Journal of Experimental Social Psychology 77 (2018) 76–88

Contents lists available at ScienceDirect

Journal of Experimental Social Psychology

journal homepage: www.elsevier.com/locate/jesp

(Pre)occupations: A data-driven model of jobs and its consequences for categorization and evaluation

Roland Imhoff[a,b,*], Alex Koch[b], Felicitas Flade[a]

[a] Social and Legal Psychology, Johannes Gutenberg University, Mainz, Germany
[b] Social Cognition Center Cologne, University of Cologne, Germany

ARTICLE INFO

Keywords:
Occupations
Stereotypes
Categorization
Multi-dimensional scaling
Evaluative contagion
Agency

ABSTRACT

We present a data-driven model of stereotypes about occupations (total $N = 3919$). Across two classification systems and national contexts (U.S.; Germany), we show remarkable convergence in the stereotype dimensions spontaneously employed to make sense of occupational groups (agency; progressiveness). Further studies show that these dimensions reflect presumed characteristics of job holders and not just describe their occupational role (Study 2), and that proximity of occupations on the emerging stereotype model increased superordinate categorization (Study 3) and contagious transfer of (positive and negative) valence from one occupation to another (Study 4). Together these studies do not only provide important insights into the perception of one of the most ubiquitous social taxonomies but also provide a rich, open access dataset for researchers seeking to employ occupational groups as a tool to better understand stereotypes and intergroup relations in general.

To simplify orienting and navigating in today's complex social world, people infer others' informative but not immediately observable characteristics from the groups they evidently belong to, also known as stereotyping. Besides gender, age, race and other stereotypes, people also form, use, and share occupational stereotypes. For example, people believe they know that librarians are shy, models are flamboyant, construction workers are tough-minded, and bankers are greedy. In the present paper we seek to provide an integrative model of dimensions on which people typically compare and position occupational groups to make sense of their social surrounding.

There are numerous examples in the social psychological literature of specific stereotypes people hold about occupational group. Knowing that someone is a scientist will likely evoke an image of this person as being maybe likeable but also robot-like and somewhat obsessed, and capable of immoral conduct (Rutjens & Heine, 2016). Politicians and lobbyists are seen as powerful and as relatively threatening (Imhoff & Bruder, 2014). Male nurses are more helpful than stockbrokers (Abele & Petzold, 1998), and physicians are seen as truthful, competent and altruistic, whereas used car salesmen are not (Rotter & Stein, 1971). Such stereotypes may be derived from the occupational activity. Originally proposed to explain gender stereotypes, social role theory (Eagly, Wood, & Diekman, 2000) posits that observed social roles held by men and women are used to infer presumed characteristics of men and women (i.e., gender stereotypes). The same principle might apply to other activities and roles people hold in society, most prominently, their occupation. If so, occupational stereotypes should closely align with what people believe to be central activities in a job.

Although it is of course relevant to describe the specific stereotype people hold of specific groups, our goal in the current research was to systematize these stereotypes. Specifically, we were interested whether occupations are more likely categorized on some dimensions than others. Existing taxonomies placed occupations on continua from hierarchy-enhancing to hierarchy-attenuating (e.g., Pratto, Stallworth, Sidanius, & Siers, 1997; Sidanius, Liu, Pratto, & Shaw, 1994), but these were typically theory-driven researcher-based sorting schema, not data-driven insights into how people spontaneously mentally arrange the occupational field.

On the lookout for dimensions of spontaneous occupational stereotypes, one both relevant and established finding is that people readily categorized some occupations as stereotypically (fe)male (Wilbourn & Kee, 2010). Particularly occupations with leadership responsibilities were construed as masculine (Koenig, Eagly, Mitchell, & Ristikari, 2011) and women's likelihood of being offered a leadership position was higher if they have stereotypically masculine attributes (Glick, Zion, & Nelson, 1988). Given the wealth of findings relating occupations to gender, it seems plausible that occupations are spontaneously compared and categorized along a stereotype dimension ranging from female/feminine to male/masculine.

Other research has simply posited that occupational stereotypes follow the same dimensionality as other stereotypes and thus suggested

* Corresponding author at: Social and Legal Psychology, Binger Str. 14-16, 55122 Mainz, Germany.
E-mail address: roland.imhoff@uni-mainz.de (R. Imhoff).

https://doi.org/10.1016/j.jesp.2018.04.001
Received 1 December 2017; Received in revised form 23 March 2018; Accepted 1 April 2018
Available online 27 April 2018
0022-1031/

the influential stereotype content model (SCM; Fiske, Cuddy, Glick, & Xu, 2002) and its two dimensions of warmth and competence as an applicable model also for occupational stereotypes (Fiske & Dupree, 2014; Imhoff, Woelki, Hanke, & Dotsch, 2013; Koenig & Eagly, 2014). One critical aspect of getting at spontaneous stereotypes, however, is to make no a priori decision in the research design that only allows specific dimensions to be applied (e.g., by asking only about warmth and competence; Fiske & Dupree, 2014). Revealing the stereotype dimensions that people spontaneously use to make sense of others based on their occupation requires ecologically valid studies (Brunswik, 1955) in which people are free to choose any desired stereotype dimension(s) to a representative sample of stimuli (for a recent approach to stereotypes about social groups see Koch, Imhoff, Dotsch, Unkelbach, & Alves, 2016).

1. The present research

We took a data-driven approach to spontaneous occupational. In two studies, people rated the similarity of two different exhaustive samples of occupations. Importantly, similarity allowed people to compare the occupations on any desired stereotype dimension(s) that spontaneously came to their mind. If that was morality, they would rate nurses and surgeons as more similar than surgeons and lawyers. If it was intelligence, they would rate nurses and surgeons as less similar than surgeons and lawyers. Importantly, if most people would consensually rate the similarity of most occupations on the same dimensions, these spontaneous occupational stereotypes could be identified by computing and interpreting a cognitive model that would visualize the pattern of the consensual similarity ratings. In Studies 1a and 1b, we computed cognitive models of 150 U.S. occupations, respectively 88 German occupations and interpreted the dimensions that spanned this "map". This also allowed an exploration whether such stereotype dimensions align with dimensions developed to characterize occupational roles. In Study 2 we sought to differentiate people's impression of a job from their impression of people who have this job as only the latter constitutes a stereotype (inferring people's characteristics from their group memberships) in the strict sense. Based on these findings, we examined automatic social categorization (Studies 3a–c) and lateral attitude change (Studies 4a–b; Glaser et al., 2015) as downstream consequences. We report all studies, as well as therein all measures, manipulations, and exclusions (if any) conducted in this research line. In lack of informed estimates of effect sizes, sample size for each study was determined by rule of thumb: 25 raters per rating and 100 participants per cell in between-subject designs. No intermediate analyses were conducted and there was no continued data collection after data analysis. All materials, data and supplemental figures and tables are available on our OSF project site (link).

2. Study 1

To identify the dimensions people spontaneously use to stereotype others based on their occupation we followed a sequence of three steps in two national contexts: USA and Germany. For both contexts, we generated an exhaustive list of occupations. We started from two different classification systems that seek to include all occupations: the U.S. Department of Labor's Occupational Employment Survey in Study 1a and the International Standard Classification of Occupations in Study 1b and adapted these lists to increase comprehension. We then collected estimates of similarity between typical holders of all occupations and subjected these similarity estimates to multidimensional scaling (MDS; for an introduction, see Hout, Papesh, & Goldinger, 2013; for an example, see Lammers, Koch, Conway, & Brandt, 2017). For both contexts, this resulted in three-dimensional spaces in which typical occupation holders stereotyped as more similar were positioned closer to one another. To understand the dimensions that spanned this space we aligned the typical occupation holders' coordinates on the three space dimensions with independent ratings of potential stereotype dimensions (property fitting analyses). As an important difference, the researchers selected potential candidate dimensions that spanned the space in Study 1a, which could introduce biases and limit generalizability of results. In Study 1b, therefore, participants themselves generated labels for the dimensions, which were then synthesized and later rated for each occupation by another group of participants.

2.1. Method and results

2.1.1. Study 1a

2.1.1.1. A complete list of U.S. occupations

We approximated a complete list of occupations based on the U.S. Department of Labor's Occupational Employment Survey (OES) of 2012 (http://www.bls.gov/oes/special.requests/oesm12all.zip). According to this survey, the highest order of the North American Industry Classification System lists 457 "broad" occupations. The list contains 16, 24, and 25 kinds of teachers, engineers, and managers, respectively, and numerous other highly similar occupations. As we were not interested in such detail, we cut down the list based on first two authors' consensual decision. We merged highly similar occupations and shortened long occupation titles to everyday equivalents (e.g., "Agents and Business Managers of Artists, Performers, and Athletes" became "Agents of Artists"). In some cases, ambiguous occupations titles were split into their components (e.g., "Physicians and Surgeons" became "Physicians" and "Surgeons"). Table osm.1 (see online supplementary materials) shows the final list of 150 occupations.

