent findings of Study 1 that did not use a spatial arrangement task, in Study 4 we sought to rule out spatial metaphors of A, B, and C and motivational forces as alternative explanations for the results obtained in Studies 2 and 3.

Another caveat that we sought to address in Study 4 refers to how participants decided on the criterion to estimate the dissimilarities between the social groups. Participants relied on group differences that map well on the groups' A and B. This can be interpreted as evidence that A and B are the two fundamental and thus most important stereotype content dimensions. It is also possible that the employed dimensions are not the most important but merely the most convenient, either because they are most accessible (which could be seen as an indirect indicator of importance), or because their metaphorical spatial representation conveniently maps on the spatial nature of the arrangement task. We sought to address this caveat by introducing a more explicit judgment of dimension importance. Specifically, in Study 4, we explicitly asked participants to label the first and second most fundamental stereotype content dimensions prior to the spatial arrangement task. This also helped addressing the issue of order of importance.

Study 4

Study 4 presented people with a sample of 40 groups, and asked them to specify the two person characteristics (i.e., stereotype content dimensions) that they thought best capture the dissimilarities between the social groups. Asking them to prioritize between the two allowed us to get an empirical hold on which dimension is seen as more primary than the other. Then, people judged the social groups precisely on the two dimensions that they had just selected as the most important ways to stereotypically compare the groups by spatially arranging them. Importantly, the first named dimension always had to be arranged on the horizontal dimension of the screen, thereby undermining spontaneous mapping on metaphorically corresponding dimensions (assuming that A would be most primary and thus mapped horizontally instead of vertically as in the previous studies). Based on the 2D ABC model, we expected people to specify stereotype content dimensions with a high relation to both A and B, and with a low relation to C.

Method

Participants and stimuli. We paid 66 MTurkers (31 women, 35 men; $M = 32.86$ years, $SD = 10.81$) \$1 to "rate 40 social groups on 2 dimensions of your choice." People spatially arranged a random sample of 40 of the 80 social groups that are representative of the structure of the U.S. society (see Study 1).

Procedure. In the middle of the first screen slide, people were presented with a table that showed 40 social groups in 10 rows and four columns. At the top, they read: "Dear participant, please name the person characteristic that best describes the differences and similarities between these 40 social groups. Ideally, you should be able to divide the 40 social groups into 10+ low scorers, 10+ average scorers and 10+ high scorers on this characteristic. Please enter this characteristic in the text box below [this instruction]." At the bottom, they read: "Now, please name another person characteristic that well describes the differences and similarities between these 40 social groups. Again, you should be able to divide the 40 social groups into 10+ low scorers, 10+ average scorers and 10+ high scorers on this other characteristic. Please enter this other (= not the same as above!) characteristic in the text box below [this instruction]." On the second screen slide, participants read: "Dear participant, your next task is to position the 40 social groups on the computer screen. More specifically, on the next slide the 40 social groups will appear in the middle of the screen, and your task is to drag-and-drop each social group to a different position on the screen. Please make use of the entire screen and position the social groups as follows. For the first person characteristic that you specified (→ [whatever they had typed in first]): position the low scorers on the left of the screen, position the average scorers in between the left and the right, and position the high scorers on the right of the screen. For the second person characteristic that you specified (→ [whatever they had typed in second]): position the low scorers at the bottom of the screen, position the average scorers in between the bottom and the top, and position the high scorers at the top of the screen. In sum, your task is to plot the 40 groups according to how they differ on [whatever they had typed in first] and [whatever they had typed in second]. If you want to exchange these person characteristics, you may do so at any time during the positioning task by rephrasing the text boxes at the screen edges." On the third slide the 40 groups appeared in 40

adjacent labels in the screen middle. The appearance of the labels and the full screen background was the same as in Studies 2 and 3, except for a horizontal axis labeled in accordance with the characteristic that the participant had specified first, and a vertical axis labeled in accordance with the characteristic that the participant had specified second (see online supplementary material, Figure osm.2). The labeling of the axes ("low scorers on . . ." and "high scorers on . . .") could be changed at any time during the rating phase. Once all groups were dragged-and-dropped to a different position on the screen, an "I finished" button appeared. Upon clicking this button, the computer program rescaled the groups' positions to a quadratic 2D space, and then, as in Studies 2 and 3, the distances between the groups were recorded as a proportion of the greatest possible distance—the diagonal of this 2D space.

Results

We first computed the mean rating distance—that is, the mean dissimilarity—for each of the 3,610 unique pairs of social groups across all people who had rated that pair ($M = 16.02$, $SD = 3.83$). These mean dissimilarities were subjected to MDS with the same settings as in Studies 1–3. The goodness of fit of the 1D, 2D, 3D, 4D, 5D, and 6D scaling solutions are shown in Table 2. As in Study 3, to ease comparing Study 4 to the other studies, we extracted and analyzed three dimensions, but interpreted only the 1D and 2D group spaces, as the scree plots of S and 1-R^2 showed that extracting a third or higher dimensions did only slightly improve S and R^2 (i.e., modeling a third independent stereotype content dimension was nonessential.). Extracting a second dimension did not substantially improve S, but substantially improved R^2, and thus we interpreted the 1D and 2D social group space.

Identical property fitting analyses as in the previous studies were carried out and replicated the findings of Study 3 that A, *R(2D axis)* = .89, $p < .001$, and B, *R(2D axis)* = .85, $p < .001$ (see Table 4),[4] were far better axial interpretations of the 2D social group space compared to C, *R(2D axis)* = .38, $p = .003$. Replicating Studies 1–3, C was a substantially suitable polar interpretation of this 2D space, r(2D pole) = .51, $p < .001$—A and B were not.

More importantly, we sought to address whether the spatial nature of our task might have prompted participants to employ metaphorically corresponding stereotype content dimensions (A for the vertical dimension, B for the horizontal dimension, C for centrality). As participants had to specify two stereotype content dimensions before learning about the spatial nature of the rating task, the task could only have influenced participants' choice of dimensions if participants changed the two dimensions once realizing that they had to spatially arrange social groups. Tracking such changes in the self-selected dimension labels revealed that this happened for only very few cases (around 7%).

To categorize the self-selected labels according to unambiguous fit with our combined candidate stereotype content dimensions, 40 additional MTurkers (18 women, 22 men; $M = 33.10$ years, $SD = 8.45$) were paid \$0.5 to "assign 66 person characteristics" one after the other, and they read: "Please select the category to which this person characteristic [e.g., wealthy] fits best. If this person characteristic does not fit well to any of the categories, check 'no match.'" The seven categories available for selection were 'no match' plus the polar opposites on A, B and C (A+: "powerful / dominating / high status / wealthy / confident / competitive," A-: "powerless/dominated/low status/poor/unconfident/unassertive," B+: "modern/science-oriented/alternative/liberal," B-: "traditional/religious/conventional/conservative," C+: "trustworthy/sincere/likable/benevolent/warm/altruistic," and C-: "untrustworthy/dishonest/repellent/threatening/cold/egoistic").

Participants assigned either the 66 labels chosen by the original participants as the "best" description of the differences and similarities between the groups (i.e., the first dimension chosen, $N = 22$), or the 66 labels chosen by the original participants as their second dimension ($N = 18$). For each of the 132 labels, we averaged percentage of assignment to categories A ±, B ±, C ±, and "no match," a measure of the labels' relatedness to A, B, C, and something else, respectively. Consistent with the property fitting analyses reported in Studies 1–4, the labels related to A ($M = 28.35\%$, $SD = 31.21$) and B ($M = 31.15\%$, $SD = 30.83$) to an equal extent, $F(1, 131) = 0.37$, $p = .54$, $\eta_p^2 = .00$, 90% CI [.00, .04], related at least by trend more to A than to C ($M = 21.03\%$, $SD = 22.65$), $F(1, 131) = 3.69$, $p = .057$, $\eta_p^2 = .03$, 90% CI [.00, .09], and related more to B than to C, $F(1, 131) = 7.43$, $p < .01$, $\eta_p^2 = .05$, 90% CI [.01, .13].

Discussion

In Study 4, people first named the two stereotype content dimensions that they thought best describe the similarities and differences between 40 randomly selected social groups from our full set of 80 groups from Study 1. A relatively large number of given dimensions did not fit either of our combined ABC candidate dimensions unambiguously. Yet, for those that did, agency/socioeconomic success was consensually seen as the most fundamental dimension to describe similarities and differences between groups, and conservative–progressive beliefs was named as the second most important dimension. Subsequently, people spatially arranged the groups on a rating board with x- and y-axes labeled precisely according to the two stereotype content dimensions that they had just named. During this task, people were free to relabel one or both of the stereotype content dimensions that they had named *before*. The very low frequency of relabeling showed that most participants stuck to the originally named dimensions, which thus cannot be prompted by the spatial nature of the sorting task. Nevertheless, the two most suitable axial interpretations of the 2D social group space were again A and B, suggesting that in Studies 2 and 3 spatial A, B, and C metaphors (verticality, horizontality, and centrality, respectively) do not provide a sufficient explanation for the empirical support for our 2D ABC model of stereotype content. Further, in Study 4, the 2D group space had a sufficient scaling fit ($S <= .15$), and the ProFit analyses clearly confirmed the more fine-grained results of Study 3 that A best described the primary stereotype content dimension, followed by B as the best description for the secondary dimension, and that C emerged as a nonlinear function of A but not B.

Study 4 also addressed another issue. The most noteworthy difference between our data-driven stereotype content model and existing theory-driven models lies in the role of communion

[4] Separate property fitting analyses for the original 24 candidate stereotype dimensions show results that are highly consistent with Table 4, and that can be seen in the online supplementary material, Table osm.3.

or, alternatively, warmth. Whereas dominant theories not only suggest that warmth is a central dimension on which social groups are compared and judged, but also postulate a primacy of this dimension (for an overview, see Abele & Wojciszke, 2014), we only found support for an emergent nature of C. More precisely, Studies 1–4 suggest people compare and judge groups on A and B, and that C is encoded by A, with average and extreme A implying high and low C, respectively. One explanation for this difference in conclusions might be that the current studies tap into spontaneously employed dimensions, whereas previous work explicitly instructed participants to judge the C or warmth of groups. Another explanation might be that in our studies participants refrained from employing C-related stereotypical information, because they saw it as relatively more socially undesirable to denigrate groups on stereotypic C (e.g., "lawyers are dishonest, homeless are repellent, military are threatening, and punks are untrustworthy") compared with stereotypic A and B ("lawyers are overconfident, homeless are powerless, military are overly conservative, and punks are too alternative"). Such hesitancy would in all likelihood emerge most strongly if we explicitly encourage participants to label the dimensions they employ as we did in Study 4. An inspection of the results, however, suggests that the opposite was true: The suitability of C as a third axial interpretation of the 3D social group space was never high, but more pronounced in Study 4 than in any of the previous studies. This speaks against the idea that participants responded in a socially desirable way that prohibited the expression of perceived group differences in communion.

In sum, the studies so far support a 2D model of the mental organization of social groups. In marked contrast to previous theorizing about the nature of these dimensions, the data suggest that the primary and secondary mapping principles are agency/socioeconomic success (i.e., the groups' perceived wealth, status, power, dominance, confidence and competitiveness) and conservative–progressive beliefs (i.e., the groups' perceived position on a continuum ranging from liberal, alternative, science-oriented and modern to traditional, conventional, religious and conservative), respectively (see Figures 1–3). Communion (i.e., perceived trustworthiness, sincerity, warmth, benevolence, likability, and altruism) was not used as a criterion for distinguishing between the representatively sampled social groups, but emerged for those social groups who are stereotyped as average on A but not so much B.[5]

Despite the consistency of our findings so far, there remain some caveats that require further attention. Specifically, despite adherence to a data-driven strategy, every study required some decisions. Our decisions of (a) how to sample social groups, (b) how to instruct the arrangement of these groups, and (c) which candidate dimensions to use for our property fitting analyses may have biased our results, leading to the observed 2D ABC model. For example, our instruction of how to sample groups might have favored groups defined by their sociopolitical ideology, our arrangement instruction might have made similarity in sociopolitical opinion particularly salient and there might be further oblique candidate dimensions that are equally fitting candidates that we never collected ratings for. Study 5 sought to address these caveats in a comprehensive design.

Study 5

Study 5 sought to generalize the 2D ABC model of stereotype content beyond our previous approaches to sampling (see Studies 1a and 3a), comparing (see Studies 1b, 2, 3b, and 4), and rating (see Studies 1c and 3c) social groups. This was done to rule out that the consistent results we obtained in the first four studies were due to unduly influences of top-down decisions we made in pursuing our bottom-up, data-driven approach. Below, we outline four potential sources of bias in stimulus sampling, dissimilarity arrangement, and property fitting and how Study 5 addressed those.

Stimulus Sampling

First, the relatively abstract instructions according to which our participants named social groups ("name . . . groups that structure society") might have primed social categories related to low–high agency/socioeconomic success and/or conservative–progressive beliefs rather than different levels of communion and/or other stereotype content dimensions. Particularly our definition that social groups ". . . are based on how people behave or see the world . . ." might be interpreted as referring to religious and political ideology or lifestyle (e.g., Christians, Muslims, Republicans, Tea Party, conservatives, hippies, hipsters, goths, Democrats, liberals, independents, etc.). This would artificially increase the salience of conservative–progressive beliefs over other stereotype content dimensions. Further, we forced participants to name 40 groups, which might be more than what is typically sufficient/necessary to mentally represent and organize society. Possibly, it is mainly this surplus of groups that relates to different levels of agency/socioeconomic success and conservative–progressive beliefs.

A truly data-driven approach might require instructions under which different types of groups, including social categories, task groups (e.g., clubs, committees), and primary groups (e.g., family, friends) are equally accessible in memory. To this end, in Study 5a

[5] It is possible that the empirical support for the 2D ABC model obtained in Studies 1–4 hinges on our criterion for defining a social group as "representative of the U.S./German society." This criterion was that at least 10% of the participants named the group as "representative of the U.S./German society" in Study 1/3. The relatively large number of groups that reached this criterion (80 and 76 compared with less than 30 in other research; e.g., Fiske et al., 2002) necessarily introduces variance on several dimensions. Such variance is a prerequisite for people to place groups on simplifying stereotype content dimensions (Ford & Stangor, 1992; Nelson & Miller, 1995). This begs the question whether a more strict criterion of what constitutes a group leads to a less diverse sample, and thus to less and/or different stereotype content dimensions, if any at all. In another study, we sought empirical support for that our 2D ABC model of stereotype content is not an artifact of the relatively infrequently named groups. This additional study was identical to Study 4, except that participants spatially arranged only the 40 groups that people saw as most representative of U.S. society in Study 1 (i.e., Table 1's left columns; in Studies 1-4 participants spatially arranged random samples of 40 out of the 80 most frequently named groups). This study provided results highly consistent with Studies 1-4: The scaling fit of the 2D group space was satisfactory ($S = .07$). A, $R(2D\ axis) = .88$, $p < .001$, and B, $R(2D\ axis) = .91$, $p < .001$ (see online supplementary material, Tables osm.4 and osm.5), were again far better axial interpretations of the 2D space compared with C, $R(2D\ axis) = .53$, $p = .002$, and C was a suitable polar interpretation of the 2D space, r(2D pole) $= .33$, $p = .04$, whereas A and B were not. These results ruled out that the consistently found 2D ABC model of stereotype content was an artifact of overly inclusive sampling of social groups.

we used a minimalist definition of groups and minimalist naming instructions that did not prime certain kinds of groups. Additionally, we allowed participants to name any number between three and 30 social groups (e.g., Fiske et al., 2002). Consistent with Studies 1a and 3a, the selection that we used included all groups named by at least 10% of all participants.

To further validate the sampled groups, we created another selection. First, we selected all groups that were named at least twice. Of those, we selected the groups that appeared most frequently in a multibillion word text corpus that contains a vast variety of digitalized books published in recent years (e.g., Akpinar & Berger, 2015; Michel et al., 2011). Thus, the relevance of groups in our second, naturalistic selection was not determined by participants but based on frequency of appearance in cultural products like books.

Dissimilarity Arrangement

Second, the dissimilarity rating, arrangement, and labeling tasks might have been too broad and abstract. That is to say, participants may have used information that goes well beyond typical stereotypic comparisons based on character traits and personal encounters (Fiske et al., 2002; Koenig & Eagly, 2014). Specifically, it is conceivable that participants also based their dissimilarity ratings, arrangements, and labels on the degree to which members of the groups typically (dis)agree with each other in the social and political arena. In Study 5b, we tested whether the 2D ABC model of stereotype content is valid even if people are instructed to compare groups based on the characters of/personal encounters with typical group members.