2.1.1.2. Computing the cognitive model of U.S. occupations

To explore the stereotype dimensions that people spontaneously use to make sense of others based on their occupation, 213 U.S. Americans recruited from Amazon's Mechanical Turk (M_{age} = 34.05, SD = 10.54; 101 women, 111 men) were instructed to "position 50 occupational groups [randomly drawn from the list of 150 occupations] on the computer screen according to how similar or dissimilar you perceive typical members of these groups to be." (Figs. osm.1–osm.2). This spatial arrangement method (SpAM; Alves, Koch, & Unkelbach, 2016; Hout, Goldinger, & Ferguson, 2013; Hout & Goldinger, 2016; Koch, Alves, Krüger, & Unkelbach, 2016) measures similarity in terms of proximity. This is a particularly efficient approach as repositioning a stimulus simultaneously adjusts the proximities/similarities between that stimulus and all other stimuli on the screen. We recorded the distance between two occupations in relation to the greatest possible distance (the screen diagonal).

For each of the 11,175 pairs that could be formed with the 150 occupations, we averaged dissimilarity across all participants who repositioned the two occupations. We subjected the 11,175 mean dissimilarity indices to MDS with an ALSCAL procedure (Young, Takane, & Lewyckyj, 1978); assuming an interval scale, we estimated coordinates for the 150 occupations in a 1D-6D model. Scaling fit was indicated by a (preferably low) scaling stress (S; 0.16, 0.14, 0.12, 0.11, 0.10, and 0.10 for the 1D, 2D, 3D, 4D, 5D, and 6D model, respectively). Balancing fit and parsimony of the scaling solution (Jaworska & Chupetlovska-Anastasova, 2009), we proceeded with interpreting the 3D cognitive model of U.S. occupations.

2.1.1.3. Interpreting the model of U.S. occupations

Rotating the "map", we searched for and selected a number of stereotype dimensions that possibly spanned the "map" (i.e., spontaneous stereotypes that people could have used to rate the occupations' similarity). These candidate dimensions inspired by the data were augmented with candidates derived from main theories of stereotype content (e.g., Fiske et al., 2002), resulting in the 41 candidates shown in Table osm.2. 1245 MTurkers (M_{age} = 34.04, SD = 11.77; 517 women,

Table 1
Representing and relating four factor-analytically derived stereotype dimensions and Holland's (1997) RIASEC in the occupations' 3D model.

	R	1.	2.	3.	4.	5.	6.	7.	8.	9.
Stereotypes										
1. Agency	0.86									
2. Progressiveness	0.72	−0.22								
3. Sociability	0.59	−0.07	0.17							
4. Communion	0.41	0.19	−0.99	−0.03						
RIASEC Ratings										
5. Realistic	0.86	−0.39	−0.40	−0.82						
6. Investigative	0.78	0.97	−0.47	−0.10		−0.25				
7. Artistic	0.58	0.43	0.78	0.17		−0.66	0.19			
8. Social	0.69	0.27	−0.38	0.82		−0.65	0.36	−0.12		
9. Enterprising	0.54	−0.05	0.68	0.83		−0.88	−0.22	0.64	0.42	
10. Conventional	0.37	−0.54	0.19	0.87		−0.48	−0.53	−0.13	0.59	0.68

Note. Strongest possible correlation between rating dimension and any dimension running through the three dimensional space (*R*) as well as all correlations of these dimensions' coordinates if projected in the space.

696 men, 1 other, 31 did not indicate gender)[1] used a 0–100 slider scale with a "no response" option to rate all 150 occupations on one of the 41 candidates (randomly chosen; see Figs. osm.3–osm.4). Raters' agreement about the occupations was very high, $ICC(2,k) \geq 0.79$, for all 41 candidates.

To systematize interpretation, we ran a principal component analysis (varimax rotation; details Table osm.3; separate property fitting analyses for all 41 candidate dimensions Table osm.2) on the 41 candidates which resulted in four factors: *agency* (e.g., powerful, assertive, high status; $\alpha = 0.96$), *progressiveness* (e.g., creative, liberal, promotion-oriented; $\alpha = 0.90$), *sociability* (e.g., outgoing, sociable, interconnected; $\alpha = 0.89$), and *communion* (e.g., sincere, trustworthy, warm; $\alpha = 0.95$).

To interpret which spontaneous stereotypes people consensually used to rate the occupations' similarity, we ran four multiple regressions with the occupations' mean rating on the four factors as criterion and their x-, y-, and z-coordinates in the model as predictors. The multiple correlation of each of these property fitting analyses (e.g., Deng, Armstrong, & Rounds, 2007) showed the extent to which respective candidate could be fitted to the model by rotating it. *R* in Table 1 shows the strongest possible correlation between the respective component and a dimension running through the space. Table 1 also shows the correlation between each dimension running through the space that best represented one component and each dimension running through the space that best represented another component. Treating $r \mid < 0.40 \mid$ as independent enough, agency, progressiveness, and sociability were indeed adequate interpretations for the three dimensions that spanned the space (Fig. 1), whereas communion did not show a sufficiently strong ($R > 0.50$) relation to any dimension in that space ($R = 0.41$).

Social role theory posits that stereotypes should align with the role people fulfill, here: their occupational duties. Based on Holland (1959, 1997)'s influential taxonomy of occupations we tested whether vocational character traits realistic, investigative, artistic, social, enterprising, and conventional (together: RIASEC) offered a better interpretation of the occupations' model, we obtained up-to-date (i.e., 2008–2013) RIASEC scores for the 150 occupations from the O*NET website (for the scores, see https://www.onetcenter.org/dl_files/database/db_20_3_excel/Interests.xlsx; for the scoring procedure, see Rounds, Armstrong, Liao, Lewis, & Rivkin, 2008). Just focusing on these six dimensions (lower half of Table 1), and Treating $|r| < 0.40$ as independent enough, the three independently represented facets in RIASEC that were best represented in the model were *investigative, artistic*, and *social*. Although *realistic* had the strongest relation to any of the dimensions, it showed substantial correlations with virtually all other dimensions, and may thus be – at least in the space – a composite of them. Thus, treating the RIASEC dimensions of vocational interest as potential candidates for stereotypes about how typical occupation holders are, the best available orthogonal model of the three spontaneous stereotypes that people consensually used to rate the occupations' similarity was thus investigative, artistic, and social.[2]

Table 1 also allows a comparison of the two models. The two models derived from our ratings as well as the RIASEC coding were equally well-fitting models of the content of spontaneous occupational stereotypes (mean *R*s = 0.72 and 0.68, respectively). The correlation between the model's two dimensions that best represented *agency* and *investigative* was extremely high, $r = 0.97$, and it was high for *progressiveness* and *artistic*, $r = 0.78$, and for *sociability* and *social*, $r = 0.82$. Thus, the two models were not just equally well-fitting but empirically almost identical. Their only differences were aspects that separate progressiveness from artistic and sociability from social. The model based on our ratings was slightly better because it came closer to orthogonal, $|r|s < 0.23$, than the model based on RIASEC ratings, $|r|s < 0.37$.

2.1.2. Study 1b

2.1.2.1. International occupations

To validate the APS model, we started out from the 111 "minor" occupations of the International Standard Classification of Occupations (http://www.ilo.org/public/english/bureau/stat/isco/docs/resol08.pdf). Based on consensus in a committee discussion, four raters agreed on dropping rare categories (e.g., "Mining and mineral processing plant operators"), breaking down broad categories (e.g., "Architects, engineers and related professionals"), and renaming technical categories (e.g., "Nursing and midwifery associate professionals"). Table osm.4 shows the final list of 84 occupations.

2.1.2.2. Computing the cognitive model of international occupations

119 students of different majors ($M_{age} = 26.45$, $SD = 12.18$; 56 women, 62 men) were presented with 42 random occupations and spatially arranged occupations whose typical members are more similar closer to one another on the screen. The only two differences between this and Study 1a's task were an initial overview of the occupations to be spatially arranged to avoid order effects and the addition of a horizontal and vertical axis to the spatial arrangement (SpAM; Hout, Papesh, & Goldinger, 2013) board (see Figs. osm.5–osm.7). We reasoned that these axes might facilitate participants to think in terms of dimensions rather than clusters, and also in their task to generate labels

[1] Another 26 raters (8 women, 17 men) judged the pleasantness of the 150 job activities per se to test if it was a better description of people's similarity estimates than stereotypes about typical holders of the 150 jobs. This was not the case (Table osm.2).

[2] Note that this 3D IAS model deviates from Holland (1997)'s 2D RIASEC model of the structure of vocational interests. In Holland (1997)'s model, RIASEC form a 2D hexagon with dimensions at opposite corners negatively correlated (e.g., realistic vs. social). Our model of the structure of occupational stereotypes was different in that no dimension running through the map represented C well, shortening RIASEC to RIASE. Second, our model was 3D with I orthogonal(ish) to the plane of RASE.