Property Fitting

Third, perhaps our empirical support for the 2D ABC model of stereotype content is contingent on the instructions according to which participants rated the social groups on agency/socioeconomic success, conservative–progressive beliefs, and communion. In particular, we asked participants for their personal belief about the groups' A, B, and C rather than to ask for the groups' A, B, and C "as viewed by society" (Cuddy et al., 2007; Fiske et al., 2002, p. 884; Kervyn et al., 2013; Kervyn, Fiske, & Yzerbyt, 2015). Society's view of groups is closer to the definition of stereotypes as socially shared views (Fiske et al., 2002). Thus—to make sure that we measure A, B, and C stereotypes—in Study 5c we asked for A, B, C ratings "as viewed by society." In addition, we employed a different measure of communion that has been reported to better capture its essence (Kervyn et al., 2015).

And fourth, despite the good statistical fit of A and B as almost entirely orthogonal axes of the 2D social group spaces extracted in Studies 1–4, it is conceivable that there are other pairs of equally orthogonal and well-fitting stereotype dimensions that we overlooked. That is, our candidates possibly did not reflect the full range of stereotype dimensions that our participants used to mentally organize the social groups, and thus we cannot be sure that A and B is the best model of the two most important stereotype dimensions that people spontaneously use to distinguish between groups. To show that this is the case, in Study 5d we asked a new sample of participants to label nine rotated, equidistant axes that run through the origins of the 2D group spaces extracted in Study 5b, and we asked another sample of participants to categorize the generated axes labels as fitting well to either A, B, or C as defined in Study 1, or as "no match" if a label "does not fit well" to any of A, B, and C. If the 2D group spaces entail stereotype dimensions that we overlooked so far, we would expect sizable amounts of "no match" responses at the corresponding rotation angles. If, however, most spontaneously generated labels for virtually all rotation angles are categorized as fitting well to A or B (i.e., as being synonyms of A and B), this provides strong support that no other, oblique 2D space provides a better description of spontaneously activated stereotype content about groups.

Our hypotheses for Studies 5a–5d were that the 2D ABC model of stereotype content holds true (a) for the new minimalist sampling instruction and the new naturalistic sample of social groups, (b) for similarity-, character- and personal encounter-based comparisons of social groups, (c) for social groups rated on the relevant dimensions "as viewed by society" rather than single persons, and that (d) there will be no evidence for overlooking an alternative model.

Method and Results

Study 5a. Creating a minimalist and a naturalistic sample of social groups. We paid 100 MTurkers (39 women, 61 men; M = 32.21 years, SD = 10.89) \$0.5 to "name up to 30 social groups." Participants read these minimalist instructions: "Off the top of your head, what various types of people do you think today's society categorizes into groups?" These were the exact same instructions as in Fiske et al. (2002, p. 883; see also Kervyn et al., 2013, 2015), except that we dropped "(i.e., based on ethnicity, race, gender, occupation, ability, etc.)" to avoid priming groups defined by the ethnicity, race, gender, occupation and/or ability of their members. Participants had the possibility to list up to 30 groups, although a minimum of three was required. On average, participants named 14.61 groups (SD = 9.32).

Table 6 shows all 42 groups named by more than 10% of all participants.[6] Forty of these 42 groups had also been named by 10% of participants in Study 1a. Further, the frequency with which the 40 groups had been named in Study 1a substantially predicted the frequency with which they were named in Study5a, r = .83, $p < .001$. Thus, the minimalist social group sample in Study 5a was very similar to the social group sample in Study 1a. If anything, in the minimalist sample extreme scorers on A ("poor," "middle class," "rich") and B ("Democrats," "Republicans," "gays," "Christians," "liberals," "conservatives") were named more, not less, frequently. Thus, in hindsight the instructions employed in Study 1a do not seem to have biased the group sample in a way that social categories defined by their religious/political ideology/lifestyle were named disproportionally often.

To create the naturalistic sample, we recorded how often the 136 social groups[7] that were named by at least two participants appear in the Google Books Corpus (Michel et al., 2011), the world's largest collection of digitized and searchable books (>5 million contain-

[6] We added the frequency of naming for the synonymic social groups Blacks/African Americans, Church/Christians, rich/wealthy, Hispanics/Latinos, elderly/old/seniors, upper class/elites, athletes/sportsmen, and atheists/nonreligious.

[7] We added the frequency of occurrence for the synonymic social groups children/kids, old/elderly/seniors, Blacks/African Americans, Church/Christians, military/veterans, rich/wealthy, Indians/Native Americans, athletes/sports, Hispanics/Latinos, elites/upper class, entrepreneurs/business owners, atheists/nonreligious, and boy/girl scouts.

Table 6
Most Frequently Named Social Groups in the U.S. (Consensus > 10%) in Study 5a

1st–21st Most frequent minimalist groups	22nd–42th Most frequent minimalist groups	1st–31st Most frequent naturalistic groups	32nd–61st Most frequent naturalistic groups
Blacks (50%)	Athletes (15%)	Children (13.12m)	Professionals (.94m)
Whites (41%)	Parents (15%)	Women (10.80m)	Muslim/s (.93m)
Poor (37%)	Nerds (14%)	Old (10.51m)	Conservative/s (.88m)
Middle class (34%)	Hippies (14%)	Family (10.43m)	Scientists (.87m)
Rich (33%)	Immigrants (14%)	Men (9.05m)	Tall (.86m)
Hispanics (31%)	Atheists (13%)	White/s (7.91m)	Republican/s (.84m)
Asians (29%)	Blue collar (13%)	Black/s (7.88m)	Artists (.84m)
Democrats (29%)	Religious (13%)	Christians (7.58m)	Lesbian/s (.70m)
Republicans (29%)	Men (12%)	Students (7.49m)	Actors (.69m)
Gays (27%)	Teenagers (12%)	Young (6.70m)	Immigrants (.60m)
Christians (26%)	White collar (12%)	Short (4.68m)	Hispanic/s (.60m)
Liberals (26%)	Politicians (12%)	Parents (4.35m)	Farmers (.59m)
Conservatives (26%) (26%)	Jocks (11%)	Poor (3.83m)	Teenagers (.52m)
Working class (22%)	Hipsters (11%)	Jewish (3.47m)	Educated (.51m)
Transgender (21%)	Celebrities (11%)	Friends (3.37m)	Elites (.50m)
Elderly (20%)	Drug addicts (11%)	Military (3.36m)	Democrat/s (.41m)
Students (19%)	Homosexuals (10%)	Religious (3.08m)	Clubs (.38m)
Lesbians (17%)	Homeless (10%)	Americans (2.43m)	Homosexual/s (.38m)
Women (16%)	Jews (10%)	Rich (2.30m)	Politicians (.38m)
Upper class (15%)	Goths (10%)	Gay/s (2.21m)	Musicians (.36m)
Muslims (15%)	Lower class (10%)	Europeans (2.19m)	Activists (.30m)
		Chinese (2.05m)	Minorities (.28m)
		Indian/s (1.89m)	Law enforcement (.26m)
		Straight (1.83m)	Alcoholics Anonymous (.26m)
		Adults (1.80m)	Entrepreneurs (.24m)
		Athletes (1.43m)	Catholics (.24m)
		Writers (1.26m)	Homeless (.24m)
		Ethnic (1.21m)	Mexicans (.20m)
		Asian/s (1.15m)	Rebels (.18m)
		Employed (1.04m)	Middle class (.17m)
		Liberal/s (.94m)	

Note. Left sample: percentage in parentheses is proportion of participants who spontaneously named this group as part of today's U.S. society. Right sample: number in parentheses is millions of occurrences in contemporary (i.e., 2000–2009) American English literature (i.e., ~27 billion words) according to the Google Books Corpus (Davies, 2011).

ing >500 billion words). To measure these groups' frequency of occurrence in contemporary American English literature, we searched only within the 2000–2009 publication period of the American English section of the Google Books Corpus (= 26.9 billion words) provided by Davies (2011; see http://googlebooks.byu.edu/). Taking the average between the Study 1a sample ($N = 80$) and the minimalist sample ($N = 42$), the naturalistic sample included the 61 groups that we found to be most prevalent in this collection of texts (see Table 6). Forty-six of these 61 groups had also been named by 10% of participants in Study 1a. However, the frequency with which the 46 groups had been named in Study 1a did not predict their frequency of occurrence in contemporary American English literature, $r = .17$, $p = .25$. Noteworthy, the most frequently occurring groups in the naturalistic sample did not score either high or low on either A or B, but rather reflected differences in race, sex, and age ("children," "women," "old," "men," "Whites," "Blacks").

Study 5b. Character-/personal encounter-based arrangement of 42/61 groups. We paid 378 MTurkers (148 women, 230 men; $M = 33.94$ years, $SD = 10.74$) \$0.75 to "sort 42 social groups on the computer screen." They arranged either the 42 minimalist groups, or a random selection of 42 of the 61 naturalistic groups. The arrangement task was the same as in Studies 2 and 3—with two exceptions. First, to give participants an overview of the groups, they appeared all at once in a random order of four columns and 11 rows in the middle of the screen. More importantly, there were three different arrangement instructions. In the similarity (control) condition, participants read: ". . . social groups whose typical members are similar should be placed closer together, while social groups whose typical members are different should be placed further apart." In the character condition, they read: ". . . social groups whose typical members have similar characters should be placed closer together, while social groups whose typical members have different characters should be placed further apart." And in the personal encounter condition, they read: ". . . social groups for which personal encounters with their typical members are similar should be placed closer together, while social groups for which personal encounters with their typical members are different should be placed further apart"[8] (for the minimalist sample, there were between 49 and 51 participants per condition; for the naturalistic sample there were between 74 and 77 partici-

[8] To ensure that participants follow these different instructions, we presented them not only before, but also during the spatial arrangement phase, namely in abbreviated from (e.g., "similar character -> closer together; different character -> further apart") left to a "Continue" button at the bottom of the screen.

pants per condition). As in Studies 2–4, the arranged distances between the groups were recorded as proportions of the screen diagonal.

Study 5c: Testing the validity of the 2D ABC model of stereotype content. Next, 201 MTurkers (70 women, 131 men; $M = 33.03$ years, $SD = 11.07$) were paid \$0.6 to "rate about 50 social groups on a stereotype dimension." They rated the 42 minimalist groups or the 61 naturalistic groups on either A, B, or one of two versions of C. Each of the eight corresponding 0–10 slider scale items was anchored with a meaning cloud (for an example, see Figure osm.3 in the online supplementary material; meaning clouds accurately measure groups' A, B, and C, see Footnote 1) of all subdimensions of either A, B, C (see Table 3), or C2. Based on recommendations in the literature (Kervyn et al., 2015), the second version of communion was anchored with "Not at all . . . –Extremely friendly/sincere/sociable/well-intentioned." Above the slider scale items, participants read: "As viewed by society, how . . . [e.g., friendly, sincere, sociable, and well-intentioned] are members of these groups?" There were between 22 and 28 raters per stereotype dimension, and as in the previous studies, raters' agreement about the groups was very high, all *ICC(2,k)*s $> = .85$, (McGraw & Wong, 1996). While the expected correlations between the two versions of communion in the minimalistic sample ($r = .87$), and in the naturalistic sample ($r = .86$) were large, ps $< .001$, there was another moderate but statistically significant correlation between agency/socioeconomic success and progressive beliefs in the minimalistic sample, $r = .34$, $p < .05$, all other $|r|$s $< .24$, ps $> .05$.[9]

Next, we computed the mean distance between each pair of spatially arranged groups, separately for the minimalist and the naturalistic sample, and separately for the similarity-, character-, and personal encounter-based instructions. For the minimalist sample, these mean distances correlated highly across the three different spatial arrangement instructions, mean $r = .90$, $SD = .06$, and the same was true for the naturalistic sample, mean $r = .80$, $SD = .06$. Thus, we collapsed mean intergroup distance across the three different spatial arrangement instructions, separately for the minimalist and the naturalistic sample of groups.[10]

The mean distances between the groups were subjected to MDS (separately for the minimalist and the naturalistic sample) with the same settings as in the previous studies. As in Studies 3 and 4, based on the goodness of fit of the 1D, 2D, 3D, 4D, 5D, and 6D scaling solutions (see Table 2), we extracted and analyzed three dimensions, but proceeded with interpreting only the 1D and 2D space of the minimalist and the naturalistic groups. Property fitting analyses confirmed the validity of the 2D ABC model of stereotype content. Agency/socioeconomic success, *R(2D axis)* $= .93$, $p < .001$, and conservative–progressive beliefs, *R(2D axis)* $= .94$, $p < .001$, were far better *axial* interpretations of the minimalist groups' 2D space than communion, *R(2D axis)* $= .13$, $p = .72$, and the second version of communion, *R(2D axis)* $= .20$, $p = .47$. The same was true for the 2D space of the naturalistic groups; A: *R(2D axis)* $= .81$, $p < .001$; B: *R(2D axis)* $= .86$, $p < .001$; C: *R(2D axis)* $= .12$, $p = .67$, and C2: *R(2D axis)* $= .15$, $p = .52$ (see Table 7; see also Figures 4 and 5). Further, as in the previous studies C and C2 were suitable *polar* interpretations of the minimalist groups' 2D space, r(2D pole) $= .60$, $p < .001$ and r(2D pole) $= .50$, $p < .001$, respectively, and the naturalistic groups' 2D space, r(2D pole) $= .61$, $p < .001$ and r(2D pole) $= .74$, $p < .001$, respectively. In contrast, A and B were not suitable as polar interpretations of these spaces.

Study 5d: Ruling out alternative 2D models of stereotype content. One-hundred and 80 additional MTurkers (82 women, 98 men; $M = 31.65$ years, $SD = 9.38$) were paid \$0.75 to "identify nine person characteristics." We rotated the 2D coordinates of the 42 minimalist groups clockwise around the origin of their space. We rotated in 18 steps of 20° (= a full rotation of 360°). At each rotation step, we formed a 1D ranking based on the groups' current x-coordinates (i.e., after nine rotation steps—a half rotation of 180°—each of these 18 rankings was reversed). We presented participants with the group rankings of nine consecutive rotation steps, one at a time and in random order. These nine rankings represented nine axes that run through (the origin of) the minimalist groups' 2D space in such a way that any so far overlooked stereotype dimension would have a maximal distance of 10° from one of these nine axes. A rotation angle of 10° corresponds to a correlation of $r > .98$, and thus the nine axes included all fundamental stereotype dimensions that we have overlooked in our previous analyses—if there are any. If agency/socioeconomic success and conservative–progressive beliefs are the only two stereotype dimensions encoded in the minimalist groups' 2D space, then the collection of labels for all pairs of reversed axes (i.e., two axes between which the rotation angle is 180°) should reflect A or B and not C or something else.

To understand the task, participants were first presented with an example in which the animal characteristic based on which "giraffe, elephant, horse, deer, dog, mouse, and bee" were ranked one atop the other was labeled as "tall" and/or "big." Then, before labeling each of the nine target axes, participants were presented with the minimalist groups one atop the other in the order of their x-coordinates on the corresponding axis, and they read: "Your task is to identify X. X is the person characteristic based on which the social groups are ranked. As viewed by society, groups at the top of the ranking are extremely X. . . . groups above the center of the ranking are above-averagely X. . . . groups at the center of the ranking are averagely X. . . . groups below the center of the ranking are below-averagely X. . . . groups at the bottom of the ranking are not at all X. Please enter the person characteristic X (an adjective) in the textbox below. If you have no idea about X, enter 'I don't know.'" Participants generated a total of 521 labels for the nine pairs of reversed axes of the minimalist groups' 2D space (due to

[9] The main difference between the group rating instructions in Studies 1 and 5 was that in Study 5 participants rated the groups "as viewed by society" rather than the self. If participants perceived mainstream society as relatively conservative, then, assuming that society trusts and likes society, participants would have rated conservative groups as scoring higher on C than progressive groups. If so, then this effect should be less pronounced in Study 1, where participants rated the groups as viewed by the self rather than society. However, the groups' B and C correlate $r = -.01$ in Study 1 and $r = .04$ (minimalist sample) and $r = .05$ (naturalistic sample) in Study 5, suggesting that participants in Study 5 perceived mainstream society as neither conservative nor progressive.

[10] Separate property fitting analyses for the 1D–3D spaces extracted from the similarity-, character-, and personal encounter-based mean intergroup distances yielded almost identical results to the property fitting analyses for the 1D–3D spaces extracted from the mean intergroup distances collapsed across the these three spatial arrangement instructions, as shown in the online supplementary material in Tables osm.6 (minimalistic groups) and osm.7 (naturalistic groups).

Table 7
Property Fitting Results for Studies 5 and 6

	Group sample	Stereotype content	*R(1D axis)*	*R(2D axis)*	*R(3D axis)*	r(1D pole)	r(2D pole)	r(3D pole)
Study 5	42 minimalist U.S. groups	Agency (A)	**.777**	**.929**	**.909**	−.188	−.093	−.097
		Beliefs (B)	**.783**	**.937**	**.935**	.136	.143	.159
		Communion (C)	.050	.129	.247	**.551**	**.595**	**.590**
		Communion (C2)	.149	.196	.272	**.521**	**.502**	**.545**
Study 5	61 naturalistic U.S. groups	Agency (A)	**.765**	**.806**	**.803**	.072	.086	.116
		Beliefs (B)	.354	**.863**	**.890**	**.487**	.195	.261
		Communion (C)	.097	.118	**.526**	**.432**	**.607**	**.540**
		Communion (C2)	.113	.150	**.569**	**.640**	**.737**	**.665**
Study 6	42 minimalist U.S. groups	Agency (A)	**.826**	**.917**	**.923**	−.184	−.100	−.078
		Beliefs (B)	**.708**	**.936**	**.945**	.186	.222	.218
		Communion (C)	.058	.055	.084	**.640**	**.642**	**.668**
		Communion (C2)	.181	.194	.327	**.567**	**.567**	**.625**
Study 6	61 naturalistic U.S. groups	Agency (A)	**.691**	**.741**	**.906**	−.064	.136	.094
		Beliefs (B)	**.657**	**.901**	**.945**	.391	.339	.328
		Communion (C)	.015	.102	.381	**.480**	**.518**	**.507**
		Communion (C2)	.054	.236	.447	**.619**	**.613**	**.604**

Note. *R(1D–3D axis)* indicate the maximal correlations between the 42/61 minimalist/naturalistic U.S. social groups' agency/socioeconomic success, conservative–progressive beliefs, and communion (standard and alternative operationalization) ratings and their projections on an axis rotated around the origin of their 1D, 2D, and 3D space; *r*(1D–3D pol) indicate correlations between the social groups' A, B, and C (standard and alternative operationalization) ratings and their proximity to the origin of these three spaces. Bold correlation coefficients are significant at $p = < .001$.

redundancy 274 unique labels; "I do not know" = 38.27% of all cases). We repeated this axes labeling procedure for the 61 naturalistic groups' 2D space. Other participants generated a total of 516 labels for the nine pairs of reversed axes of the naturalistic groups' 2D space (265 unique labels; "I do not know" = 33.33% of all cases).

One-hundred and eight additional MTurkers (47 women, 61 men; $M = 30.47$ years, $SD = 9.05$) were paid \$0.5 to "assign 100 person characteristics" one after the other, and they read: "Please select the category to which this person characteristic [e.g., wealthy] fits best. If this person characteristic does not fit well to any of the categories, check 'no match.'" The seven categories available for selection were "no match" plus the polar opposites on A, B, and C (A+: "powerful/dominating/high status/wealthy/confident/competitive;" A-: "powerless/dominated/low status/poor/unconfident/unassertive;" B+: "modern/science-oriented/alternative/liberal;" B-: "traditional/religious/conventional/conservative;" C+: "trustworthy/sincere/likable/benevolent/warm/altruistic;" and C-: "untrustworthy/dishonest/repellent/threatening/cold/egoistic")—the candidate stereotype dimensions examined in Studies 1–5. Participants assigned either 100 random of the 274 different labels generated for the (nine pairs of reversed axes =) 18 axes of the minimalist groups' 2D space, or 100 random of the 265 different labels generated for the 18 axes of the naturalistic groups' 2D space. On average, each label generated for an axis of minimalist and naturalistic groups' 2D space was assigned by 21.10 ($SD = 3.68$) and 18.86 ($SD = 3.41$) participants, respectively. For each of the 512 and 516 labels generated for one of the axes of the minimalist and naturalistic groups' 2D space, respectively, we recorded the percentage of assignments to categories A ±, B ±, C ±, and "no match," a measure of the labels' relatedness to A, B, C, and something else (rel:A, rel:B, rel:C, and rel:Else), respectively. Finally, for each of the nine pairs of reversed axes of the minimalist and naturalistic groups' 2D space, we averaged rel:A, rel:B, rel:C, and rel:Else across all labels generated for that pair of axes.

Table 8 shows mean relatedness of the participant-generated labels to A, B, C, and something else (rel:A, rel:B, rel:C, and rel:Else), separately for the nine pairs of reversed axes of the minimalist groups' 2D space and the naturalistic groups' 2D space. All pairs of axes in both 2D group spaces predominantly related to agency/socioeconomic success or conservative–progressive beliefs rather than communion or something else. Therefore, according to participants in the label generation and assignment studies, the 2D space of both the minimalist and the naturalistic groups does *not* encode a fundamental, spontaneously employed stereotype dimension other than A and B.

Figures 4 (minimalist groups) and 5 (naturalistic groups) illustrate this. The left side of these figures shows the groups' coordinates in their 2D space plus the two axes that best represent A and B, which together form the best available explanation of the variance contained in this space (see the property fitting results in Study 5). The right side shows the same space (aligned in the same direction) plus the 18 axes for which we collected labels and label assignments to the stereotype dimensions A, B, C, and something else (= "no match"). For each of these 18 axes, the far end shows the label most often generated for that axis. All these most consensual axes labels are relatable to A or B. More importantly, for each of the 18 axes, the stretch from the spaces' origin to the far end of that axis indicates the percentage of participants who assigned the labels generated for that axis to A, B, C, and something else that "does not fit well" to A, B, and C. As is immediately evident, all nine pairs of reversed axes reflect A or B rather than C or something else, and the axes that reflect A and B are orthogonal to one another and run through the space more or less exactly at the angles that best represent A and B (see the left side of the figures) according to the property fitting analyses in Study 5. That is, in Study 5 we took a data-driven approach not just to scaling the groups' 2D space, but also to interpreting this space, and results showed that A and B is the one and only pair of orthogonal stereotype dimensions that underlies this space.

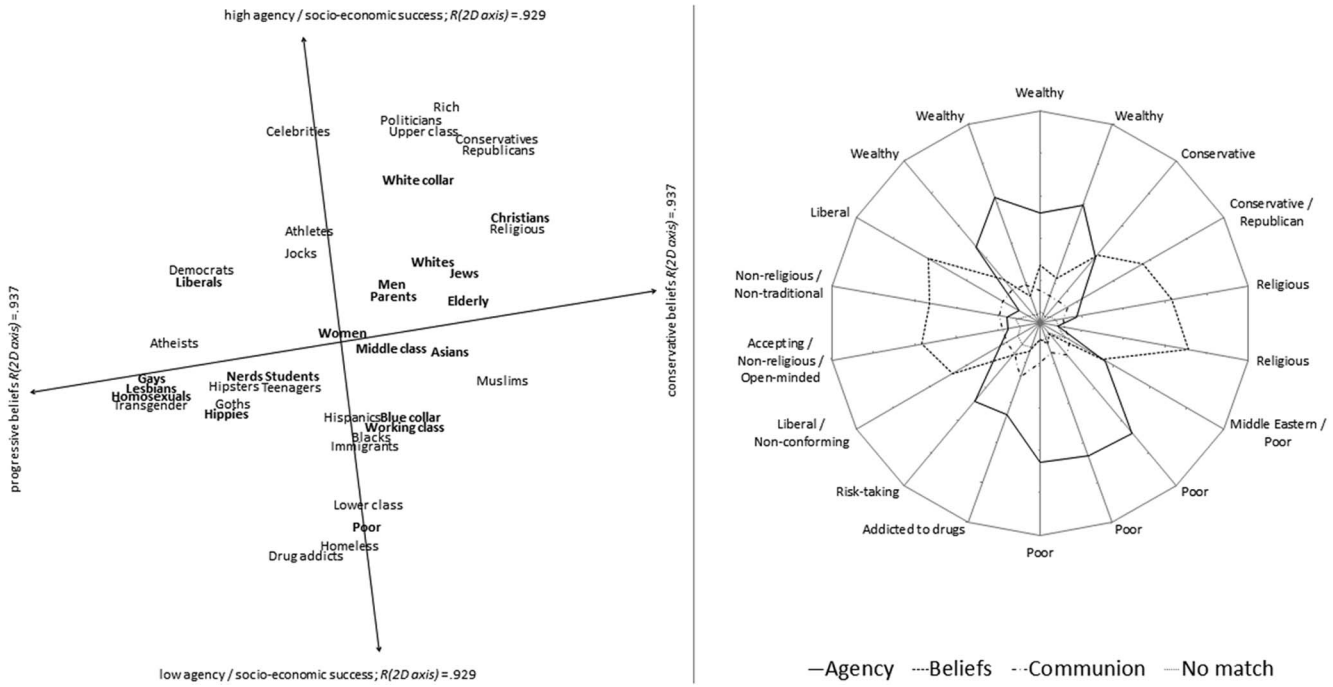

Figure 4. Study 5: The left side illustrates that the minimalist groups' 2D stereotype space is made up of the axes agency/socioeconomic success and conservative–progressive beliefs and the centrally located pole communion (21 most/least communal groups = bold/not bold). The right side plots the most frequent labels for the nine pairs of reversed axes of this space, and the percentage of labels for these axes that were assigned to A, B, C, and "no match." All axes mainly reflect A or B at the angle where these two run through the space (see property fitting results on the left side).

Discussion

To rule out that the 2D ABC model of stereotype content is limited to the detailed and thus possibly biased instructions under which participants in Studies 1 and 3 named groups, in Study 5 participants received the minimalist naming instructions used by Fiske et al. (2002, p. 883; see also Kervyn et al., 2013, 2015), except that we dropped "(i.e., based on ethnicity, race, gender, occupation, ability, etc.)" to avoid priming groups defined by the ethnicity, race, gender, occupation, and/or ability of their members. Moreover, to omit groups that are not essential to participants' view of society, participants were free to name any desired number between three and 30 groups (see Kervyn et al., 2013, 2015). Both the groups in this "minimalist" sample and their frequency of naming were highly similar to the Study 1 sample of groups, and Study 5 shows that the 2D ABC model generalizes well from the Study 1 sample to the minimalist sample of groups.

To generalize the model to a naturalistic sample of groups, we recorded the frequency with which all groups that were named at least twice in Study 5 appear in a large text corpus that is arguably representative of contemporary American English literature (Davies, 2011; Michel et al., 2011; for another example of such a linguistic approach to personality and social psychology research, see Akpinar & Berger, 2015). The groups mentioned most often between 2000 and 2009 formed our "naturalistic" sample, which is somewhat different from the Study 1's sample of groups. The results of Study 5 showed that the 2D ABC model generalizes to this naturalistic sample as well.

To rule out that the 2D ABC model is limited to the instructions under which participants in Studies 1–4 compared groups, in Study 5 we instructed participants to spatially arrange the minimalist or the naturalistic groups either based on the global dissimilarity of their typical members (see Studies 1–4), based on the dissimilarity of the character of their typical members, or based on the dissimilarity of personal encounters with their typical members. The latter two types of instructions may better reflect the essence of stereotypic social group comparisons (Fiske et al., 2002; Koenig & Eagly, 2014). However, the three types of instructions yielded almost identical group comparisons for both the minimalist and the naturalistic sample, and thus the validity of the 2D ABC model of stereotype content generalizes from unspecified to character- and personal encounter-based group comparisons (see also online supplementary material, Tables osm.6/osm.7).

To rule out that the 2D ABC model is limited to the instructions under which participants in Studies 1–4 rated groups, in Study 5 participants rated the minimalist and the naturalistic groups' stereotypic agency/socioeconomic success, conservative–progressive beliefs, and communion as viewed by society (Cuddy et al., 2007; Fiske et al., 2002; Kervyn, Fiske, & Yzerbyt, 2013, 2015) rather than themselves. Rating groups from the perspective of society rather than the self is arguably closer to the definition of stereotypes as socially shared views, and additionally circumvents single individuals' tendency to respond in a socially desirable way that eliminates meaningful variance on valence-related stereotype dimensions such as communion (Fiske et al., 2002). Further, in

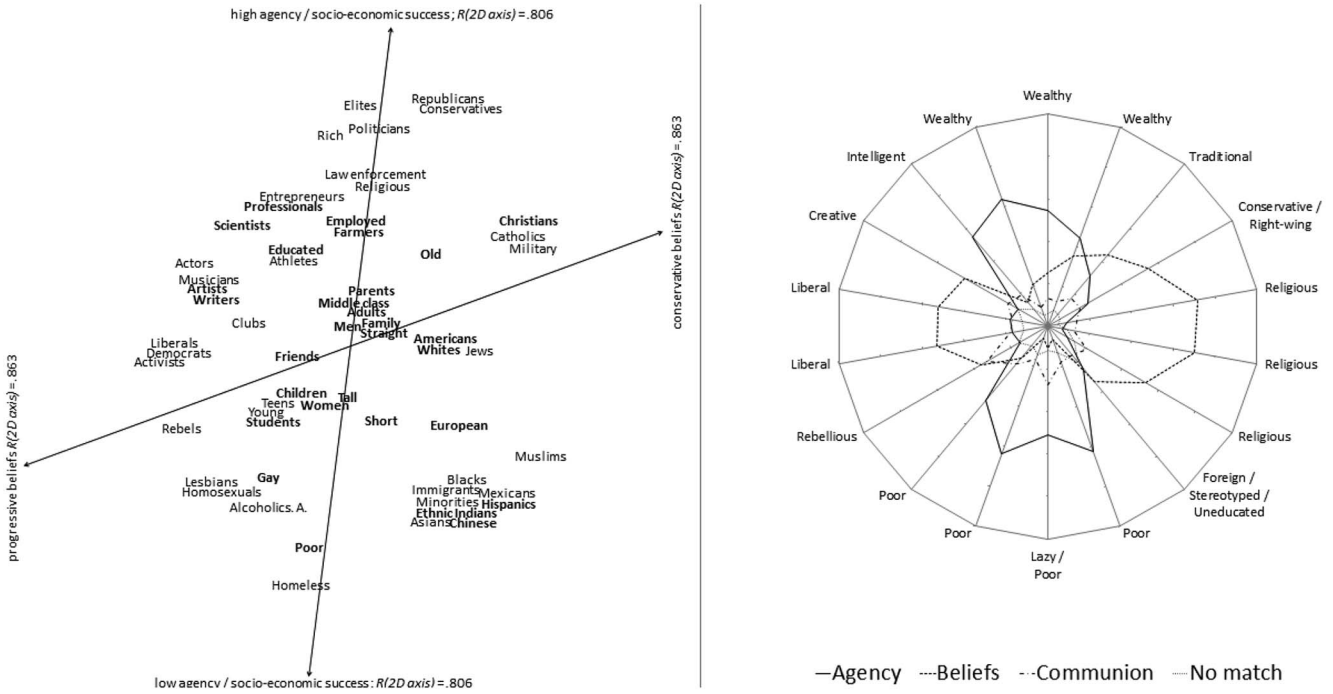

Figure 5. Study 5: The left side illustrates that the naturalistic groups' 2D stereotype space is made up of the axes agency/socioeconomic success and conservative–progressive beliefs and the centrally located pole communion (most/least communal groups = bold/not bold). The right side plots the most frequent labels for the nine pairs of reversed axes of this space, and the percentage of labels for these axes that were assigned to A, B, C, and "no match." All axes mainly reflect A or B at the angle where these two run through the space (see property fitting results on the left side).

Study 5 participants also rated both the minimalist and the naturalistic groups on another set of communion subscales that has been reported to better reflect its essence, namely "friendly/sincere/sociable/well-intentioned." Results showed that this alternative our version of communion are almost identical in meaning ($rs >= .86$), and that the ABC model generalizes from groups' A, B, and C as viewed by the self to their A, B, and C/C2 "as viewed by society" (Fiske et al., 2002, p. 884).[11]

Furthermore, we aimed to rule out that there are other models of stereotype content that we might have overlooked when selecting candidate stereotype dimensions based on a visual inspection of the 2D group space computed in Study 1. To that end, we asked new participants to label the stereotype dimensions that underlie different group rankings that together reflect all axes (i.e., stereotype dimensions) that run through the origin of the groups' 2D space. Yet other participants then assigned the generated stereotype dimension labels to our candidates A, B, or C as defined in Study 1, and they were instructed to select "no match" if a label "does not fit well" to A, B, or C. For each axis/stereotype dimension of the groups' 2D space, our results showed that participants predominantly assigned the labels generated for that axis/stereotype dimension to A or B rather than C or something else (= "no match"), a pattern that was found for both the 2D space of the minimalist and the naturalistic groups (see Table 8 and Figures 4 and 5). Thus, participants spontaneously used A and B (not C or something else) to stereotypically compare the groups.