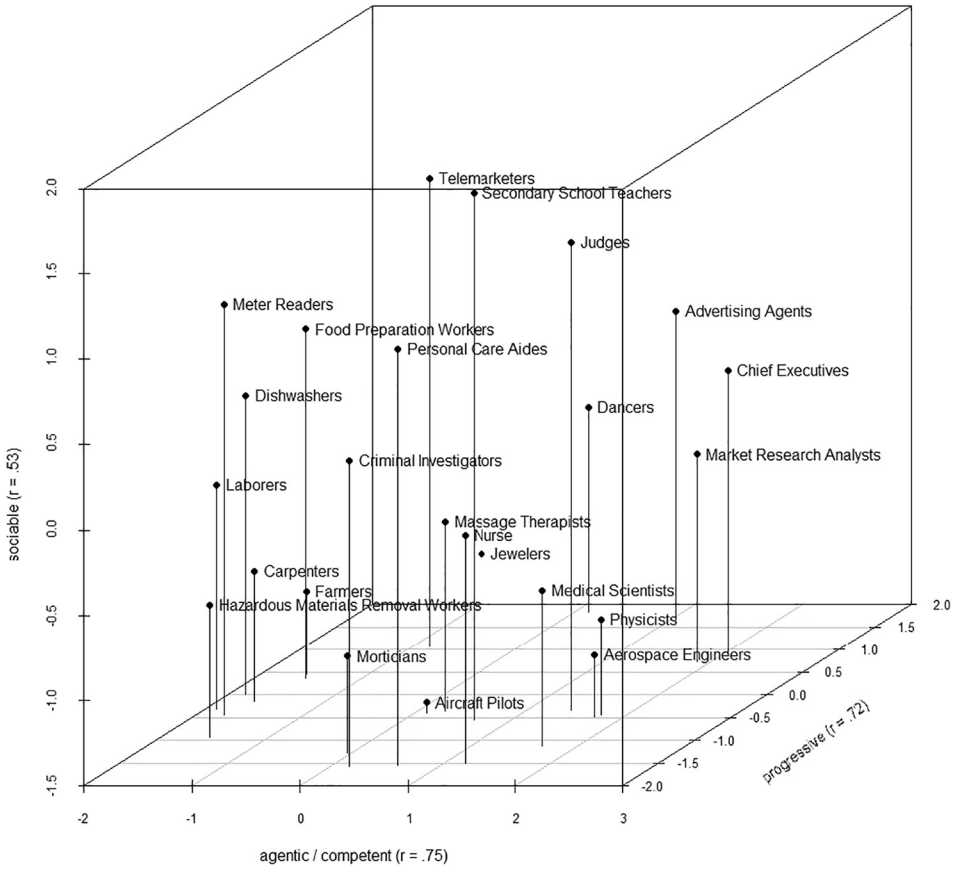

Fig. 1. 3D model of 150 U.S. occupations (24 occupations are shown). Abandoning orthogonality in favor of model fit yielded agency, progressiveness and sociability as the best available interpretations of people's spontaneous occupational stereotypes. These dimensions were almost orthogonal, $rs \mid < 0.23|$, and spanned the model close to where shown.

for dimensions. Upon pressing "I finished", the occupations' spatially arranged positions were fixed, and people were instructed to interpret the occupations' distribution along the horizontal and vertical axis by entering 1 noun label and 1–9 adjective labels for each axis into text boxes.

We obtained dissimilarity indices for the occupations in the same way as in Study 1a and multidimensionally scaled them in the same way as in Study 1a. Scaling fit was 0.16, 0.12, 0.10, 0.10, 0.09, 0.09, and 0.08 for the 1D, 2D, 3D, 4D, 5D, and 6D model, respectively. The 3D model's fit ($S = 0.10$) was satisfactory, indicating that three orthogonal dimensions sufficed to describe the spontaneous stereotypes that people consensually used to rate the occupations' similarity. As in Study 1a, extracting a fourth dimension did not improve the fit, and thus we proceeded with interpreting the 3D cognitive model of international occupations.

2.1.2.3. Interpreting the model of international occupations

Participants entered > 700 labels to interpret the horizontal and vertical axis of the spatial arrangement board. Many of the labels were identical or synonymous. Five raters systematized the labels (see Table osm.5) and selected ten stereotype dimensions that best represented them based on consensus reached in discussion. 208 additional participants ($M_{age} = 25.55$, $SD = 5.58$, 149 women, 59 men) used a 0–10 slider scale to rate all 84 occupations on two of the ten candidate dimensions (randomly chosen; Fig. osm.8). There were between 36 and 47 raters per candidate, and their agreement on the occupations was high, $ICC(2,k) \geq 0.94$.

To systematize interpretation, we ran a principal component analysis on the ten candidates with the same settings as in Study 1, resulting in two components (see Table osm.6) so similar to Study 1a's agency (e.g., high status, assertive; $\alpha = 0.88$) and progressiveness (e.g., liberal, creative; $\alpha = 0.71$) (see Table osm.3) that we decided to name them identically. Because the internal consistency of the third component was far from satisfactory ($\alpha = 0.41$), we decided to discard it and proceed with single items that represented sociability (extraverted) and communion (trustworthy).

As in Study 1a, we ran property fitting analysis on these four dimensions (for a property fitting analysis on the ten separate candidates, see Table osm.7). *R* in Table 2 shows the strongest possible correlation between the respective component and a dimension running through the model as well the correlation between the dimensions running through the model.

As in Study 1a, Agency and Progressiveness were the best-fitting candidates (albeit slight violating the standard of independence of $|r| < 0.40$). The single items representing sociability and communion were equally suited candidates to represent a dimension, but communion was more orthogonal to the other two dimensions.

2.2. Discussion

In Study 1, we took a data-driven approach to the stereotype dimensions that people spontaneously use to compare others based on their occupations in two different national context (USA, Germany) and two different sets of occupations. People rated the similarity of typical members of exhaustively sampled occupations, allowing comparing them with respect to any desired dimension(s). Multidimensional

Table 2
Study 1b: representing and relating four stereotypes in the occupations' 3D model.

Candidate/component	*R*	1.	2.	3.
1. Agentic/competent	0.89			
2. Progressive	0.77	0.44		
3. Social	0.67	0.57	0.67	
4. Communal	0.65	0.17	0.34	−0.41

scaling, principal components, and property fitting analyses showed that across both studies people's first independent occupational stereotype was *agency* with surgeons, software developers, and aerospace engineers at the top of the occupational ladder and cashiers, telemarketers, and parking lot attendants at the bottom. People's second independent stereotype was best described by *progressiveness*. Ambulance drivers, firefighters, and police officers scored at the conservative, conventional, and prevention-oriented end of this dimension; musicians, athletes, and designers were stereotyped as most liberal, alternative, and promotion-oriented. For the third dimension, there was slightly less consistency across studies. Whereas Study 1a suggested *sociability* (with teachers, clergy, and tellers scoring high and pilots, fishers, and farmers scoring low), Study 1b did not fully replicate this finding with an arguably sub-optimal one-item measure of sociability (extraverted). Instead, communion seemed to be the better (i.e., less overlapping candidate for the third dimension). The reasons for this inconsistency may be manifold. First, there may indeed be cultural differences at play here, but it is also conceivable that slight changes to the paradigm (with two axial lines suggesting a two-dimensional solution) might be responsible for this. In light of the fact that a) this only rests on single-item measures on Study 1b and b) the third dimension did not particularly strongly reduce the stress of the MDS solution in either study, we would refrain from putting too much emphasis on this inconsistency and instead focus on the remarking consistency for the first two dimensions. In two different countries, with two different sets of occupations and two different approaches to interpret the final dimensions (consensual candidates vs. idiosyncratic labelling) virtually the same cognitive model emerged.

Occupational stereotypes are not the only domain in which the two dimensions that best described this space, *agency* and *progressiveness*, appear. To our knowledge, highly similar constructs appeared as assertive and loose in Peabody's (1967) personality model, as economic development and political alignment in Wish, Deutsch, and Biener's (1970) nations model, as competitive and modern in Jones and Ashmore's (1973) personality model, as self-enhancement and openness to change in Schwartz and Bilsky (1987)'s values model, as agency (A) and progressive beliefs (B) in Koch and colleagues' (2016) recent social group model, and as prosperity and ideology in Koch, Kervyn, Kervyn, and Imhoff's (in press) new U.S. states model. The current research thus adds further substance to the ubiquity of these two dimensions.

Further, the cognitive model of occupational stereotypes is (as specifically tested for Study 1a) virtually identical to a model resting on well-established dimensions of vocational interest (RIASEC). From the perspective of social role theory, this seems to suggest that people infer jobholder's *agency* from the extent to which their job is *investigative*, their *progressiveness* from the extent to which it is *artistic* and their *sociability* from the extent to which it is *social*. Given the plausibility of these stereotype links, the two models of spontaneous occupational stereotypes validated each other. This is particularly noteworthy as the RIASEC model (Holland, 1959, 1997) was developed based on self-rated occupational interest, whereas Study 1a's model was developed based on comparisons of typical job holders (i.e., stereotypical similarity).