Study 6

Study 6 addresses another caveat. The 2D ABC model of stereotype content may only apply to distinguishing between the entirety of groups that together form society. In Studies 1–5, participants always compared either all groups in the respective sample, or a randomly drawn set of groups that is more or less representative of all groups in the respective sample. Thus, for their comparisons participants had to spontaneously select stereotype dimensions on which all groups in the respective sample can be meaningfully placed. However, in real life people likely compare self-selected rather than representative or complete sets of groups, maybe because they want to compare some groups on a stereotype dimension on which only those and not all groups can be placed well, for example because they have no idea about the other groups' position or construe them as highly heterogeneous regarding this dimension. In principle it may be that participants predominantly process communion/warmth information when ste-

[11] Forthy minimalist and 46 naturalistic groups are also part of the Study 1 sample, which allows correlating these groups' stereotypic A, B, and C "as viewed by society" (measured in Study 5) with their A, B, and C as viewed by single persons (measured in Study 1). For the minimalist groups, these A, B, and C correlations are $r = .98$, .97, and .92, respectively. For the naturalistic groups, the correlations are $r = .98$, .97, and .88, respectively. Thus, it does not make a difference whether group stereotypes are measured as viewed by society or by single persons.

Table 8
Label Assignment Results for Studies 5 and 6

Relatedness to stereotype content	Axes 0/180°	Axes 20/200°	Axes 40/220°	Axes 60/240°	Axes 80/260°	Axes 100/280°	Axes 120/300°	Axes 140/320°	Axes 160/340°
Study 5									
Minimalist 2D space									
Agency (A)	**59**	**52**	45	23	17	12	22	**57**	**65**
Beliefs (B)	18	18	30	**52**	**60**	**63**	49	18	12
Communion (C)	17	21	16	16	14	16	19	20	17
Something else	7	9	9	9	9	10	10	6	6
Naturalistic 2D space									
Agency (A)	**53**	**63**	44	14	12	13	18	39	**54**
Beliefs (B)	18	14	22	**48**	**62**	**61**	44	31	20
Communion (C)	20	13	19	22	15	16	26	20	14
Something else	9	9	15	16	11	9	12	10	11
Study 6									
Minimalist 2D space									
Agency (A)	**58**	**53**	**49**	23	19	12	11	44	**58**
Beliefs (B)	7	16	17	**47**	**46**	**53**	53	14	14
Communion (C)	27	21	26	24	26	27	30	36	20
Something else	8	9	8	6	9	8	7	6	7
Naturalistic 2D space									
Agency (A)	**50**	**43**	36	21	11	19	19	24	35
Beliefs (B)	18	19	19	35	**57**	**55**	**52**	**49**	34
Communion (C)	21	22	25	29	21	17	20	18	20
Something else	12	15	20	15	11	8	9	8	10

Note. Axes 0°/180°–160°/340° indicate the mean percentage of participants who assigned the labels generated for the respective pair of reversed axes of the respective 2D group space to the categories agency/socioeconomic success, conservative–progressive beliefs, communion, and something else. Participants saw all nine pairs of reversed axes of both 2D group spaces as related to A or B rather than communion or something else. Because the nine pairs of reversed axes include all stereotype dimensions encoded in the respective 2D group space, these results add to the corresponding property fitting results that A and B are the *only* suitable axial interpretations of both 2D group spaces. Bold numbers indicate paired *t*-tests of the highest against the second highest percentage that are significant at $p = < .001$.

reotyping groups but forcing them to rate a large number of groups for which they have no clear stereotype about their communion undermines spontaneously employing this dimension. Thus, by omitting a phase in which participants self-select groups to be stereotypically compared, in Studies 1–5 we might have artificially limited the range of stereotype dimensions that participants could spontaneously select to only those dimensions that can be meaningfully applied to all social groups.

Study 6 therefore added a phase in which participants could freely choose the groups that they would subsequently compare. If participants decide to compare different groups on different stereotype dimensions, the scaling of 1D–3D group spaces to be interpreted will entail a poor statistical fit, and these group spaces will not entail meaningful stereotype dimensions that can be interpreted based on candidate stereotype dimensions (see property fitting analyses in Studies 1–5). If, however, participants decide to compare different groups on more or less the same few stereotype dimensions, we will be able to reveal the nature of these fundamental stereotype dimensions if they are among our candidates A, B, and C. We hypothesized that participants decide to compare different groups on A and B rather than C and/or something else.

Method and Results

We paid 751 MTurkers (240 women, 411 men; $M = 32.30$ years, $SD = 10.45$) \$0.75 to "select and sort 21 social groups on the computer screen." Participants were presented with either the 42 minimalist groups or 42 random naturalistic groups (out of 61), and were instructed to select at least 21 of these groups to spatially arrange them based on either the global dissimilarity of their typical members, the dissimilarity of the character of their typical members, or the dissimilarity of personal encounters with their typical members. We set a minimum of 21 groups, because selecting half of the available groups holds a balance between increasing the number of stereotype dimensions on which the groups can be placed (i.e., compared) and decreasing the number of participants required to obtain reliable dissimilarity estimates for all possible pairs of groups (861 and 1,830 in the minimalist and naturalistic sample, respectively), a necessity for an accurate scaling of the 1D–3D group spaces to be interpreted (Borg & Groenen, 2005). On average, participants selected 21.80 minimalistic and 21.58 naturalistic groups (for frequency of selection of all minimalist and naturalistic groups averaged across the three spatial arrangement instruction conditions, see online supplementary material, Table osm.8). For the minimalist groups, frequency of selection correlated with frequency of naming in Study 5, $r = .40$, $p < .01$; for the naturalistic groups, frequency of selection correlated with frequency of appearance in contemporary (2000–2009) American English literature according to the Google Books Corpus (Davies, 2011), $r = .42$, $p = .001$. Apparently, in both the minimalist and the naturalistic condition participants most often selected groups defined by the race, sex, and age of their members, and groups perceived as either high or low on either A or B.

Next, participants spatially arranged the self-selected groups (the instructions and procedure were the same as in Study 5; between 97 and 100 participants per condition for the minimalist sample and between 151 and 154 participants per condition for the

naturalistic sample). As in Studies 2–5, the spatially arranged distances between the groups were recorded as proportions of the screen diagonal.

Next, we computed the mean distance between each pair of spatially arranged groups, separately for the minimalist and the naturalistic sample, and separately for the similarity-, character-, and personal encounter-based instructions. As in Study 5, these mean distances correlated highly across the three different spatial arrangement instructions (mean $r = .82$, $SD = .09$ for the minimalist groups, and mean $r = .62$, $SD = .03$ for the naturalistic groups), and thus we collapsed mean intergroup distance across the three different spatial arrangement instructions, separately for the minimalist and the naturalistic sample of groups.[12]

The mean distances between the groups were subjected to MDS (separately for the minimalist and the naturalistic groups) with the same settings as in the previous studies (Table 2 shows the goodness of fit of the 1D, 2D, 3D, 4D, 5D, and 6D scaling solutions). Property fitting analyses confirmed the validity of the 2D ABC model of stereotype content. A, *R(2D axis)* = .92, $p < .001$, and B, *R(2D axis)* = .94, $p < .001$, were far better *axial* interpretations of the minimalist groups' 2D space than C, *R(2D axis)* = .06, $p = .94$, and the second version of C, *R(2D axis)* = .19, $p = .47$. The same was true for the 2D space of the naturalistic groups; A: *R(2D axis)* = .74, $p < .001$; B: *R(2D axis)* = .90, $p < .001$; C: *R(2D axis)* = .10, $p = .74$, and C2, *R(2D axis)* = .24, $p = .19$ (see Table 7, see also Figures 6 and 7). Further, as in the previous studies, C and C2 were suitable *polar* interpretations of the minimalist groups' 2D space, r(2D pole) = .64, $p < .001$ and r(2D pole) = .57, $p < .001$, respectively, and the naturalistic groups' 2D space, r(2D pole) = .52, $p < .001$ and r(2D pole) = .61, $p < .001$, respectively. In contrast, A and B were not suitable as polar interpretations of these spaces.

Two-hundred and 19 additional MTurkers (88 women, 131 men; $M = 31.35$ years, $SD = 9.51$) were paid \$0.75 to "identify nine person characteristics." As in Study 5, to find out if our candidates A, B, and C do not include one or more fundamental stereotype dimensions encoded in the groups' 2D space, we asked participants to label nine pairs of reversed group rankings that represent nine axes that run through (the origin of) this space in such a way that one of them correlates at least $r = .98$ with any stereotype dimension not included in A, B, and C. If, as predicted, A and B are the only two stereotype dimensions encoded in the groups' 2D space, then the collections of labels for all nine axes should predominantly reflect A and B.

Participants generated a total of 758 labels for the nine axes of the minimalist groups' 2D space (due to redundancy 347 unique labels; "I do not know" = 22.25% of all cases), and other participants generated a total of 730 labels for the nine axes of the naturalistic groups' 2D space (332 unique labels; "I do not know" = 33.16% of all cases).

One-hundred and 20 additional participants (54 women, 66 men; $M = 33.02$ years, $SD = 10.82$) were paid \$0.5 to "assign 100 person characteristics" to A, B, C, or "no match" (see Studies 4 and 5). Participants assigned either 100 random of the 347 different labels generated for the nine axes of the minimalist groups' 2D space, or 100 random of the 332 different labels generated for the nine axes of the naturalistic groups' 2D space. On average, each label generated for an axis of minimalist and naturalistic groups' 2D space was assigned by 16.92 ($SD = 3.39$) and 18.36 ($SD = 3.41$) participants, respectively. For each of the 758 and 730 labels generated for one of the axes of the minimalist and naturalistic groups' 2D space, respectively, we recorded the percentage of assignments to categories A, B, C, and "no match," a measure of the labels' relatedness to A, B, C, and something else, respectively. Finally, for each of the nine axes of the minimalist and naturalistic groups' 2D space, we averaged relatedness to A, B, C, and something else across all labels generated for that axis.

Table 8 shows mean relatedness of the participant-generated labels to A, B, C, and something else, separately for the nine axes of the minimalist groups' 2D space and the naturalistic groups' 2D space. As in Study 5, all axes in both 2D group spaces predominantly relate to agency/socioeconomic success or conservative–progressive beliefs rather than communion or something else (see Figures 6 and 7). Thus, according to the participants in the label generation and assignment studies, the two 2D spaces do *not* encode a fundamental stereotype dimension other than A and B.

Discussion

Study 6 examined if the 2D ABC model generalizes from comparing all or representative samples of groups to individually tailored samples of groups. In real life people stereotypically compare self-selected rather than representative or complete samples of groups. In doing so, they may compare different group samples on different stereotype dimensions. If this had been the case, then the MDS algorithm (Young, Takane, & Lewyckyj, 1978; see also Borg & Groenen, 2005) applied throughout this article would have resulted in a statistically poor-fitting and thus uninterpretable 2D group space.

Quite to the contrary, the results of Study 6 showed that people stereotypically compare different selections of groups on the same dimensions, namely agency/socioeconomic success and conservative–progressive beliefs, and that communion again emerges as centrality in the well-fitting and thus interpretable 2D stereotype space spanned by A and B. Moreover, Study 6 fully replicated Study 5, providing further empirical support for our conclusions that the 2D ABC model is valid across different approaches to sampling groups (from memory vs. text corpora), comparing groups (globally vs. character- vs. personal encounter-based), and rating groups (as viewed by society vs. single persons). Finally, as in Study 5, we applied a data-driven approach to scaling the groups' 2D space (see left side of Figures 6 and 7), and to interpreting this space. Results show that there is no other 2D model of stereotype content that we have overlooked (see Table 8 and the right side of Figures 6 and 7; see also the Discussion of Study 5).

Before we will elaborate on these results in detail, we put the model to a test that goes beyond the description of a similarity structure and shows that the groups' positions on the two fundamental stereotype content dimensions A and B have downstream consequences.

[12] Separate property fitting analyses for the 1D–3D spaces extracted from the similarity-, character-, and personal encounter-based mean intergroup distances yielded almost identical results to the property fitting analyses for the 1D–3D spaces extracted from the mean intergroup distances collapsed across these three spatial arrangement instructions, as shown in the online supplementary material in Tables osm.9 (minimalistic groups) and osm.10 (naturalistic groups).

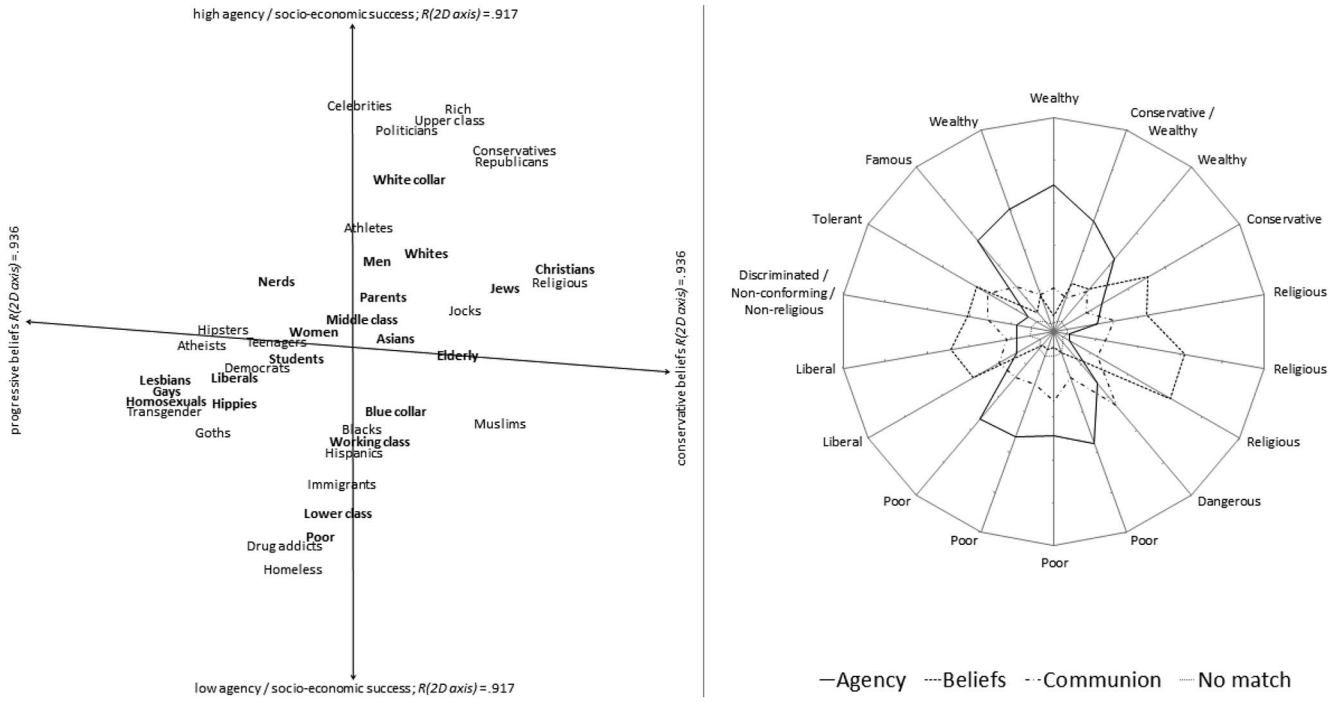

Figure 6. Study 6: The left side illustrates that the minimalist groups' 2D stereotype space is made up of the axes agency/socioeconomic success and conservative–progressive beliefs and the centrally located pole communion (21 most/least communal groups = bold/not bold). The right side plots the most frequent labels for the nine pairs of reversed axes of this space, and the percentage of labels for these axes that were assigned to A, B, C, and "no match." All axes mainly reflect A or B at the angle where these two run through the space (see property fitting results on the left side).