Before turning to the issue of how people use these dimensions when thinking about occupational groups in spontaneous manner, it seems important to discuss whether the space dimensions identified in Study 1 really reflect stereotypes about individuals. Specifically, it is conceivable that the space dimensions reflect how people see occupations per se rather than the professionals who hold them. For instance, the evidence in support of *agency* in Study 1 might actually reflect people's stereotypes about how much money others make in different jobs. Likewise, the evidence in support of *progressiveness* might actually reflect stereotypes about what people do in their job in the sense of either conserving the status quo by preventing certain negative things (crime, decay, death etc.) or progressing the status quo by promoting certain arguably positive things (art, advertisement, athletic performance). To show that the space dimensions identified in Study 1 indeed reflect stereotypes about individuals and not just the role they fulfill in their job we conducted two further studies.

3. Studies 2a and 2b

In the first study, participants were instructed to rate the similarity between different occupations based on how they perceive typical exemplars to be. We therefore have interpreted the resulting dimensions as the fundamental dimensions on which occupational group are stereotyped. Stereotypes imply that group membership serves as a cue to infer stereotype-consistent characteristics of the individual. It is therefore crucial that these beliefs really concern the individuals who hold the respective profession and are not just part of their role in that profession (as posited by social role theory). As an example, people might think that a nuclear safety officer will have to be extremely prevention-oriented in his job, trying to contain the nuclear chain reaction to the reactor chamber – therefore behaving rather conservative than progressive and rather reactive than creative. Simultaneously, they might think that this is restricted to his or her role as an engineer and has nothing to do with how that person is. At the end of the workday, this person might entertain an extremely liberal lifestyle and improvise little travesty sketches on a bar stage. Although real life provides anecdotal examples of art-loving programmers or conservative and narrow-minded musicians, our argument about stereotypes necessitates that people do not dissociate the occupational role from the individual to such an extent.

We tested this in Studies 2a and 2b by inquiring about the likelihood of a) occupational flexibility and b) two identical twins holding the same occupation. If the characteristics are just part of a role that can be taken off like a lab coat, distance on the dimensions should not affect the likelihood that people change from one job to another or the likelihood of two twins holding similar jobs. If, however, the occupational stereotype is used to infer something about the persons working on the job, then these should be judged as less likely to change into another job, the further away it moves from their original job. Likewise, identical twins should seem less likely to have two jobs that are distant from each other on the model.

3.1. Method

We selected 25 occupations from Study 1a to represent a great diversity on all three dimensions (see Fig. 1). Each of these was paired each of the other 24 occupations, resulting in 600 pairs (each pairing appearing twice, but differing which of the two was mentioned first). Study 2a (N = 183 MTurk workers; 90 men and 93 women aged between 19 and 74, M = 36.14, SD = 12.71) and Study 2b (N = 203 MTurk workers; 104 men, 97 women, and 2 other gender aged between 18 and 74, M = 30.78, SD = 10.19) were identical except for the judgment participants had to make about these pairs. In Study 2a, participants rated for 100 randomly selected pairs of occupations how difficult it is for persons who work in one occupation to work in the other occupation (e.g., “How difficult is it for Personal Care Aides to become Secondary School Teachers?”) on a 11-point scale from extremely easy (1) to extremely difficult (11). For Study 2b, they were asked to imagine two identical twins and estimate how likely it is that one of them has a certain occupation (e.g., baker), given that the other twin has a certain other occupation (e.g., cab driver) for 100 randomly selected pairs (e.g., “How likely is it that one twin is a personal care aid given that his or her identical twin is a criminal investigator?”) on a 11-point scale from extremely likely (1) to extremely unlikely (11).

3.2. Results

3.2.1. Study 2a

The difficulty of each occupational move was rated by between a minimum of 16 and a maximum of 49 raters (M = 30.50, SD = 5.22) and there was considerable – albeit not perfect – agreement between the

raters, Krippendorff's Alpha = 0.41, 95% CI [0.38; 0.44] in 5000 bootstrap samples (Hayes & Krippendorff, 2007). We thus averaged each estimation across raters. For each of the 600 pairs and each dimension[3] we subtracted the coordinate from Study 1 of the departing occupation from that of the occupation of destination. Higher scores thus reflect moving up on the respective dimension, whereas negative scores reflected moving down on that dimension. Based on these change scores we computed the distance in the three-dimensional space and correlated with the averaged difficulty ratings. Results indicated that the difficulty of changing occupations was perceived as greater, the further away the two respective occupations were in the three-dimensional space, $r = 0.26$, $p < .001$, independent of the direction (up or down).

To have a more fine-grained analysis, we then predicted the difficulty ratings with hierarchical multiple regressions. In a first step, we entered the raw difference scores to assess directional effects of either moving up or down on one of the three dimensions, $R^2 = 0.44$, $p < .001$. Moving up on the first dimension increased difficulty, $\beta = 0.56$, $p < .001$, whereas difficulty of job change was greater for moving down on the second, $\beta = -0.13$, $p < .001$, and third dimension, $\beta = -0.30$, $p < .001$. Thus, participants saw greater difficulty to change to a job that is typically done by people that are more intelligent, more conservative, and less sociable. Importantly, we added the (orthogonal) squared distances in a second step to test whether above and beyond these directional effects distance added incremental explanatory value which it did indeed, $\Delta R^2 = 0.06$, $p < .001$. Positive regression weights indicative of greater difficulty for greater distance were found for dimension 1, $\beta = 0.20$, $p < .001$, dimension 2, $\beta = 0.16$, $p < .001$, as well as dimension 3, $\beta = 0.13$, $p < .001$.

3.2.2. Study 2b

For each of the 600 pairs of occupations the likelihood of two identical twins having the respective occupations was rated by between a minimum of 18 and a maximum of 53 raters ($M = 36.65$, $SD = 5.52$). Agreement between the raters was low, Krippendorff's Alpha = 0.20, 95% CI [0.16; 0.24] in 5000 bootstrap samples, suggesting that there was only little (albeit significant) agreement on the estimated likelihood. Nevertheless, we proceeded by averaging across all raters to get an estimate of the overall average perceived likelihood for each pair.

Using the same distance estimates as in the previous study, it was perceived as more likely that two identical twins have two different occupations if the two were less distant in the three-dimensional space, $r = -0.64$, $p < .001$. Predicting the estimated likelihood with the raw differences on all three dimensions yielded no explained variance, $F(3, 596) = 0.06$, $p = .982$, $R^2 = 0.00$, but adding the quadratic terms did, $F(6, 593) = 61.89$, $p < .001$, $\Delta R^2 = 0.39$. Specifically, high distance on the first, $\beta = -0.61$, $p < .001$, the second, $\beta = -0.08$, $p = .011$, as well as the third dimension, $\beta = -0.15$, $p < .001$, independently predicted lower likelihood of two identical twins having these two occupations.

3.3. Discussion

Conjointly, Studies 2a and 2b show that participants did not just rely on perceived similarity between occupational roles when judging similarity between jobs but indeed thought of the persons who held these occupations. Perceived difficulty of changing from one occupation to another was greater the more dissimilar typical representatives of the respective occupations were seen. Additionally, this pattern was not fully symmetrical for any of the three extracted dimensions. One possible speculation to make sense of this asymmetry is that participant construed the ease of job switching as a function of required skills (for each dimension, either going up or down is easier than in the opposite direction because it requires less or more common skills). One could argue that if the effect was due to the need to acquire required skills, stereotypes are still more a function of the mere job description (what are the skills needed for this jobs) than the persons working on it.

This problem, however, was not apparent in Study 2b. The fact that participants saw the likelihood of identical twins holding two specific jobs as lower (vs. higher) when the respective two occupations were far apart (close together) in the occupational model speaks against the notion that the similarity judgments that we used to feed our MDS analyses are merely based on characterizations of the occupational role, detached from the individuals who fill that role. If these were indeed merely based on job descriptions that could be done by anybody, it would seem implausible that occupations for which people believe typical representatives to be particularly dissimilar are very unlikely to be held by identical twins. Although Study 2b has other problems (e.g., the extremely low interrater agreement), we could not conceive how these might introduce alternative explanations for the obtained results. If anything, low rater agreement, and thus low reliability should attenuate a real effect, not create one. In the following studies, we aimed at moving beyond the mere description of the mental model of occupations to implications this might have for basic processes like categorization and evaluative contagion.

4. Studies 3a–c

The mere existence of underlying dimensions as shown in the MDS solutions does not necessarily imply that they are also used in making sense of the world. One very basic process of everyday sense-making is categorization, i.e. the lumping of people into distinct categories to reduce complexity of the social world. If people do indeed recruit their presumed knowledge about the characteristics of certain occupations to make sense of the world around them, we would expect them to form higher order, superordinate categories of people situated high or low, respectively, on the basic dimensions. This of course is a complex endeavor as unlike other cues used for categorization (e.g., age, race, gender), occupations typically have no clear visual cues. In traditionally applied categories, e.g. black and white skin colors are a quite deterministic cue of the race dimension from early age on (Dunham, Stepanova, Dotsch, & Todorov, 2015). In our case, however, a selection of occupations could be arranged on many more meaningful dimensions that might be just as applicable as “agency” or “progressiveness”. It is thus crucial to show that people indeed rely on these dimensions to form inclusive superordinate categories (e.g., people high on dimension 1 vs people low on dimension 1).

Do people use the latent stereotype dimensions when they engage in spontaneous social categorization? The most widely and successfully used paradigm measuring spontaneous social categorization is the Who-said-What task (Taylor, Fiske, Etcoff, & Ruderman, 1978). It consists of two phases: learning phase and surprise recall task. In the learning or discussion phase, several “speakers” are displayed, who engage in a discussion. These speakers usually belong to one of two categories, e.g. female and male. In the subsequent surprise recall task, participants are asked “Who said that?” for each of the discussion statements and have to reassign the speakers to their respective statements. The main logic behind this paradigm is that more within-category errors relative to between-category errors are made when participants, confronted with a sentence they cannot reallocate to the correct speaker, use speaker category as proxy to increase their chance at guessing the correct speaker (Taylor et al., 1978). This is traditionally assessed by the error-difference measure that compares the sums of within- and between-category errors. A higher within-category error rate would be attributed to the application of social categories in the memory task. Although more recent and refined statistical methods to account for guessing processes have been proposed (Klauer & Wegener, 1998), the basic logic of the method is identical.

[3] We focused on the actual MDS dimension coordinates rather than the ratings to make sure that our results still hold when we or another scholar find a better candidate to interpret the MDS space. All results henceforth are thus independent of how to interpret the respective dimensions.

In Studies 3a–c we thus tested for each of the three dimensions whether the high, respectively low ends of the underlying continuum are used to construe superordinate categories. We used an adapted WSW-Task (to allow for multinomial processing tree analyses sensu Klauer & Wegener, 1998) by randomly drawing four occupations from pool of the 15 most extreme occupations at each end of the respective dimension. Eight persons introduced by first name and the randomly drawn occupation label then had a small talk conversation before the surprise recognition task appeared. We expected for all three dimensions more within category errors than between category errors.

4.1. Method

4.1.1. Participants

A different set of one hundred US-Americans took part in each of the three online studies on Mechanical Turk. An automatic filter only allowed them to participate if they had not participated in a similar previous study. If participants indicated at the end of the survey that they either saw their data not fit for analysis or that they had taken notes during the experiment, their data were not analyzed. Thus, for the first dimension, the data of 86 participants (35 men, 50 women, 1 did not indicate, M_{age} = 33.56, SD_{age} = 10.62) were included in the analysis. The data of 81 participants (43 men, 38 women, M_{age} = 32.25, SD_{age} = 9.22) were included in the analysis for the second dimension, while N = 90 (45 men, 45 women, M_{age} = 35.63, SD_{age} = 12.52) provided the sample for the third dimension.

4.1.2. The who-said-what task

To rule out that specific features of certain occupations or similarities between them would distort the observed effect of categorization along the dimension, speakers were randomly selected for each participant. For each of the three studies, the 15 highest and 15 lowest ranking occupational groups were selected for the respective dimension (see Table osm.1). For each participant, four occupations were randomly drawn from these pools of low- and high-ranking occupations respectively, to form the set of 8 speakers. We used a statement set and a distractor set, consisting of 48 demographics sentences each. Each speaker stated their name first, then each speaker stated their age and so on. Order of speakers and statements in each round was random.

4.1.3. Procedure

After giving demographic information, participants were instructed that they were about to see several people identified exclusively by their professions “meeting for the first time and engaging in small talk”. Participants saw all eight occupation titles and imagined each of them briefly to assign them some distinct visual features.

Then, the participants were presented with successive paired presentations consisting of a speaker (for 9 s) and a statement (after 1.5 s delay, for 7.5 s) each. After observing all 48 pairings (without inter-trial-breaks), participants moved on to the surprise recall task. All statements from the presentation phase and distractor set were shown in random order, and participants had to indicate “Who said that?” by ticking one of nine answer options (the eight speakers plus the option “None”). Participants also completed additional similarity ratings for the included speakers and dimension ratings for all 30 occupations in the pool of occupations. These largely replicated the findings of Study 1 and will not receive further notice. Finally, participants were debriefed and had to indicate whether they considered their own data fit for analysis or whether they had taken notes during the presentation phase.

4.2. Results and discussion

Comparing within-category errors and between-category errors (corrected for the higher base-rate likelihood to randomly make such an error; Taylor et al., 1978) yielded significantly more within-category errors in the first dimension, $F(1,85) = 9.25$, $p = .003$, the second dimension, $F(1,80) = 3.96$, $p = .050$, as well as the third dimension, $F(1,89) = 4.17$, $p = .044$ (Fig. 2).

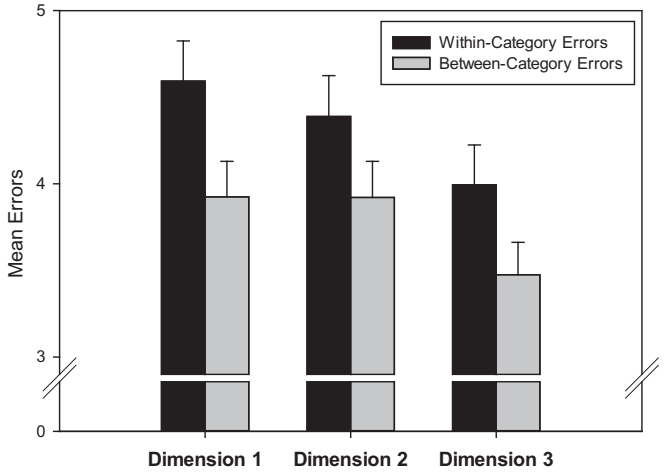

Fig. 2. Within- and between category Errors (+SE) in Who-Said-What paradigm for categories randomly sampled from top and low 10% on dimension 1 (Study 3a), dimension 2 (Study 3b), and dimension 3 (Study 3c) according multi-dimensional scaling in Study 1a. Between-category errors are corrected for higher base rate likelihood to make such an error (four instead of three possibilities to make an error).

To account for potential guessing artefacts, we also computed the multinomial processing tree (MPT) analysis by means of the R package TreeBUGS (Heck, Arnold, & Arnold, 2018). In Study 3a the model did not show appropriate model fit when the parameter for low dimension 1 categorization strength (d_b) was free to vary, but fit when it was restricted to 0 ($T_1^{observed} = 0.077$, $T_1^{predicted} = 0.082$, $p = .54$, $T_2^{observed} = 19.44$, $T_2^{predicted} = 13.94$, $p = .22$). In Study 3b the standard model restricting $d_a = d_b$ fit the data best ($T_1^{observed} = 0.084$, $T_1^{predicted} = .083$, $p = .48$, $T_2^{observed} = 10.21$, $T_2^{predicted} = 14.09$, $p = .73$). In Study 3c the model did not show appropriate model fit when the parameter for high dimension 3 categorization strength (d_a) was free to vary, but fit when it was restricted to 0 ($T_1^{observed} = 0.097$, $T_1^{predicted} = 0.075$, $p = .33$, $T_2^{observed} = 32.22$, $T_2^{predicted} = 13.11$, $p = .01$). As customary, results of the best fitting model are reported. The MPT results largely confirmed the classical error measure with the unexpected finding that for two of the three dimensions, only one of the two categories showed significant categorization strength (as indicated by categorization parameter d), whereas the opposite pole did not. Specifically, the MPT analysis suggested that participants categorized occupations high on dimension 1, $d_a = 0.49$ [95% Credibility Interval:0.31;0.64], but not low on that dimension, they categorized occupations along dimension 2, $d = 0.34$ [95% Credibility Interval:0.11;0.50], and they categorized occupations low on dimension 3, $d_b = 0.27$ [95% Credibility Interval:0.03;0.53], but no those high on that dimension.

Although we currently have no explanation for the apparent asymmetry on dimensions 1 and 3, the MPT as well as the classic analyses suggest that the occupational dimensions are indeed utilized to form higher order categories. Together with the previous studies this leads us to summarize that humans not only meaningfully and consensually place occupational groups on a “map” but they use these stereotypes to make sense of the world by forming larger superordinate categories. This is all the more surprising as our categories (high vs. low on a hypothetical dimension) are far more abstract than most previously employed categories in that paradigm. Primitive categories like race (e.g., Biernat & Vescio, 1993; Stangor, Lynch, Duan, & Glas, 1992) or gender (Stangor et al., 1992) are not only much more salient in everyday encounters both also come along with perceptual cues. The fact that all three dimensions produced categorization at least on one end on the dimension is thus all

the more remarkable. One remaining questions, however, concerns the consequences for actual social perception. We argue that the valence of neighboring occupations matter and predict contagious effects of valence information through social space.