Study 7

Stereotypes about groups are an effective and efficient tool to plan social interactions (Gilbert & Hixon, 1991; Hamilton, Sherman, & Ruvolo, 1990; Pattyn et al., 2013; Sherman, Lee, Bessenoff, & Frost, 1998; Unkelbach, Forgas, & Denson, 2008). If the 2D ABC model of stereotype content is valid, then people should rely predominantly on A and B to make predictions about the states and dynamics of their social environment. In Study 7, we explored how people make predictions about the likelihood of members of different social groups being in the same place at the same time (judgments of time–space proximity), and about the likelihood of members of different groups being friends with one another (judgments of interpersonal liking). Particularly the latter constitutes a critical test of the relatively greater weight of A and B compared with C. Interpersonal liking is clearly a judgment about a communal aspect of interpersonal behavior, allowing the assumptions that it is particularly likable people who like each other. If, however, interpersonal liking is seen as a function of A and/or B this would translate into participants' estimation that people like each other when they share power, status, and dominance, and/or conservative or progressive beliefs.

Method

Participants and stimuli. We paid 214 MTurkers (84 women, 130 men; $M = 34.72$ years, $SD = 11.55$) \$0.6 to "sort 40 social groups into five categories of social groups." As in Studies 2 and 4, people were presented with a random sample of 40 of the 80 social groups that had been named by at least 10% of all people who had named U.S.-representative social groups in Study 1. Participants were randomly assigned to one of two conditions. In one condition, participants sorted groups according to their *time–space proximity*, while in the other condition, participants sorted social groups according to their *interpersonal liking*.

Procedure. On the first screen slide, people read: ". . . please drag-and-drop each of these 40 social groups into one of the category boxes presented below." Thereafter, in the time–space proximity condition, 112 participants read on: "Members of social groups that are likely to be in the same place at the same time should be placed into the same category box. Members of social groups that are unlikely to be in the same place at the same time should be placed into different category boxes." In the interpersonal liking condition, 102 participants read on: "Members of social groups that like one another should be placed into the same category box. Members of social groups that do not like one another should be placed into different category boxes." Below, on the left side of the screen were 40 groups below one another in random order. On the rights side were five unlabeled category boxes below one another. To finish the task, participants were instructed to sort all 40 groups into any number between two and all five category boxes.

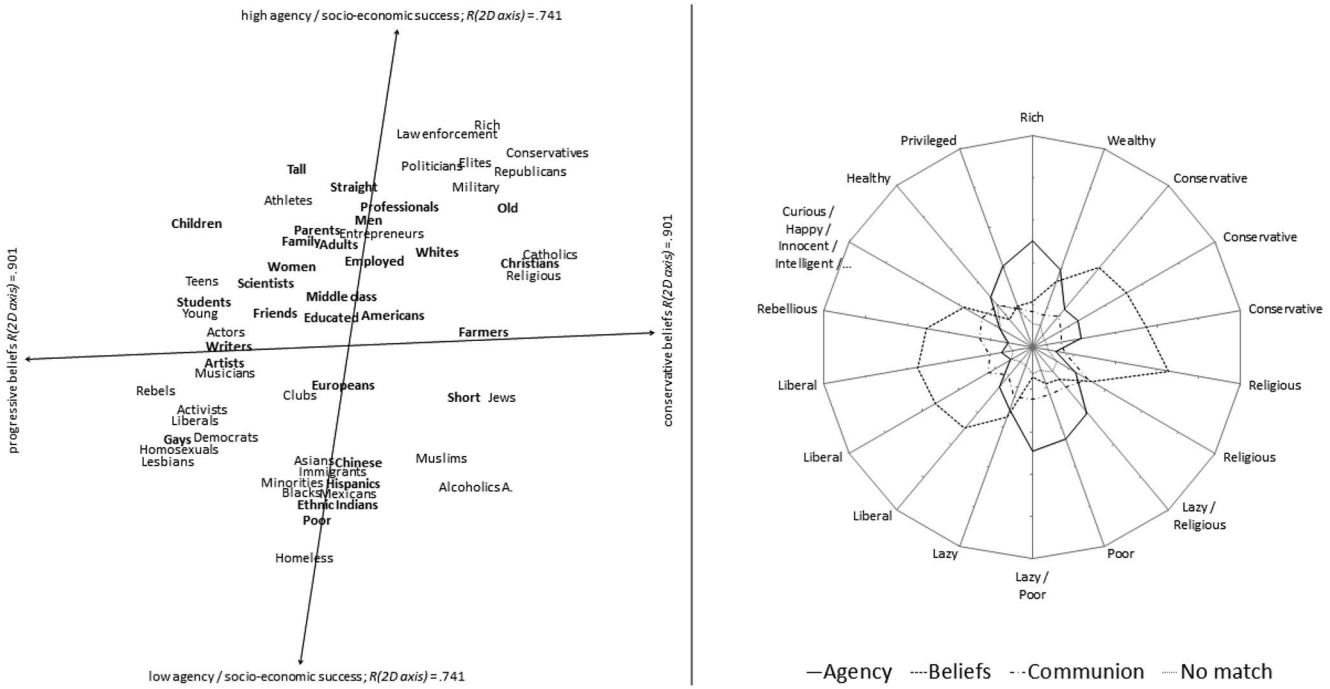

Figure 7. Study 6: The left side illustrates that the naturalistic groups' 2D stereotype space is made up of the axes agency/socioeconomic success and conservative–progressive beliefs and the centrally located pole communion (most/least communal groups = bold/not bold). The right side plots the most frequent labels for the nine pairs of reversed axes of this space, and the percentage of labels for these axes that were assigned to A, B, C, and "no match." All axes mainly reflect A or B at the angle where these two run through the space (see property fitting results on the left side).

Results

First, we calculated the likelihood of being sorted into the same category separately for each unique pair of social groups (n = 3,160) across all participants who had sorted that pair, and separately for the 112 and 102 people who had categorized the groups based on time–space proximity (M = 27.64 participants per pair, SD = 3.89) and interpersonal liking (M = 25.15 participants per pair, SD = 3.69), respectively. Second, based on the data from Study 1c, we calculated the absolute A, B, and C difference for all unique pairs of groups (e.g., the agency of students and scientists is 41.62 and 67.75 on a 0–100 scale, respectively, and thus their absolute agency difference is 26.13). And third, to test whether people sorted groups with similar and different A, B, and C ratings into the same and different time-space proximity and interpersonal liking categories, respectively, we ran two multiple linear regressions with the 3,160 group pairs' likelihood of being sorted into the same time-space proximity and interpersonal liking category as criterion and with the group pairs' absolute A, B, and C differences as predictors.

The results displayed in Table 9 (all regression weights were highly significant due to the high number of degrees of freedom) showed that similarity in both A and B was used as an indicator of a high likelihood of time–space proximity in real life. Thus, participants estimated that groups that share a position on either of the two dimensions are also particularly likely to meet in real life. Importantly, both effects were additive in that they explained unique variance. C did have a comparably small impact on whether groups were seen as particularly likely to be in the same place at the same time. Results for the second criterion, the likelihood that persons from the 80 groups were judged to like one another, mirrored the results for time–space proximity: similarity in A and B strongly and independently influenced whether group members were judged to like one another (i.e., the more similar two social groups are on these two dimension, the more likely their members are judged to like one another). Similarity in C again had a much smaller influence on judgments of interpersonal liking.

Discussion

As predicted, participants predominantly relied on stereotypic A and B, and substantially less on C to predict the time–space proximity and interpersonal liking of members of 80 representatively sampled social groups. More precisely, participants thought that members of different groups can likely be found in the same place at the same time if the groups are similar in A and B. Further, they thought that members of different groups are likely to like one another depending on the groups' similarity regarding A and B as well.

This is consistent with the similarity breeds liking literature, which shows that similarity in values and attitudes (i.e., B) is particularly relevant for interpersonal liking (e.g., Byrne, 1971; Collisson & Howell, 2014). Likewise, assortative mating (i.e., a mating pattern in which similar partners mate more frequently with

Table 9
Study 7: Pairwise Time–Space Proximity and Interpersonal Liking Simultaneously Predicted by Pairwise Absolute Rating Distance on Agency/Socioeconomic Success, Conservative–Progressive Beliefs, and Communion

Predictors	β	*t*	*p* <=	*r*	pr^2	sp^2
Criterion: time–space proximity						
Distance on Agency (A)	−.431	−28.74	.001	−.421	−.456	−.426
Distance on Beliefs (B)	−.361	−24.34	.001	−.341	−.398	−.360
Distance on Communion (C)	−.038	−2.56	.01	−.105	−.046	−.038
Criterion: interpersonal liking						
Distance on Agency (A)	−.365	−23.39	.001	−.359	−.384	−.361
Distance on Beliefs (B)	−.344	−22.30	.001	−.327	−.369	−.344
Distance on Communion (C)	−.059	−3.78	.001	−.116	−.067	−.058

Note. 3,160 unique U.S. social group pairs' likelihood of being sorted into the same time–space proximity or interpersonal liking category simultaneously predicted by the group pairs' absolute rating distance on A, B, and C, our candidate stereotype content dimensions; *r*, pr^2, and sp^2 denote zero order, partial, and semi-partial correlation, respectively.

one another that would be expected by chance) has been shown to occur not only along dimensions of physical traits (e.g., attractiveness, height), but also along socioeconomic status, intelligence, religious beliefs, and political ideology (Kail & Cavanaugh, 2010).

In sum, Study 7 provided further support for A and B as fundamental dimensions that people spontaneously employ to distinguish between social groups.

General Discussion

Previous research shows that people are able to employ estimates of warmth and competence to distinguish between social groups (e.g., Bergsieker et al., 2012; Cuddy et al., 2007; Fiske et al., 2002; Kervyn et al., 2013, 2015), and that a group's position on these two stereotype dimensions matters, because it has consequences for people's emotional and behavioral reactions and responses to that group (Becker & Asbrock, 2012; Cikara & Fiske, 2012; Cuddy et al., 2007). However, regardless of their undisputed importance, there is no evidence that warmth and competence are fundamental in the sense that people *spontaneously* use these two and not other stereotype dimensions to distinguish between groups. To test this, people need to be free to use any stereotype dimensions that they want to use. This article presents such a data-driven approach to the assessment of the dimensionality and content of spontaneous stereotypes about groups.

Complying with Brunswik's (1956) call for representative designs, in our studies participants freely selected not only stereotype dimensions, but also groups, because stimulus samples selected by researchers are often biased toward their theories (Fiedler, 2011). To implement this research design, we asked participants to name a fixed or self-chosen amount of types of people that today's U.S./German society categorizes into groups (Fiske et al., 2002). In six studies, other participants spatially arranged either the most frequently named groups, or the groups that appear most often in contemporary American English literature on stereotype dimensions of their free choice. According to our multidimensional scaling analyses, the spatially arranged mean distances between these groups could best be described by a 2D space. According to our property fitting analyses as well as several studies in which other participants generated and categorized labels for virtually all axes that run through the origin of this space, the single best pair of more or less orthogonal stereotype axes that underlie the space was *agency/socioeconomic success* (A; powerless-powerful, dominated-dominant, low status-high status, poor-wealthy, unconfident-confident, and unassertive-competitive) and *conservative–progressive beliefs* (B; traditional-modern, religious-science-oriented, conventional-alternative, and conservative-liberal). In other words, variation in spontaneous stereotype content about groups could best be described by A and B. We obtained evidence for this 2D model from U.S. online and German lab samples and from various data-driven approaches to measure spontaneous representations of groups: sequential similarity judgment, simultaneous spatial arrangement with respect to global, character-, and personal encounter-based similarity, as well as spatial arrangement with prior labeling of two similarity axes. Moreover, our data suggests that of the two fundamental dimensions A is primary, and B is secondary. Across studies, agency/socioeconomic success was regularly the best interpretation of a 1D group distances scaling solution (i.e., a 1D stereotype space), and was most often named as most important for distinguishing between the groups.

Speculating Why People Compare Groups in Terms of Their Stereotypic A and B, not C

Why agency/socioeconomic success? Social hierarchies are millions of years old and even today ubiquitous, not just in adults, but also in children and species other than humans. Social hierarchies satisfy people's need for structure, stability, identity, and safety (Jost & van der Toorn, 2012), and satisfy people's need to maintain a shared reality that coordinates social interaction for the common good (Magee & Galinsky, 2008). Specifically, keeping track of social rank is instructive about who needs to be concerned with whose perspective and feelings (Galinsky, Magee, Inesi, & Gruenfeld, 2006; van Kleef, De Dreu, Pietroni, & Manstead, 2006), about who is constrained and who is free to do and speak their mind (Berdahl & Martorana, 2006), about who speaks and who listens (DePaulo & Friedman, 1998), and about who tells whom what to do.

Perhaps more importantly, people keep track of social rank, because doing so is critical for their individual good. Social groups that are higher in rank hold the key to what people need and

want—be it health (e.g., doctors), wealth (e.g., managers), entitlement (e.g., lawyers), insight (e.g., teachers), or voice (e.g., politicians). Thus, to reach their goals, people must keep track of and connect well with groups of higher rank. Also, people want to rise in social rank to have greater access to what they need and want, and to increase their influence on other groups. In a nutshell, distinguishing between groups based on their A might be essential for feeling secure, for managing cooperation, for reaching goals, and for climbing up the social ladder by approaching, attaching to, and blending in with groups of higher rank (Magee & Galinsky, 2008).

Previous research on fundamental dimensions of social perception has come to similar conclusions: A or competence (which is correlated with A, but distinct) is an integral part of virtually any such model, be it under these labels or under labels like instrumentality (Parsons & Bales, 1955), intellectual desirability (Rosenberg et al., 1968), self-profitability (Peeters, 1992), or self-enhancement (Schwartz, 1994). Ultimately, A is considered to be functional both evolutionarily and culturally (Anderson, Hildreth, & Howland, 2015; Cheng, Tracy, Foulsham, Kingstone, & Henrich, 2013; Fiske et al., 2007).

Why conservative–progressive beliefs? Less consistent with previous research is our finding that the second fundamental dimension on which people spontaneously distinguish social groups is whether they are engines of change or preservers of the status quo—that is, their position on the dimension of *progressive-conservative beliefs*. We speculate that knowing whether the ideological beliefs of a group are conservative or progressive comes with a lot of valuable insights about the ways in which that group intends to use the influence that it has on other groups, and about the ways in which members of that group think, feel, and behave. In line with the idea that humans are intention detectors and often prioritize intentions over outcomes (e.g., Ames & Fiske, 2013, 2015), B may inform individuals about the general intentions of groups at a societal level. Generally speaking, conservative groups (e.g., Christians, Republicans, elderly, and the military) want things to be uniform and stay the way they are, and thus they emphasize religion, traditions, conventions, and conformity. Interacting with conservatives provides people with feelings of stability, predictability, control, safety, comfort, and belonging (for a review, see Jost, Federico, & Napier, 2009; Schwartz & Bilsky, 1987, 1990). In contrast, progressive groups (e.g., techies, actors, hipsters, and homosexuals) want things to change and diversify, and thus they emphasize freedom, autonomy, creativity, innovation, (technological, economic, legal, etc.), modern subculture (art, music, literature, etc.) and media, and alternative views and lifestyles. Interacting with progressives provides people with feelings of curiosity, stimulation, expansion, entertainment, and distinctiveness (Leonardelli, Pickett, & Brewer, 2010; Schwartz, 1994). Thus, keeping track of the ideological beliefs of groups might serve at least two functions: It helps people to anticipate and handle the form and content of social interactions (e.g., politely agreeing with somebody vs. dressing up in an outrageous way), and it enables people to strike a balance between cognitive, emotional, and behavioral exploitation (conservative groups) and exploration (progressive groups; Jost et al., 2009).

Managing the trade-off between exploiting available resources of certain quality and quantity and exploring alternative resources of uncertain quality and quantity is fundamental to self-regulatory success (Cohen, McClure, & Yu, 2007; Hills, Todd, & Goldstone, 2010; Inzlicht, Schmeichel, & Macrae, 2014) both culturally (i.e., in the last couple of millennia) and evolutionarily (i.e., since the beginning of life). That is, adults, children, other primates, and many other beings have always had to choose between current and alternative habitats, shelters, occupations, foods, mates, and so forth, and these choices have always been important to survive and thrive. Based on our results it could be argued that in today's society it is the conservative and the progressive groups that provide access to current, certain and alternative, uncertain resources, respectively. Therefore, to successfully manage the ancient and ubiquitous exploitation-exploration trade-off, today's citizens might mentally organize groups along the stereotype dimension of conservative–progressive beliefs.

To further explore whether B is in fact a fundamental stereotype dimension that informs individuals about group-specific opportunities for exploitation and exploration, we asked additional participants to rate the groups that we examined in Studies 1–4 on seven stereotype dimensions that map onto exploitation-exploration.[13] With one exception ("prevention-promotion"), these stereotype dimensions ("familiarity-novelty," "safety-freedom," "comfort-stimulation," "loyalty-autonomy," "preservation-change," and "uniformity-diversity") correlated strongly with B[14] (mean $r = .68$, all $ps < .001$)—in fact as strongly as the correlations between the four stereotype dimensions that form B (mean $r = .70$, all $ps < .001$).