5. Study 4

There is an abundance of research in the intergroup attitude domain that shows exactly such contagion, transfer or generalization effects. The positive effect of contact with an outgroup does not end at the boundaries of that group but generalizes to other outgroups (e.g., Pettigrew, 1997; Tausch et al., 2010), particularly if these outgroups are similar to the contacted one (Harwood, Paolini, Joyce, Rubin, & Arroyo, 2011). Such "lateral attitude change" (Glaser et al., 2015) is a well-documented phenomenon, not only for contact-induced attitude change but also for effects of evaluative conditioning (Walther, 2002) and affective learning (Verosky & Todorov, 2010). We thus tested the relevance of similarity (or proximity) in our occupational space by experimentally manipulating the valence of one target group and measuring the effect on all other occupational groups as a function of their similarity to the target group. The dependent variables in all studies were perceived valence of the other groups, operationalized as likeability or valence of expectations. Likeability as a communal trait has frequently not only been argued, but also been shown to be an almost perfect correlate of valence proper (with correlations up to 0.93 between communal traits like likeability and positivity ratings; see Imhoff & Koch, 2017). We chose police officers (Study 4a) and legislators (Study 4b) as target groups as there are numerous examples of very positive as well as very negative attitudes towards them that made it likely that both negative and positive experiences can be made salient experimentally.[4] To bolster the representativeness of our approach, in a final preregistered study, we sampled 28 target groups along the three dimensions and tested the distribution of the predicted contagion effects across these 28 conditions (Study 4c).

5.1. Study 4a

In Study 4a we tested whether an experimental manipulation of police officers' perceived valence affects other occupations' perceived likeability in a contagious way (likeability as a communal trait can be considered an item of social evaluation; Abele & Wojciszke, 2007; Imhoff & Koch, 2017). As we were interested in generalization effects (in contrast to displacement effect; see Glaser et al., 2015), we first piloted a manipulation to successfully change the evaluation of police officers.

5.1.1. Method

5.1.1.1. Pilot study. To be able to test a contagion effect it was crucial to experimentally manipulate participants' attitude towards police officers. To evoke the respective evaluation we presented participants in a pilot study ($N = 125$ MTurk workers; 53 women, 72 men; $M_{age} = 36.4$ years, $SD_{age} = 11.3$) with screenshots from two different Facebook group sites "A Warmer Shade of Blue. Stories About Good Things Cops Do" (positive evaluation), resp. "Stop Police Brutality" (negative evaluation) and asked them to write down an example of how they had a similar experience with police to the ones reported on the website. They then wrote how easy it was for them to come up with such an example (from 1 to 11) and rated police officers on six warmth-related items (e.g., trustworthy, sincere, benevolent; $\alpha = 0.969$) on a 11-point scale from 0 to 10.

Although there was a small effect of higher impressions of warmth after reading about and reporting positive experiences, $M = 6.21$, $SD = 2.55$, than after reading and reporting negative experiences, $M = 5.31$, $SD = 2.32$, $t(123) = 2.07$, $p = .041$, it became apparent from their essays that many participants had difficulties of generating concordant experiences (e.g., "I do not have any experiences with law enforcement."). This difficulty was also reflected in the fact that a relatively large proportion of participants (31.2%) experienced great difficulty generating examples as exhibited in extremely low ease ratings of 1 or 2 on the 11-point scale. Importantly, and in line with ease-of-retrieval heuristics (Schwarz et al., 1991) this perceived ease moderated the effect of experimental condition: The explained variance in warmth by effect-coded experimental condition, $B = 0.37$, $SE = 022$, $p = .096$, and standardized ease, $B = 0.34$, $SE = 0.22$, $p = .127$, $R^2 = 0.05$, $p = .038$, was significantly increased by the inclusion of their interaction, $B = 0.88$, $SE = 0.21$, $p < .001$, $\Delta R^2 = 0.12$, $p < .001$. The manipulation only worked as intended for those who experienced a relative ease generating consistent examples and even backfired for those who had great difficulties (Fig. 3, left panel). For the main study we thus slightly adapted our experimental manipulation.

5.1.2. Main study

5.1.2.1. Independent variable. To manipulate the valence of police officers we used the manipulation described in the pilot study with some changes. First, as the attribution to difficulty might just have been an excuse for participants who wanted to click fast through the study, the "continue" button appeared only after one minute. Second, we stressed that the generated experience did not have to be first-hand but could stem from friends or media.

5.1.2.2. Dependent variable. To estimate the effect of this manipulation on the perception of all occupational groups, participants indicated the perceived warmth of all 150 occupational groups (including police officers) on the same scale as in the pilot study in random order. To save time, however, they did not complete each of the scale items separately, but completed a single composite measure ranging from 1 (untrustworthy, dishonest, threatening, cold, repellent, unfriendly, and egoistic) to 10 (trustworthy, sincere, benevolent, warm, likable, friendly, and altruistic).

5.1.2.3. Participants. In total, $N = 226$ US-based MTurk workers (121 men, 105 women; 80.5% White) between the age of 18 and 68 ($M = 35.67$, $SD = 11.82$) completed a study for small monetary compensation.

5.1.2.4. Procedure. The study allegedly consisted of two parts. The first part, participants were told, dealt "with media portrayals of occupations in the times of social networking. The occupation that was randomly selected for you is: police officers." Contingent on condition, they saw screenshots reflecting either positively or negatively on police and then had to come up with an example of their own (experienced or heard about through friends or media) that was consistent with the portrayal on the facebook page. They indicated how easy this was on the same scale as in the pilot study, before the second part of the study began on global impressions of occupational groups with the instruction to indicate their personal feelings and attitudes towards occupational groups based on gut.

5.1.3. Results and discussion

As expected, reading about and elaborating on negative examples of police conduct evoked less positive ratings of police officers on the warmth item, $M = 5.52$, $SD = 2.36$, than reading about and elaborating on positive examples, $M = 6.33$, $SD = 2.58$, $t(224) = 2.49$, $p = .014$. This main effect, however, was again moderated by the perceived ease of generating examples consistent with the description in the excerpts. Although, there were main effects of both the experimental condition, $B = 0.51$, $SE = 0.16$, $p = .002$, and perceived ease, $B = 0.54$, $SE = 0.16$, $p = .001$, $R^2 = 0.07$, $p < .001$, including the interaction,

[4] We had also conducted an additional study on lawyers but failed to achieve an experimental effect on their perceived warmth.

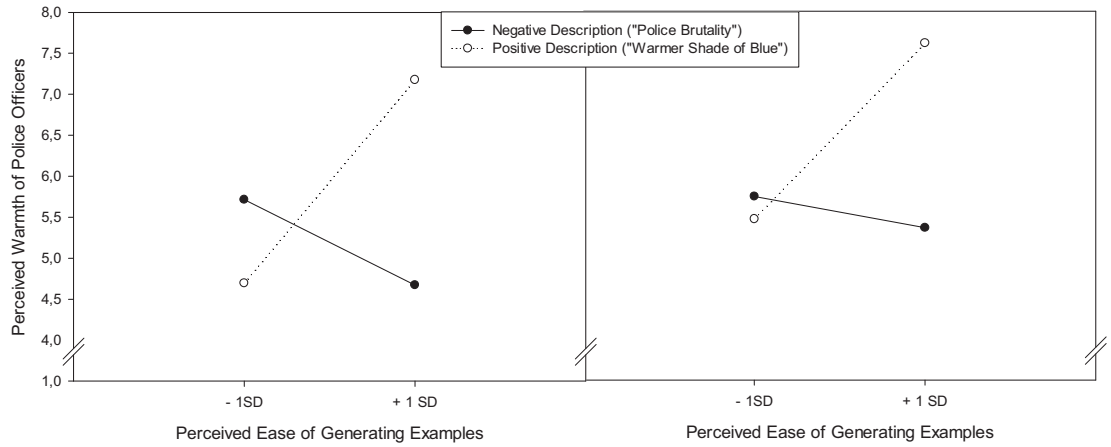

Fig. 3. Effect of Experimental manipulation of rated warmth of police officers as a function of perceived ease of generating examples in pilot study to Study 4a (left panel) and Study 4a (right panel).