Further, we combined the four stereotype dimensions that form B with the six exploitation–exploration stereotype dimensions. *Exploitation–exploration* and B were equally suitable for disambiguating the U.S. groups' 2D stereotype spaces reported in this article, mean *R(2D axis)* = .89 and .90 (*SD*s = .04 and .04), respectively, $ps < .001$. However, because the participant-generated labels for the horizontal and vertical axis of the 2D arrangement board in Study 4 and the participant-generated labels for the nine pairs of reversed axes of the 2D stereotype spaces scaled in Studies 5 and 6 mainly reflected B ("religious," "traditional," "conservative," "nonreligious," "nontraditional," and "liberal," see Figures 4–7) and not exploitation–exploration, it seems that on the manifest level individuals spontaneously use B to distinguish between groups. However, striking a balance between exploitation and exploration might be the latent regulatory function that distinguishing between conservative and progressive groups tries to serve.

[13] We paid 166 participants (67 females, 99 males; $M = 42.80$ years, $SD = 7.91$) \$0.75 to rate the 80 Studies 1–4 U.S. groups on one of seven stereotype dimensions that map onto exploitation-exploration: "familiarity–novelty," "safety–freedom," "comfort–stimulation," "loyalty–autonomy," "preservation–change," "uniformity–diversity," and "prevention–promotion." Participants read: "To what extent do these 80 groups stand for . . . [e.g., safety vs. freedom]. We are not interested in your personal view, but in how you think these 80 groups are viewed by today's society." Then, as in all studies reported here, they used 0–100 slider scales to rate the groups one atop the other in random order. There were between 21 and 27 raters per stereotype dimension, and as in the previous studies, raters' agreement about the groups was very high, all $ICC(2,k) > = .79$, (McGraw & Wong, 1996).

[14] The exploitation–exploration stereotype dimensions "familiarity–novelty," "safety–freedom," "comfort–stimulation," "loyalty–autonomy," "preservation–change," and "uniformity–diversity" correlated strongly with B (mean $r = .68$, all $ps < .001$), but not with A (mean $r = .36$, four out of six $ps < .001$) and C (mean $r = .24$, none of the $ps < .001$).

Why not communion? Lastly, our data-driven model deviates from existing theoretical approaches in the role of *communion* or warmth. Classical models construe C as an orthogonal stereotype dimension (Cuddy et al., 2007; Fiske et al., 2002) that has processing priority over all other information (Abele & Bruckmüller, 2011; Brambilla, Rusconi, Sacchi, & Cherubini, 2011; Wojciszke, 1994; Wojciszke, Bazinska, & Jaworski, 1998). Following the functional logic developed above, one could of course ask why individuals should pay attention to whether the intentions of a group are communal or not, if that group does not have the A to implement its intentions (e.g., children, homeless, drug users, and agnostics; for a previously posed similar question, see Fiske et al., 2002). Consistent with the order of priority suggested by this question, our data showed that C is an emergent quality that is not independent from other stereotype dimensions but follows from A. Groups that are seen as particularly unagentic (e.g., homeless, welfare recipients) or overly agentic (e.g., rich, managers) are also seen as least trustworthy, sincere, likable, warm, benevolent, and altruistic. Perhaps those groups are seen as contributing too little to society and profiting too much from society, respectively. As communion emerges at the center of the A dimension, it can be reconciled with the 2D AB model of stereotype content.

Importantly, this finding is not an artifact of asking for spatially arranged dissimilarity judgments and ratings on two self-labeled stereotype dimensions. Even if we completely ignore the multidimensionally scaled and property fitted dissimilarity data presented in Studies 1–6, and consider only on the ratings of A and C, it becomes apparent that these dimensions are not independent. Groups' C is the higher the more average their A is: $r = .40$, $p < .001$ (U.S. groups in Studies 1–4), $r = .44$, $p < .001$ (German groups in Study 3), $r = .51$, $p < .001$ (minimalist U.S. groups in Studies 5 and 6; $r = .35$, $p < .01$ (naturalistic U.S. groups in Studies 5/6).[15] This new look on communion as average agency is entirely consistent with the abundant literature that people trust and like typical, average faces and trait scores more than atypical, extreme faces and trait scores (Langlois & Roggman, 1990; Potter, Corneille, Ruys, & Rhodes, 2007; Rhodes, Halberstadt, & Brajkovich, 2001; Sofer et al., 2015).

Theoretical Implications of the ABC Model of Stereotypes About Groups

The 2D ABC model allows for a new perspective on the well-established effects of compensation between warmth and competence (Kervyn, Yzerbyt, Judd, & Nunes, 2009; Yzerbyt, Kervyn, & Judd, 2008). Although warmth and competence are conceptualized as orthogonal dimensions in Fiske et al. (2002) stereotype content model, individuals who are described as particularly competent are systematically inferred to be relatively cold (Kervyn, Bergsieker, & Fiske, 2012). Although none of the dimensions in the 2D ABC model was best described as competence, we observed a similar relation between C (high overlap with warmth) and A (related to competence) with one important qualification. In the 2D ABC model, the compensation between A and C should only hold for the upper half of the A dimension: Moderately agentic groups are more communal than highly agentic groups because C is inferred from centrality on the A dimension. Importantly, our model makes further predictions that are in contradiction to general compensation effects. Groups less agentic than average will also be less communal. Starting from a very low position on A (e.g., drug users, homeless), an increase in a group's A toward the average will also lead to more favorable C impressions.

Given that stereotypic C (but not so much A and B) can be taken as a proxy for stereotypic valence,[16] this new perspective on stereotypic C also allows further delineations. If C emerges as average A, then after a certain point (i.e., being exactly average on A), social groups cannot be stereotyped as more communal, while they can always be stereotyped as less communal, because there is no limit to being more extreme in terms of A. This is consistent with the notion that negative stimuli are stronger (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001), more dominant/contagious (Rozin & Royzman, 2001), and more mobilizing (Taylor, 1991) than positive stimuli. Finally, maximal C and thus the highest positive valence at average agency is also consistent with the *density hypothesis* (Unkelbach, 2012; Unkelbach, Fiedler, Bayer, Stegmüller, & Danner, 2008), that is, the notion that positive stimuli are more similar to one another than negative stimuli (see also Alves, Koch, & Unkelbach, in press; Alves, Unkelbach, Burghardt, Koch, Krüger, & Becker, 2015; Koch et al., in press). More precisely, if increasingly communal groups are increasingly close to the center of the A dimension, then they must be increasingly similar to one another, just because by necessity they are also increasingly close to one another. If so, then the groups' C should be related to their average similarity to all other social groups. This was indeed the case: $r = .38$, $p < .001$ (Study 1); $r = .59$, $p < .001$ (Study 2); $r = .79$, $p < .001$ (Study 3); $r = .53$, $p < .001$ (Study 4); $r = .64$ and $r = .67$, both ps $< .001$ (Study 5; minimalist and naturalistic groups, respectively); and $r = .53$ and $r = .42$, both ps $< .001$ (Study 5; minimalist and naturalistic groups, respectively. In Tolstoy's (1873–1877/2001)- terms: Communal social groups are all alike (i.e., they are all average on A); but every noncommunal social group is noncommunal in its own way (being either higher or lower than average on A).

Limitations and Future Directions

The studies described here leave open whether there are spontaneous/fundamental stereotype content dimensions other than agency/socioeconomic success and conservative–progressive beliefs. With the exception of Study 1, in all studies reported in this article, the 2D spatial arrangement board (Hout et al., 2013) that

[15] Consistent with our finding that C emerges from A but not B, the social groups' C is not the higher the more average their B is, $r = -.17$, $p = .13$ (U.S. groups in Studies 1–4), $r = .06$, $p = .60$ (German groups in Study 3), $r = -.10$, $p = .55$ (minimalist U.S. groups in Studies 5 and 6), and $r = -.28$, and $p < .05$ (naturalistic U.S. groups in Studies 5 and 6; this correlation is the only exception, and it is rather weak).

[16] We paid 25 MTurkers (16 females, nine males; $M = 42.80$ years, $SD = 7.91$) \$1 to rate the valence ("worse–better") of the 80 U.S. groups examined in Studies 1–4. Valence correlated with C, $r = .78$, $p < .001$, but neither with A, $r = -.01$, $p = .90$, nor with B, $r = .07$, $p = .55$. Note that this does not contradict the linear relation between valence and A as found by Abele and Wojciszke (2007) or Suitner and Maas (2008). These and other studies lack extremely agentic stimuli (e.g., "aggressive," "reckless," and "conceited" rather than just "assertive," "brave," and "confident"), and thus they found a linear relation. In our stimulus sample, there are extremely agentic groups (e.g., "rich," "celebrities," "elites," "upper class," "managers," "politicians," "lawyers"), and thus we find a quadratic relation between valence and A (more precisely, a linear relation between valence and averageness on A), $r = .31$, $p < .01$.

people used prompted them to spontaneously select no more than two unrelated stereotype content dimensions. Thus, it is possible that there is consensus about a third, fourth, fifth, and so forth spontaneous/fundamental stereotype content dimension that our research designs did not reveal. The third dimension might actually be communion, as communion was (not highly, but) to some extent suitable as a third independent dimension in Studies 1, 4, and, in part, 5. Although the question of whether there are more than the two spontaneous/fundamental stereotype content dimensions is informative, insights about additional dimensions would not speak against our assertion that the two most spontaneous/fundamental ones are A and B. C as the third, fourth, fifth, and so forth fundamental stereotype content dimension would also be compatible with C as average A, as was found in Studies 1–6.

Despite the highly consistent results, our studies speak to the relatively abstract question of how individuals distinguish between all societal groups. In motivating our research we have labeled this approach to stereotyping the "lay sociologist" perspective and related it to previous research that (at least at the core of its empirical contribution) has followed a similar aim (e.g., Cuddy et al., 2007; Fiske et al., 2002). Stereotypes might, however, not only guide how people distinguish between all societal groups, but may also serve as knowledge structures that individuals recruit in social interactions with members of proximal groups (the "relational" perspective; e.g., Cambon, Yzerbyt, & Yakimova, 2015). It may be that stereotypical knowledge about the communion/warmth of such proximal groups receives relatively greater processing priority in social interactions compared with people's perception of more remote groups.

More precisely, encountering members of proximal groups may elicit an affective, evaluative response that leads people to spontaneously construe these groups in terms of their perceived C rather than A and/or B. The results of Studies 5 and 6 speak against this idea, as participants spontaneously employed A and B rather than C to distinguish between encounters with members of all societal groups. However, maybe participants in Study 5 and 6 did not identify strongly and/or did not strongly oppose identifying with many of these groups, reducing the salience of C as a dimension. To test if C is most important for distinguishing between encounters with members of predominantly proximal groups, in future studies people should spatially arrange mostly proximal groups that they strongly identify with and/or strongly oppose identifying with. Nevertheless, we argue that distinguishing between all societal groups, and possibly also between proximal groups, in terms of their perceived A and B is functional and important, too. As stated above, groups' A and B are informative about opportunities for reaching goals and climbing up the social ladder (A) and opportunities for exploitation versus exploration (B).

Even without the relational aspect, individuals may think differently about individuals than about social groups. Future studies might thus consider prompting people to name representatives of groups (e.g., Natalie Portman for actors and Pope Francis for Catholics), and to then spatially arrange these representatives rather than the groups they represent. Alternatively, people might spatially arrange individuals that do not markedly represent any particular group(s). Such individualized processing may bring communal information to the forefront so that participants spontaneously judge the group representatives/individuals primarily along the line of how trustworthy and friendly they perceive them to be. This would be one possibility to reconcile our group-based 2D ABC model with the finding that communion enjoys a privileged position in processing of information about individuals (e.g., Abele & Bruckmüller, 2011; Willis & Todorov, 2006; Wojciszke & Abele, 2008).

Given that data-driven approaches to modeling face perception show that dominance (or agency/socioeconomic success), youthfulness versus agedness (one could argue that people with a youthful and aged face are likely to hold progressive and conservative beliefs, respectively), and trustworthiness (or communion) are fundamental dimensions (Oosterhof & Todorov, 2008; Sutherland et al., 2013), it seems promising to explore the extent to which the space of facial stereotypes also follows the 2D ABC pattern developed in this article. There is already initial evidence that faces with more average features are perceived as more communal (Sofer et al., 2015; Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015). However, it is also conceivable that faces prompt a more individualized social information processing than abstract group labels, so that C is an independent dimension that is given more weight than A and B (Sutherland et al., 2013). In any case, the 2D ABC model suggested here based on a bottom-up, data-driven approach must now be tested in a top-down, theory-driven research program.

Our results consistently show very high overlap (up to $R >= .90$) between the axis rotated around the origin of the social groups' 1D–3D spaces and the independently gathered ratings of the groups on candidate stereotype content dimensions. Although this overlap is almost suspiciously high (cf. Vul, Harris, Winkielman, & Pashler, 2009), it should be stressed that we correlated data on a very high level of aggregation. Specifically, we correlated dissimilarity averaged across individuals with stereotypic A, B, and C averaged across individuals, which removed all variance due to interindividual differences in judging the dissimilarities between the social groups and their A, B, and C. Thus, our data reflect correlations of group-level averages (social groups as cases), and not averaged individual-level correlations (participants as cases). Thus, our group-level effect sizes do not allow conclusions about individual-level effect sizes (Brand & Bradley, 2012; Brand, Bradley, Best, & Stoica, 2008; Monin & Oppenheimer, 2005). This does *not* threaten the validity of our 2D ABC model of stereotype content about groups, as stereotypes are defined as group-level effects (i.e., groups "as viewed by society;" Fiske et al., 2002, p. 884).

Finally, just like the SCM (e.g., Fiske et al., 2002), the 2D ABC model does not address how people's group identities influence stereotypes about groups (Smith, 1993; Mackie, Smith, & Ray, 2008). Obviously, in-group versus out-group memberships must influence stereotype content. That is, individual (or intergroup) differences in group stereotypes are lost in averaging across raters. The model therefore addresses stereotypes as shared knowledge structures. Nevertheless, individual (or intergroup) differences in group stereotypes are a fascinating topic for future research. For example, it could be that communion remains a centrally located polar dimension also at the individual level (existing data suggest that this is so; Imhoff & Koch, 2016). It may be that even raters who are extreme on A or B still see groups average on these dimensions as most trustworthy. Alternatively, it might be that for these raters C transforms into an axial dimension that is more or

less identical with the axial dimension on which they are extreme, with high C being located where the raters' groups are extreme. If the latter scenario holds true, then, for example, artists should perceive progressives and conservatives as high and low on C, whereas groups that differ in A should not differ in C for artists. These and other empirical questions are interesting and important avenues for future research on the 2D ABC model of stereotype content about groups.

Conclusion

We presented a data-rather than theory-driven answer to the nature and order of the stereotype content dimensions that people spontaneously employ to distinguish between social groups. Our analyses indicate that people mentally organize groups primarily based on their stereotypic *agency/socioeconomic success* (A), and secondarily based on their stereotypic *conservative–progressive beliefs* (B). Further, social groups that are thought to be average on A are inferred to be communal, whereas social groups that are thought to be extreme (high or low) on A are inferred to be as noncommunal (C), resulting in a 2D *ABC model* of stereotype content.