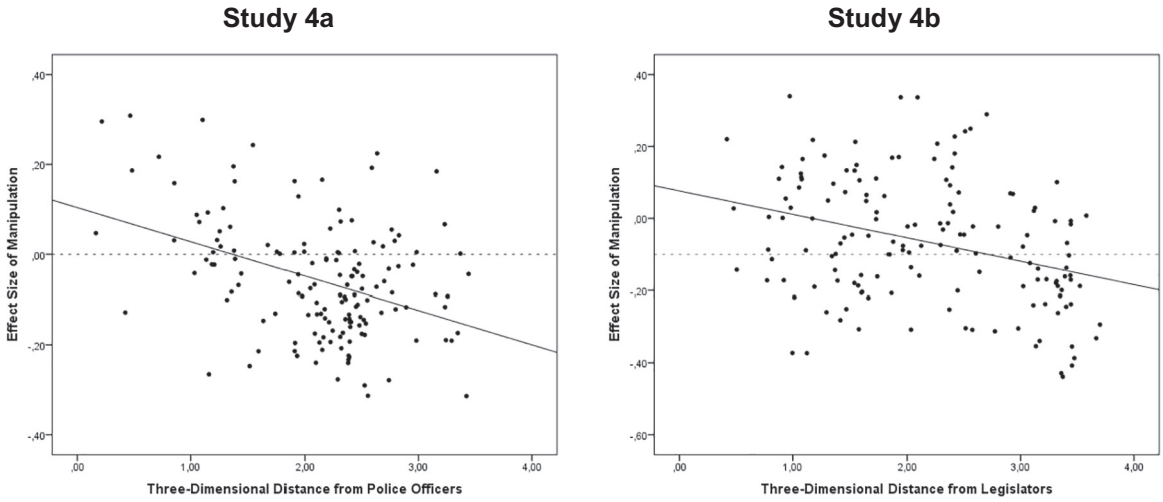

Fig. 4. Effect (in Cohen's *d*) of valence manipulation of focal group in Study 4a (police officers) and Study 4b (legislators) on valence of all others groups as a function of three-dimensional distance to focal group. Positive scores indicate valence contagion, negative values contrast effects.

$B = 0.63$, $SE = 0.16$, led to a significant increase in explained variance, $\Delta R^2 = 0.06$, $p < .001$. The plotted means again suggested that for people who experienced difficulty in generating examples, the manipulation had no effect (Fig. 3, right panel).

More relevant to the current study was whether this effect would also affect the perception of other occupational groups in a contagious way. To test this, we calculated for each occupation the effect the manipulation had on it, independent of statistical significance. Specifically, we calculated the mean warmth ratings for each occupation separately for each condition, as well as the pooled standard deviation. We then calculated a Cohen's *d* for each occupation and correlated these with the distance of each occupation to police officers in the MDS solution of Study 1a.

Results provided support for our hypothesis. For the 149 other occupations, the smaller the three-dimensional distance to police officers, the larger the effect, $r = -0.36$, $p < .001$ (Fig. 4). Looking at the unique contribution of the three dimensions, we then proceeded to conduct a multiple regression analysis with distance to police officers on each dimension as separate predictors of the effect the manipulation had on perceived warmth. Conjointly, the three distance vectors significantly predicted the effect, $R^2 = 0.16$, $p < .001$, with the strongest contribution from distance on the second dimension (interpreted as progressiveness in Study 1a), $\beta = -0.375$, $p < .001$, followed by the third (interpreted as sociability in Study 1a), $\beta = -0.158$, $p = .043$, and the first (interpreted as agency in Study 1a), $\beta = -0.152$, $p = .057$. Given that the second dimension was the only one in which police officers clearly deviated from the mean value (being seen as particularly conservative), it seems that the effect exhibited lateral contagion on the most salient dimension.

Although the previous analyses provided support for our hypotheses, they constitute a somewhat conservative test. For a considerable part of our participants, our experimental manipulation did not even change attitudes towards police officers and thus can hardly be expected to show a generalization effect beyond. As a less conservative test we repeated the analyses but excluded participants who scored more than one standard deviation below the mean on perceived ease (i.e., participants who scored a 1, 2, or 3 on the 10-point scale, 23% of the sample). As the effect of the manipulation was now stronger ($M_{neg} = 5.47$, $SD_{neg} = 2.34$, $M_{pos} = 6.96$, $SD_{pos} = 2.25$, $t(172) = 4.22$, $p < .001$), there was a greater chance to find contagion effects as well. The correlation between three-dimensional distance and effect size was now $r = -0.396$, $p < .001$, and the effect size was separately predicted by distance on dimension 1, $\beta = -0.204$, $p = .008$, dimension 2, $\beta = -0.369$, $p < .001$, and dimension 3, $\beta = -0.299$, $p < .001$, $R^2 = 0.22$, $p < .001$.

Our data thus suggest that evaluation of occupational groups is indeed contagious across the social space of occupational stereotypes. Neighboring groups are affected to a stronger degree than distant groups. This of course, is merely a first illustration and the effect may hinge on the exact group chosen. To bolster the generalizability of our claim, we conducted an almost identical study with a different target group.

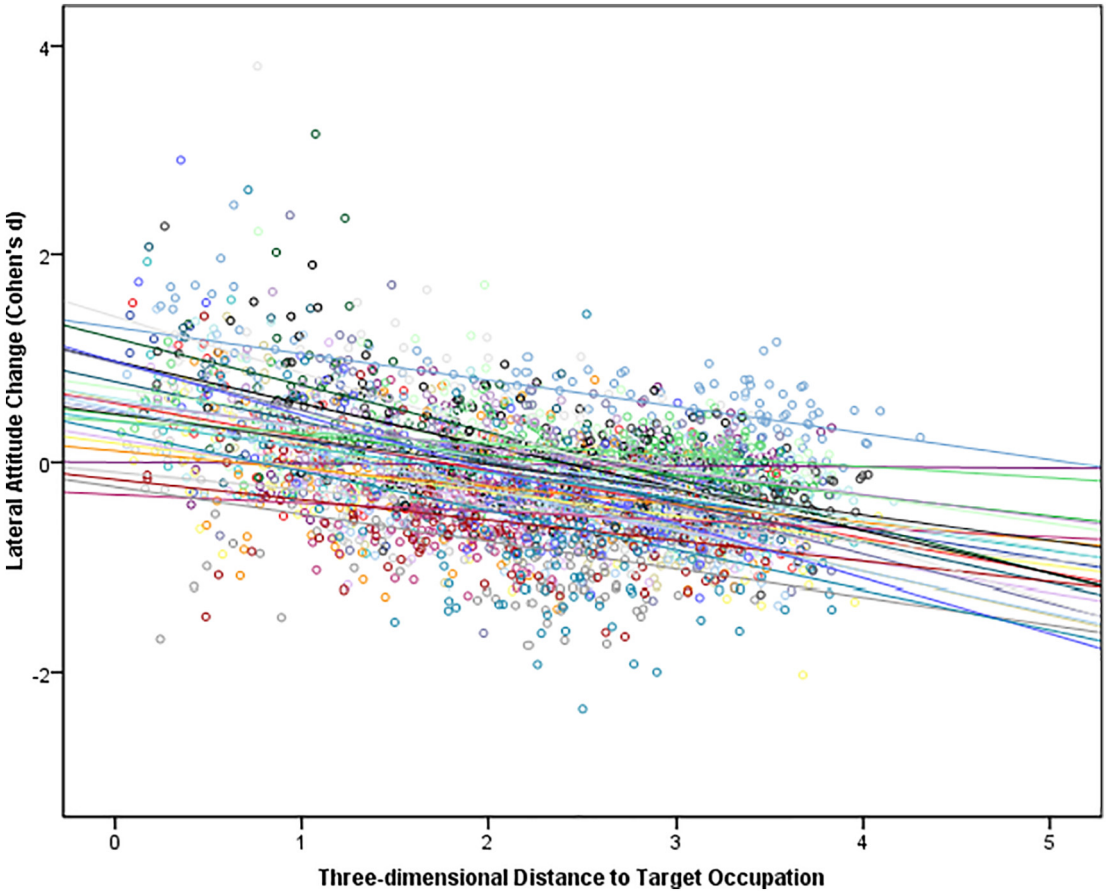

Fig. 5. Effect size (in Cohen's d) of valence manipulation of 28 target groups in Study 4c on valence of all others groups as a function of three-dimensional distance to respective focal group. Positive scores indicate valence contagion, negative values contrast effects.

5.2. Study 4b

Study 4b was a close replication of Study 4a in many regards, with the exception that we used a different occupational group: legislators or politicians. As this group scores particularly high on the first dimension (agency/status) this also allowed us to explore whether contagion effects were again strongest along the most salient dimension.

5.2.1. Method

5.2.1.1. Procedure. To manipulate the warmth of legislators, we followed the same procedure as in Study 6. Participants either read negative online reports about politicians (e.g., “Citizens against corrupt politicians”, “Whitehouse scandals”, or “sick of lying politicians”) or positive reports of politicians' prosocial behavior (“Politicians Care”) and gave fitting examples afterwards. Everything that followed was identical to Study 6: Participants rated the ease of generating the examples and rates all 150 occupations on a composite measure of warmth.

5.2.1.2. Participants. A total of 202 predominantly White (72.8%) MTurk workers (107 men, 95 women) ranging in age from 20 to 71 ($M = 35.28$, $SD = 11.80$) participated for small monetary compensation.