References

Aaker, J., Garbinsky, E. N., & Vohs, K. D. (2012). Cultivating admiration in brands: Warmth, competence, and landing in the golden quadrant. *Journal of Consumer Psychology, 22,* 191–194. http://dx.doi.org/10.1016/j.jcps.2011.11.012

Abele, A. E., & Bruckmüller, S. (2011). The bigger one of the "Big Two?" Preferential processing of communal information. *Journal of Experimental Social Psychology, 47,* 935–948. http://dx.doi.org/10.1016/j.jesp.2011.03.028

Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology, 93,* 751–763. http://dx.doi.org/10.1037/0022-3514.93.5.751

Abele, A. E., & Wojciszke, B. (2014). Communal and agentic content in social cognition: A dual perspective model. *Advances in Experimental Social Psychology, 50,* 195–255. http://dx.doi.org/10.1016/B978-0-12-800284-1.00004-7

Agerström, J., Björklund, F., Carlsson, R., & Rooth, D. (2012). Warm and competent Hassan = Cold and incompetent Eric: The harsh equation of real-life hiring discrimination. *Basic and Applied Social Psychology, 34,* 359–366. http://dx.doi.org/10.1080/01973533.2012.693438

Akpinar, E., & Berger, J. (2015). Drivers of cultural success: The case of sensory metaphors. *Journal of Personality and Social Psychology, 109,* 20–34. http://dx.doi.org/10.1037/pspa0000025

Alves, H., Koch, A., & Unkelbach, C. (2016). My friends are all alike – the relation between liking and perceived similarity in person perception. *Journal of Experimental Social Psychology, 62,* 103–117. http://dx.doi.org/10.1016/j.jesp.2015.10.011

Alves, H., Unkelbach, C., Burghardt, J., Koch, A. S., Krüger, T., & Becker, V. D. (2015). A density explanation of valence asymmetries in recognition memory. *Memory & Cognition, 43,* 896–909. http://dx.doi.org/10.3758/s13421-015-0515-5

Ames, D. L., & Fiske, S. T. (2013). Intentional harms are worse, even when they're not. *Psychological Science, 24,* 1755–1762. http://dx.doi.org/10.1177/0956797613480507

Ames, D. L., & Fiske, S. T. (2015). Perceived intent motivates people to magnify observed harms. *Proceedings of the National Academy of Sciences of the United States of America, 112,* 3599–3605.

Anderson, C., Hildreth, J. A. D., & Howland, L. (2015). Is the desire for status a fundamental human motive? A review of the empirical literature. *Psychological Bulletin, 141,* 574–601. http://dx.doi.org/10.1037/a0038781

Ashton-James, C. E., Richardson, D. C., Williams, A. C., Bianchi-Berthouze, N., & Dekker, P. H. (2014). Impact of pain behaviors on evaluations of warmth and competence. *Pain, 155,* 2656–2661. http://dx.doi.org/10.1016/j.pain.2014.09.031

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype-activation on action. *Journal of Personality and Social Psychology, 71,* 230–244. http://dx.doi.org/10.1037/0022-3514.71.2.230

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology, 5,* 323–370. http://dx.doi.org/10.1037/1089-2680.5.4.323

Becker, J. C., & Asbrock, F. (2012). What triggers helping versus harming of ambivalent groups? Effects of the relative salience of warmth versus competence. *Journal of Experimental Social Psychology, 48,* 19–27. http://dx.doi.org/10.1016/j.jesp.2011.06.015

Berdahl, J. L., & Martorana, P. (2006). Effects of power on emotion and expression during a controversial group discussion. *European Journal of Social Psychology, 36,* 497–509. http://dx.doi.org/10.1002/ejsp.354

Bergsieker, H. B., Leslie, L. M., Constantine, V. S., & Fiske, S. T. (2012). Stereotyping by omission: Eliminate the negative, accentuate the positive. *Journal of Personality and Social Psychology, 102,* 1214–1238. http://dx.doi.org/10.1037/a0027717

Berinsky, A. J., Huber, G. A., & Lenz, G. S. (2012). Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. *Political Analysis, 20,* 351–368. http://dx.doi.org/10.1093/pan/mpr057

Borg, I., & Groenen, P. (2005). *Modern multidimensional scaling: Theory and applications*. New York, NY: Springer-Verlag.

Brambilla, M., Rusconi, P., Sacchi, S., & Cherubini, P. (2011). Looking for honesty: The primary role of morality (vs. sociability and competence) in information gathering. *European Journal of Social Psychology, 41,* 135–143. http://dx.doi.org/10.1002/ejsp.744

Brand, A., & Bradley, M. T. (2012). More voodoo correlations: When average-based measures inflate correlations. *Journal of General Psychology, 139,* 260–272. http://dx.doi.org/10.1080/00221309.2012.703711

Brand, A., Bradley, M. T., Best, L. A., & Stoica, G. (2008). Accuracy of effect size estimates from published psychological research. *Perceptual and Motor Skills, 106,* 645–649. http://dx.doi.org/10.2466/pms.106.2.645-649

Brewer, M. B. (1988). A dual process model of impression formation. In R. S. Wyer & T. K. Srull (Eds.), *Advances in social cognition* (Vol. 1, pp. 1–36). Hillsdale, NJ: Erlbaum.

Brewer, M. D. (2005). The rise of partisanship and the expansion of partisan conflict within the American electorate. *Political Research Quarterly, 58,* 219–229. http://dx.doi.org/10.1177/106591290505800203

Broverman, I. K., Vogel, S. R., Broverman, D. M., Clarkson, F. E., & Rosenkrantz, P. S. (1972). Sex-role stereotypes: A current appraisal. *Journal of Social Issues, 28,* 59–78. http://dx.doi.org/10.1111/j.1540-4560.1972.tb00018.x

Bruner, J. S. (1957). Going beyond the information given. In J. S. Bruner, E. Brunswik, L. Festinger, F. Heider, K. F. Muenzinger, C. E. Osgood, & D. Rapaport (Eds.), *Contemporary approaches to cognition* (pp. 41–69). Cambridge, MA: Harvard University Press.

Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review, 62,* 193–217. http://dx.doi.org/10.1037/h0047470

Brunswik, E. (1956). *Perception and the representative design of psychological experiments*. Berkeley, CA: University of California Press.

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Per-*

spectives on Psychological Science, 6, 3–5. http://dx.doi.org/10.1177/1745691610393980

Byrne, D. (1971). *The attraction paradigm*. New York, NY: Academic Press.

Cambon, L., Yzerbyt, V., & Yakimova, S. (2015). Compensation in intergroup relations: An investigation of its structural and strategic foundations. *British Journal of Social Psychology, 54,* 140–158. http://dx.doi.org/10.1111/bjso.12067

Carrier, A., Louvet, E., Chauvin, B., & Rohmer, O. (2014). The primacy of agency over competence in status perception. *Social Psychology, 45,* 347–356. http://dx.doi.org/10.1027/1864-9335/a000176

Carroll, J. D., & Wish, M. (1974). Multidimensional perceptual models and measurement methods. *Handbook of Perception, 2,* 391–447.

Casler, K., Bickel, L., & Hackett, E. (2013). Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior, 29,* 2156–2160. http://dx.doi.org/10.1016/j.chb.2013.05.009

Chang, J. J., & Carroll, J. D. (1969). *How to use PROFIT, a computer program for property fitting by optimizing nonlinear or linear correlation*. Murray Hill, NJ: Bell Telephone Laboratories.

Cheng, J. T., Tracy, J. L., Foulsham, T., Kingstone, A., & Henrich, J. (2013). Two ways to the top: Evidence that dominance and prestige are distinct yet viable avenues to social rank and influence. *Journal of Personality and Social Psychology, 104,* 103–125.

Cikara, M., & Fiske, S. T. (2012). Stereotypes and Schadenfreude: Behavioral and physiological markers of pleasure at others' misfortunes. *Social Psychological & Personality Science, 3,* 63–71. http://dx.doi.org/10.1177/1948550611409245

Clow, K. A., & Leach, A. M. (2015). Stigma and wrongful conviction: All exonerees are not perceived equal. *Psychology, Crime & Law, 21,* 172–185. http://dx.doi.org/10.1080/1068316X.2014.951645

Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences, 362,* 933–942. http://dx.doi.org/10.1098/rstb.2007.2098

Collisson, B., & Howell, J. L. (2014). The liking-similarity effect: Perceptions of similarity as a function of liking. *The Journal of Social Psychology, 154,* 384–400. http://dx.doi.org/10.1080/00224545.2014.914882

Cuddy, A. J., Fiske, S. T., & Glick, P. (2007). The BIAS map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology, 92,* 631–648. http://dx.doi.org/10.1037/0022-3514.92.4.631

Cuddy, A. J., Fiske, S. T., Kwan, V. S., Glick, P., Demoulin, S., Leyens, J. P., . . . Ziegler, R. (2009). Stereotype content model across cultures: Towards universal similarities and some differences. *British Journal of Social Psychology, 48,* 1–33. http://dx.doi.org/10.1348/014466608X314935

Davies, M. (2011). *Google books (American English) Corpus (155 billion words, 1810–2009)*. Retrieved from http://googlebooks.byu.edu

Deaux, K., & Lewis, L. L. (1984). Structure of gender stereotypes: Interrelationships among components and gender label. *Journal of Personality and Social Psychology, 46,* 991–1004. http://dx.doi.org/10.1037/0022-3514.46.5.991

DePaulo, B. M., & Friedman, H. S. (1998). Nonverbal communication. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (pp. 3–40). Boston, MA: McGraw-Hill.

Durante, F., Volpato, C., & Fiske, S. T. (2010). Using the stereotype content model to examine group depictions in fascism: An archival approach. *European Journal of Social Psychology, 40,* 465–483.

Falbo, T. (1977). Multidimensional scaling of power strategies. *Journal of Personality and Social Psychology, 35,* 537–547. http://dx.doi.org/10.1037/0022-3514.35.8.537

Farias, A. R., Garrido, M. V., & Semin, G. R. (2013). Converging modalities ground abstract categories: The case of politics. *PLoS ONE, 8,* e60971. http://dx.doi.org/10.1371/journal.pone.0060971

Farrell, D. (1983). Exit, voice, loyalty, and neglect as responses to job dissatisfaction: A multidimensional scaling study. *Academy of Management Journal, 26,* 596–607. http://dx.doi.org/10.2307/255909

Fiedler, K. (2011). Voodoo correlations are everywhere—Not only in social neurosciences. *Perspectives on Psychological Science, 6,* 163–171. http://dx.doi.org/10.1177/1745691611400237

Fiske, S. T., Cuddy, A. J., & Glick, P. (2007). Universal dimensions of social cognition: Warmth and competence. *Trends in Cognitive Sciences, 11,* 77–83. http://dx.doi.org/10.1016/j.tics.2006.11.005

Fiske, S. T., Cuddy, A. J., 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,* 878–902. http://dx.doi.org/10.1037/0022-3514.82.6.878

Fiske, S. T., & Dupree, C. (2014). Gaining trust as well as respect in communicating to motivated audiences about science topics. *Proceedings of the National Academy of Sciences of the United States of America, 111,* 13593–13597. http://dx.doi.org/10.1073/pnas.1317505111

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. *Advances in Experimental Social Psychology, 23,* 1–74. http://dx.doi.org/10.1016/S0065-2601(08)60317-2

Fiske, S. T., & Pavelchak, M. A. (1986). Category-based versus piecemeal-based affective responses: Developments in schema-triggered affect. In R. M. Sorrentino & E. T. Higgins (Eds.), *Handbook of motivation and cognition: Foundations of social behavior* (pp. 167–203). New York, NY: Guilford Press.

Ford, T. E., & Stangor, C. (1992). The role of diagnosticity in stereotype formation: Perceiving group means and variances. *Journal of Personality and Social Psychology, 63,* 356–367. http://dx.doi.org/10.1037/0022-3514.63.3.356

Forgas, J. P. (1976). The perception of social episodes: Categorical and dimensional representations in two different social milieus. *Journal of Personality and Social Psychology, 34,* 199–209. http://dx.doi.org/10.1037/0022-3514.34.2.199

Frable, D. E. (1993). Dimensions of marginality: Distinctions among those who are different. *Personality and Social Psychology Bulletin, 19,* 370–380. http://dx.doi.org/10.1177/0146167293194002

Galinsky, A. D., Magee, J. C., Inesi, M. E., & Gruenfeld, D. H. (2006). Power and perspectives not taken. *Psychological Science, 17,* 1068–1074. http://dx.doi.org/10.1111/j.1467-9280.2006.01824.x

Ghavami, N., & Peplau, L. A. (2012). An intersectional analysis of gender and ethnic stereotypes testing three hypotheses. *Psychology of Women Quarterly, 37,* 113–127. http://dx.doi.org/10.1177/0361684312464203

Giessner, S. R., & Schubert, T. W. (2007). High in the hierarchy: How vertical location and judgments of leaders' power are interrelated. *Organizational Behavior and Human Decision Processes, 104,* 30–44. http://dx.doi.org/10.1016/j.obhdp.2006.10.001

Giguère, G. (2006). Collecting and analyzing data in multidimensional scaling experiments: A guide for psychologists using SPSS. *Tutorials in Quantitative Methods for Psychology, 2,* 27–38.

Gilbert, D. T., & Hixon, J. G. (1991). The trouble of thinking: Activation and application of stereotypic beliefs. *Journal of Personality and Social Psychology, 60,* 509–517. http://dx.doi.org/10.1037/0022-3514.60.4.509

Goldstone, R. (1994). An efficient method for obtaining similarity data. *Behavior Research Methods, Instruments, & Computers, 26,* 381–386. http://dx.doi.org/10.3758/BF03204653

Goldstone, R. L., Medin, D. L., & Halberstadt, J. (1997). Similarity in context. *Memory & Cognition, 25,* 237–255. http://dx.doi.org/10.3758/BF03201115

Haidt, J. (2012). *The righteous mind: Why good people are divided by politics and religion*. New York, NY: Pantheon Books.

Hamilton, D. L., Sherman, S. J., & Ruvolo, C. M. (1990). Stereotype-based expectancies: Effects on information processing and social behavior. *Journal of Social Issues, 46,* 35–60. http://dx.doi.org/10.1111/j.1540-4560.1990.tb01922.x

Harris, L. T., & Fiske, S. T. (2006). Dehumanizing the lowest of the low: Neuroimaging responses to extreme out-groups. *Psychological Science, 17,* 847–853. http://dx.doi.org/10.1111/j.1467-9280.2006.01793.x

Henley, N. M. (1969). A psychological study of the semantics of animal terms. *Journal of Verbal Learning and Verbal Behavior, 8,* 176–184. http://dx.doi.org/10.1016/S0022-5371(69)80058-7

Hills, T. T., Todd, P. M., & Goldstone, R. L. (2010). The central executive as a search process: Priming exploration and exploitation across domains. *Journal of Experimental Psychology: General, 139,* 590–609. http://dx.doi.org/10.1037/a0020666

Hout, M. C., Goldinger, S. D., & Ferguson, R. W. (2013). The versatility of SpAM: A fast, efficient, spatial method of data collection for multidimensional scaling. *Journal of Experimental Psychology: General, 142,* 256–281. http://dx.doi.org/10.1037/a0028860

Imhoff, R., & Koch, A. (2016). How orthogonal are the "Big Two?" On the curvilinear relation between agency and communion. *Manuscript in preparation*.

Imhoff, R., Woelki, J., Hanke, S., & Dotsch, R. (2013). Warmth and competence in your face! Visual encoding of stereotype content. *Frontiers in psychology, 4,* 386.

Inzlicht, M., Schmeichel, B. J., & Macrae, C. N. (2014). Why self-control seems (but may not be) limited. *Trends in Cognitive Sciences, 18,* 127–133. http://dx.doi.org/10.1016/j.tics.2013.12.009

Jaworska, N., & Chupetlovska-Anastasova, A. (2009). A review of multidimensional scaling (MDS) and its utility in various psychological domains. *Tutorials in Quantitative Methods for Psychology, 5,* 1–10.

Jolliffe, I. (2002). *Principal component analysis*. New York, NY: Springer.

Jones, R. A., & Ashmore, R. D. (1973). The structure of intergroup perception: Categories and dimensions in views of ethnic groups and adjectives used in stereotype research. *Journal of Personality and Social Psychology, 25,* 428–438. http://dx.doi.org/10.1037/h0034162

Jost, J. T., Federico, C. M., & Napier, J. L. (2009). Political ideology: Its structure, functions, and elective affinities. *Annual Review of Psychology, 60,* 307–337. http://dx.doi.org/10.1146/annurev.psych.60.110707.163600

Jost, J. T., Glaser, J., Kruglanski, A. W., & Sulloway, F. J. (2003). Political conservatism as motivated social cognition. *Psychological Bulletin, 129,* 339–375. http://dx.doi.org/10.1037/0033-2909.129.3.339

Jost, J. T., & van der Toorn, J. (2012). System justification theory. In P. A. M. van Lange, A. W. Kruglanski, & E. T. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 313–343). London, UK: Sage. http://dx.doi.org/10.4135/9781446249222.n42

Kail, R. V., & Cavanaugh, J. C. (2010). *Human development: A life-span view* (5th ed.). Belmont, CA: Wadsworth Cengage Learning.

Kervyn, N., Bergsieker, H. B., & Fiske, S. T. (2012). The innuendo effect: Hearing the positive but inferring the negative. *Journal of Experimental Social Psychology, 48,* 77–85. http://dx.doi.org/10.1016/j.jesp.2011.08.001

Kervyn, N., Chan, E., Malone, C., Korpusik, A., & Ybarra, O. (2014). Not all disasters are equal in the public eye: The negativity effect on warmth in brand perception. *Social Cognition, 32,* 256–275. http://dx.doi.org/10.1521/soco.2014.32.3.256

Kervyn, N., Dolderer, M., Mahieu, T., & Yzerbyt, V. Y. (2010). Atypicality and the two fundamental dimensions: Applying the negativity effect on warmth to group perception. *European Journal of Social Psychology, 40,* 484–489.

Kervyn, N., Fiske, S. T., & Yzerbyt, V. Y. (2013). Integrating the stereotype content model (warmth and competence) and the Osgood semantic differential (evaluation, potency, and activity). *European Journal of Social Psychology, 43,* 673–681. http://dx.doi.org/10.1002/ejsp.1978

Kervyn, N., Fiske, S. T., & Yzerbyt, Y. (2015). Foretelling the primary dimension of social cognition: Symbolic and realistic threats together predict warmth in the stereotype content model. *Social Psychology, 46,* 36–45. http://dx.doi.org/10.1027/1864-9335/a000219

Kervyn, N., Yzerbyt, V. Y., Judd, C. M., & Nunes, A. (2009). A question of compensation: The social life of the fundamental dimensions of social perception. *Journal of Personality and Social Psychology, 96,* 828–842. http://dx.doi.org/10.1037/a0013320

Koch, A., Alves, H., Krüger, T., & Unkelbach, C. (in press). A general valence asymmetry in similarity: Good is more alike than bad. *Journal of Experimental Psychology: Learning, Memory, and Cognition*.

Koenig, A. M., & Eagly, A. H. (2014). Evidence for the social role theory of stereotype content: Observations of groups' roles shape stereotypes. *Journal of Personality and Social Psychology, 107,* 371–392. http://dx.doi.org/10.1037/a0037215

Kriegeskorte, N., & Mur, M. (2012). Inverse MDS: Inferring dissimilarity structure from multiple item arrangements. *Frontiers in Psychology, 3,* 245. http://dx.doi.org/10.3389/fpsyg.2012.00245

Krumhansl, C. L. (1978). Concerning the applicability of geometric models to similarity data: The interrelationship between similarity and spatial density. *Psychological Review, 85,* 445–463. http://dx.doi.org/10.1037/0033-295X.85.5.445

Kruskal, J. B., & Wish, M. (1978). *Multidimensional scaling*. Newbury Park, CA: Sage.

Langlois, J. H., & Roggman, L. A. (1990). Attractive faces are only average. *Psychological Science, 1,* 115–121. http://dx.doi.org/10.1111/j.1467-9280.1990.tb00079.x

Layman, G. C., & Carsey, T. M. (2002). Party polarization and "conflict extension" in the American electorate. *American Journal of Political Science, 46,* 786–802. http://dx.doi.org/10.2307/3088434

Leonardelli, G. J., Pickett, C. L., & Brewer, M. B. (2010). Optimal distinctiveness theory: A framework for social identity, social cognition, and intergroup relations. *Advances in Experimental Social Psychology, 43,* 63–113. http://dx.doi.org/10.1016/S0065-2601(10)43002-6

Mackie, D. M., Smith, E. R., & Ray, D. G. (2008). Intergroup emotions and intergroup relations. *Social and Personality Psychology Compass, 2,* 1866–1880. http://dx.doi.org/10.1111/j.1751-9004.2008.00130.x

Macrae, N. C., Milne, A. B., & Bodenhausen, G. V. (1994). Stereotypes as energy-saving devices: A peek inside the cognitive toolbox. *Journal of Personality and Social Psychology, 66,* 37–47. http://dx.doi.org/10.1037/0022-3514.66.1.37

Magee, J. C., & Galinsky, A. D. (2008). Social hierarchy: The self-reinforcing nature of power and status. *The Academy of Management Annals, 2,* 351–398. http://dx.doi.org/10.1080/19416520802211628

Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods, 44,* 1–23. http://dx.doi.org/10.3758/s13428-011-0124-6

McCarty, N., Poole, K. T., & Rosenthal, H. (2008). Polarized America: The dance of ideology and unequal riches. *Political Science Quarterly, 123,* 687–689. http://dx.doi.org/10.1002/j.1538-165X.2008.tb01815.x

McCrae, R. R., & Costa, P. T., Jr. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology, 52,* 81–90. http://dx.doi.org/10.1037/0022-3514.52.1.81

McCrae, R. R., & John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality, 60,* 175–215. http://dx.doi.org/10.1111/j.1467-6494.1992.tb00970.x

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1,* 30–46. http://dx.doi.org/10.1037/1082-989X.1.1.30

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for similarity. *Psychological Review, 100,* 254–278. http://dx.doi.org/10.1037/0033-295X.100.2.254

Meier, B. P., Hauser, D. J., Robinson, M. D., Friesen, C. K., & Schjeldahl, K. (2007). What's "up" with God? Vertical space as a representation of the divine. *Journal of Personality and Social Psychology, 93,* 699–710. http://dx.doi.org/10.1037/0022-3514.93.5.699

Meier, B. P., & Robinson, M. D. (2004). Why the sunny side is up: Association between affect and vertical position. *Psychological Science, 15,* 243–247. http://dx.doi.org/10.1111/j.0956-7976.2004.00659.x

Michel, J. B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., Pickett, J. P., . . . Aiden, E. L. (2011). Quantitative analysis of culture using millions of digitized books. *Science, 331,* 176–182. http://dx.doi.org/10.1126/science.1199644

Monin, B., & Oppenheimer, D. M. (2005). Correlated averages vs. averaged correlations: Demonstrating the warm glow heuristic beyond aggregation. *Social Cognition, 23,* 257–278. http://dx.doi.org/10.1521/soco.2005.23.3.257

Nelson, L. J., & Miller, D. T. (1995). The distinctiveness effect in social categorization: You are what makes you unusual. *Psychological Science, 6,* 246–249. http://dx.doi.org/10.1111/j.1467-9280.1995.tb00600.x

Nosofsky, R. M. (1992). Similarity scaling and cognitive process models. *Annual Review of Psychology, 43,* 25–53. http://dx.doi.org/10.1146/annurev.ps.43.020192.000325

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of the United States of America, 105,* 11087–11092. http://dx.doi.org/10.1073/pnas.0805664105

Oppenheimer, D. M., & Trail, T. E. (2010). Why leaning to the left makes you lean to the left: Effect of spatial orientation on political attitudes. *Social Cognition, 28,* 651–661. http://dx.doi.org/10.1521/soco.2010.28.5.651

Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making, 5,* 411–419.

Parsons, T., & Bales, R. F. (1955). *Family socialization and interaction processes*. Glencoe, IL: Free Press.

Pattyn, S., Rosseel, Y., & Van Hiel, A. (2013). Finding our way in the social world: Exploring the dimensions underlying social classification. *Social Psychology, 44,* 329–348. http://dx.doi.org/10.1027/1864-9335/a000130

Peeters, G. (1983). Relational and informational patterns in social cognition. In W. Doise & S. Moscovici (Eds.), *Current issues in European social psychology* (Vol. 1, pp. 201–237). Cambridge, UK: Cambridge University Press.

Peeters, G. (1992). Evaluative meanings of adjectives invitro and in context: Some theoretical implications and practical consequences of positive-negative asymmetry and behavioral-adaptive concepts of evaluation. *Psychologica Belgica, 32,* 211–231.

Potter, T., Corneille, O., Ruys, K. I., & Rhodes, G. (2007). "Just another pretty face:" A multidimensional scaling approach to face attractiveness and variability. *Psychonomic Bulletin & Review, 14,* 368–372. http://dx.doi.org/10.3758/BF03194079

Reyna, C., Dobria, O., & Wetherell, G. (2013). The complexity and ambivalence of immigration attitudes: Ambivalent stereotypes predict conflicting attitudes toward immigration policies. *Cultural Diversity & Ethnic Minority Psychology, 19,* 342–356. http://dx.doi.org/10.1037/a0032942

Rhodes, G., Halberstadt, J., & Brajkovich, G. (2001). Generalization of mere exposure effects to averaged composite faces. *Social Cognition, 19,* 57–70. http://dx.doi.org/10.1521/soco.19.1.57.18961

Rosenberg, S., Nelson, C., & Vivekananthan, P. S. (1968). A multidimensional approach to the structure of personality impressions. *Journal of Personality and Social Psychology, 9,* 283–294. http://dx.doi.org/10.1037/h0026086

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology Review, 5,* 296–320. http://dx.doi.org/10.1207/S15327957PSPR0504_2

Rusbult, C. E., & Zembrodt, I. M. (1983). Responses to dissatisfaction in romantic involvements: A multidimensional scaling analysis. *Journal of Experimental Social Psychology, 19,* 274–293. http://dx.doi.org/10.1016/0022-1031(83)90042-2

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39,* 1161–1178. http://dx.doi.org/10.1037/h0077714

Schiffman, S. S., Reynolds, M. L., & Young, E. W. (1981). *Introduction to multidimensional scaling: Theory, methods and applications*. New York, NY: Academic Press.

Schubert, T. W. (2005). Your highness: Vertical positions as perceptual symbols of power. *Journal of Personality and Social Psychology, 89,* 1–21.

Schwartz, S. H. (1994). Are there universal aspects in the structure and contents of human values? *Journal of Social Issues, 50,* 19–45. http://dx.doi.org/10.1111/j.1540-4560.1994.tb01196.x

Schwartz, S. H., & Bilsky, W. (1987). Toward a universal psychological structure of human values. *Journal of Personality and Social Psychology, 53,* 550–562. http://dx.doi.org/10.1037/0022-3514.53.3.550

Schwartz, S. H., & Bilsky, W. (1990). Toward a theory of the universal content and structure of values: Extensions and cross-cultural replications. *Journal of Personality and Social Psychology, 58,* 878–891. http://dx.doi.org/10.1037/0022-3514.58.5.878

Shaver, P., Schwartz, J., Kirson, D., & O'Connor, C. (1987). Emotion knowledge: Further exploration of a prototype approach. *Journal of Personality and Social Psychology, 52,* 1061–1086. http://dx.doi.org/10.1037/0022-3514.52.6.1061

Sherman, J. W., Lee, A. Y., Bessenoff, G. R., & Frost, L. A. (1998). Stereotype efficiency reconsidered: Encoding flexibility under cognitive load? *Journal of Personality and Social Psychology, 75,* 589–606. http://dx.doi.org/10.1037/0022-3514.75.3.589

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86,* 420–428. http://dx.doi.org/10.1037/0033-2909.86.2.420

Slepian, M. L., Masicampo, E. J., & Ambady, N. (2015). Cognition from on high and down low: Verticality and construal level. *Journal of Personality and Social Psychology, 108,* 1–17. http://dx.doi.org/10.1037/a0038265

Smith, E. R. (1993). Social identity and social emotions: Toward new conceptualizations of prejudice. In D. M. Mackie & D. L. Hamilton (Eds.), *Affect, cognition and stereotyping: Interactive processes in group perception* (pp. 297–315). San Diego, CA: Academic Press. http://dx.doi.org/10.1016/B978-0-08-088579-7.50017-X

Sofer, C., Dotsch, R., Wigboldus, D. H. J., & Todorov, A. (2015). What is typical is good: The influence of face typicality on perceived trustworthiness. *Psychological Science, 26,* 39–47. http://dx.doi.org/10.1177/0956797614554955

Sutherland, C. A., Oldmeadow, J. A., Santos, I. M., Towler, J., Burt, D. M., & Young, A. W. (2013). Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition, 127,* 105–118. http://dx.doi.org/10.1016/j.cognition.2012.12.001

Suitner, C., & Maass, A. (2008). The role of valence in the perception of agency and communion. *European Journal of Social Psychology, 38,* 1073–1082.

Tajfel, H. (1969). Cognitive aspects of prejudice. *Journal of Social Issues, 25,* 173–191. http://dx.doi.org/10.1017/S0021932000023336

Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis. *Psychological Bulletin, 110,* 67–85. http://dx.doi.org/10.1037/0033-2909.110.1.67

Todorov, A., Dotsch, R., Wigboldus, D. H. J., & Said, C. P. (2011). Data-driven methods for modeling social perception. *Social and Personality Psychology Compass, 5,* 775–791. http://dx.doi.org/10.1111/j.1751-9004.2011.00389.x

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology, 66,* 519–545. http://dx.doi.org/10.1146/annurev-psych-113011-143831

Tolstoy, L. (1873–1877/2001). *Anna Karenina* (R. P. L. Volokhonsky, Trans.). New York, NY: Viking Penguin. (Original work published 1875–1877)

Torgerson, W. S. (1965). Multidimensional scaling of similarity. *Psychometrika, 30,* 379–393. http://dx.doi.org/10.1007/BF02289530

Tversky, A. (1977). Features of similarity. *Psychological Review, 84,* 327–352. http://dx.doi.org/10.1037/0033-295X.84.4.327

Unkelbach, C. (2012). Positivity advantages in social information processing. *Social and Personality Psychology Compass, 6,* 83–94. http://dx.doi.org/10.1111/j.1751-9004.2011.00407.x

Unkelbach, C., Fiedler, K., Bayer, M., Stegmüller, M., & Danner, D. (2008). Why positive information is processed faster: The density hypothesis. *Journal of Personality and Social Psychology, 95,* 36–49. http://dx.doi.org/10.1037/0022-3514.95.1.36

Unkelbach, C., Forgas, J. P., & Denson, T. F. (2008). The turban effect: The influence of Muslim headgear and induced affect on aggressive responses in the shooter bias paradigm. *Journal of Experimental Social Psychology, 44,* 1409–1413. http://dx.doi.org/10.1016/j.jesp.2008.04.003

van Elk, M., van Schie, H. T., & Bekkering, H. (2010). From left to right: Processing acronyms referring to names of political parties activates spatial associations. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 63,* 2202–2219. http://dx.doi.org/10.1080/17470218.2010.495160

van Kleef, G. A., De Dreu, C. K., Pietroni, D., & Manstead, A. S. (2006). Power and emotion in negotiation: Power moderates the interpersonal effects of anger and happiness on concession making. *European Journal of Social Psychology, 36,* 557–581. http://dx.doi.org/10.1002/ejsp.320

Vul, E., Harris, C., Winkielman, P., & Pashler, H. (2009). Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspectives on Psychological Science, 4,* 274–290. http://dx.doi.org/10.1111/j.1745-6924.2009.01125.x

Wells, G. L., & Windschitl, P. D. (1999). Stimulus sampling and social psychological experimentation. *Personality and Social Psychology Bulletin, 25,* 1115–1125. http://dx.doi.org/10.1177/01461672992512005

Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General, 143,* 2020–2045. http://dx.doi.org/10.1037/xge0000014

Wheeler, S. C., & Petty, R. E. (2001). The effects of stereotype activation on behavior: A review of possible mechanisms. *Psychological Bulletin, 127,* 797–826. http://dx.doi.org/10.1037/0033-2909.127.6.797

Williams, J. E., & Best, D. L. (1990). *Measuring sex stereotypes: A multination study*. Newbury Park, CA: Sage.

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after a 100-ms exposure to a face. *Psychological Science, 17,* 592–598. http://dx.doi.org/10.1111/j.1467-9280.2006.01750.x

Wojciszke, B. (1994). Multiple meanings of behavior: Construing actions in terms of competence and morality. *Journal of Personality and Social Psychology, 67,* 222–232. http://dx.doi.org/10.1037/0022-3514.67.2.222

Wojciszke, B., & Abele, A. E. (2008). The primacy of communion over agency and its reversals in evaluations. *European Journal of Social Psychology, 38,* 1139–1147. http://dx.doi.org/10.1002/ejsp.549

Wojciszke, B., Bazinska, R., & Jaworski, M. (1998). On the dominance of moral categories in impression formation. *Personality and Social Psychology Bulletin, 24,* 1251–1263. http://dx.doi.org/10.1177/01461672982412001

Young, F. W., Takane, Y., & Lewyckyj, R. (1978). ALSCAL: A nonmetric multidimensional scaling program with several individual-differences options. *Behavior Research Methods & Instrumentation, 10,* 451–453. http://dx.doi.org/10.3758/BF03205177

Yzerbyt, V. Y., Kervyn, N., & Judd, C. M. (2008). Compensation versus halo: The unique relations between the fundamental dimensions of social judgment. *Personality and Social Psychology Bulletin, 34,* 1110–1123. http://dx.doi.org/10.1177/0146167208318602

Received October 14, 2015
Revision received December 29, 2015
Accepted December 30, 2015 ■