5.2.2. Results and discussion

As in Study 4a, we present the conservative analyses with all participants first and the analyses without participants who indicated very low ease in coming up with examples afterwards. The manipulation was successful, as after reporting negative experiences with politicians, participants saw them as less likeable, $M = 3.38$, $SD = 2.02$, than after reporting positive experiences, $M = 4.28$, $SD = 2.32$, $t(200) = 2.93$, $p = .004$. A small contagion effect was observable from a (non-significant) negative correlation between the effect this manipulation had on other occupations and the respective occupation's three-dimensional distance to legislators in the MDS space, $r = -0.15$, $p = .063$ (Fig. 4). When we included the separate distance estimates per dimension in a multiple regression predicting the effect size, $R^2 = 0.09$, $p = .004$, it became apparent that the contagion effect was predominantly driven by the third dimension, $\beta = -0.181$, $p = .026$, but unexpectedly attenuated by a reverse effect on the second dimension, $\beta = 0.189$, $p = .033$, with no significant prediction from distance on the first dimension, $\beta = -0.046$, $p = .594$.

As in Study 4a, we proceeded with the same analyses excluding those participants for whom the manipulation was difficult. The stronger effect on the evaluation of legislators ($M_{neg} = 3.47$, $SD_{neg} = 2.06$, $M_{pos} = 4.63$, $SD_{pos} = 2.44$, $t(138) = 3.04$, $p = .003$) also translated in a stronger overall contagion effect, $r = -0.33$, $p < .001$. Broken down by dimension, the first dimension now had the stronger contribution to the effect, $\beta = -0.249$, $p = .003$, than the third dimension, $\beta = -0.147$, $p = .058$, but the unexpected reverse effect for the second dimension was still there, $\beta = 0.197$, $p = .020$, $R^2 = 0.17$, $p < .001$.

Study 4b largely replicated the effect with two minor wrinkles. First, there was an unexpected effect of distance on the second dimension, whereby groups that were close to legislators on this dimension were particularly unaffected (or even in the reverse direction). At present, we do not understand well, why this happens, but it may be attributable to the fact that legislators were exactly in the center of the progressiveness dimension (0.00104 in the MDS coordinates), and centrality is typically a strong cue to warmth (Imhoff & Koch, 2017). Independent of this, however, it might also be raised that the overall contagion effect was

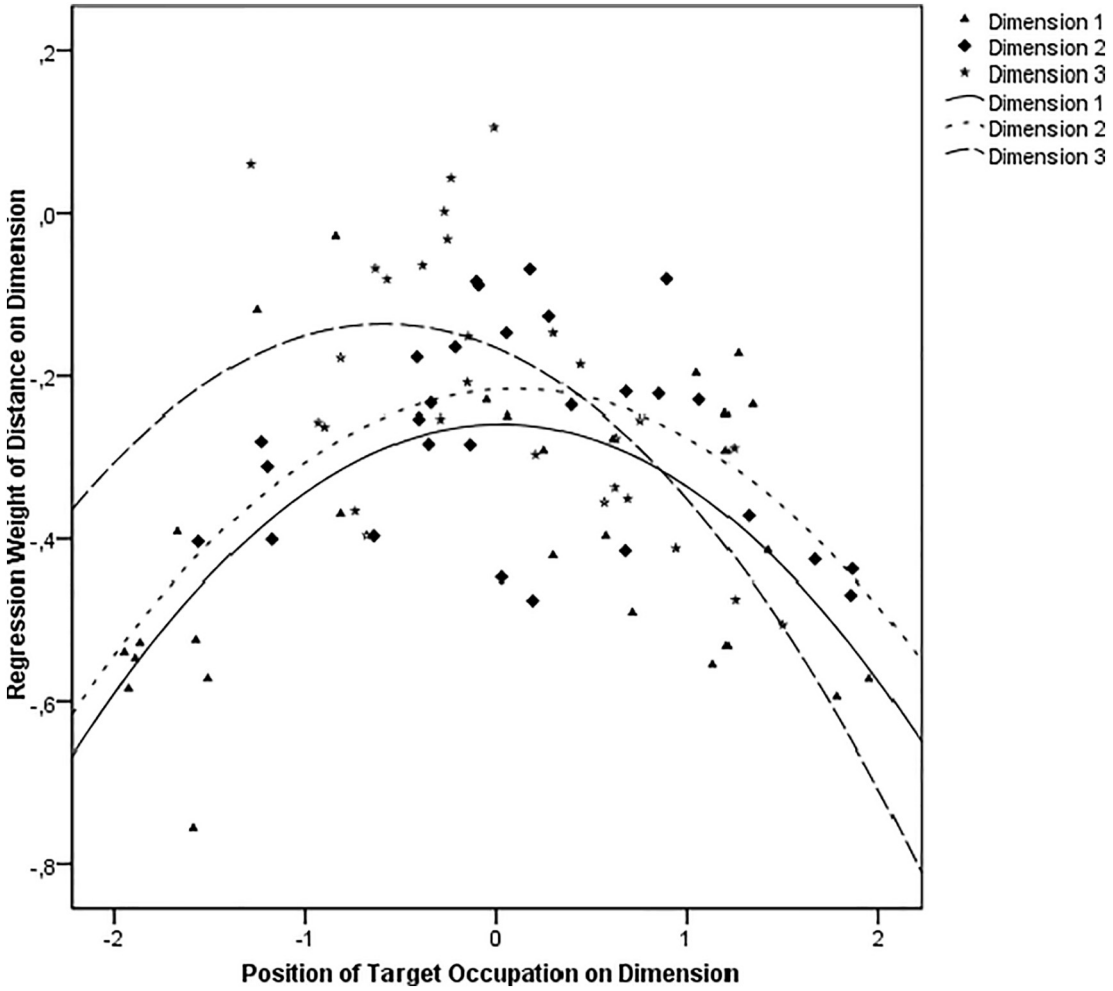

Fig. 6. Influence of distance to target occupation on three dimensions in Study 4c (as standardized regression weights) as a function of target dimension's position on the respective dimension.

not significantly different from zero ($p = .063$). There are two ways to address this. First, we are convinced that the less conservative way of analyses also reported here is actually a more accurate approach to the data, as it excludes participants who do not even show the basic effect that should then contagiously travel through the similarity space. In terms of lateral attitude change, there is a marked difference between showing an effect on other than the target attitude in the presence (generalization) or the absence (transference) of a focal effect on the target attitude. Both are theorized to follow different principles and show on different (e.g., explicit vs. implicit) level (Glaser et al., 2015). The fact that the contagion effect was very weak for the overall sample but clearly there ($p < .001$) when we filtered out those participants who described the manipulation as too difficult to yield a focal effect, suggests that the described contagious effect is restricted to generalization. Nevertheless, we refrained from presenting only this (in hindsight superior) analysis as we had not preregistered this exclusion criterion. Second, speaking to the broader evidence provided in Study 4, even if we only took the unfiltered, conservative estimates of the effect as real, the average effect is still meta-analytically different from zero, $r = -0.278$ [95%CI:-0.374;-0.181], $p < .001$. We thus are thus confident that the effect reported here is not spurious. To be sure, however, we conducted a large final study.

5.3. Study 4c

Our final study was designed to take care of three issues: first, and arguably most importantly, we pre-registered all of our analytic steps before collecting data (http://aspredicted.org/blind.php?x=tt4yx3). Second, we did not cherry-pick occupations that we saw as particularly potent to help the effect, but followed a sampling rational. Finally, we sampled occupations along the whole continuum of all three dimensions to allow a fair test of the influence of distance on all three dimensions.

5.3.1. Method

5.3.1.1. Target occupations. To generate a sample of target occupations without hand selecting seemingly fitting groups (and thereby potentially introducing researcher-based bias) we applied the same sampling criterion to all three dimensions: First, we sorted all 150 occupations along the respective dimension. Then we picked 10 equidistant occupations, center-aligned (i.e., number 3, 19, 35, 51, 67, 83, 99, 115, 131, 147). Among the 30 obtained occupations, two appeared twice, leaving us with 28 unique occupations.

5.3.1.2. Procedure. Attitudes towards the target occupation were manipulated for each occupation in an identical way. Participants were either assigned to the positive or the negative condition and read "Imagine you had a really terrible/wonderful experience with [target occupation]. Even worse, this was not the first time that you felt treated negatively/positively by [target occupation], but you have a continuous history of negative/positive interactions with [target occupation]. The most recent incident, however, was by far the worst/best and left a

persistent mark on you." Participants were then asked to estimate how negative vs. positive they expected experience with job holders of all 150 occupations (in random order) to be. Afterward, they were asked to remember the target occupation in an open-ended fashion, judged how pleasant the episode was they imagined (from −50 very negative to 50 very positive), indicated how easy it was to imagine such an experience on an 11-point scale, gave demographic information and judged their own data quality.

5.3.1.3. Participants. We had aimed for 800 participants (see pre-registration), but due to a technical error with missing completion codes a total of $N = 938$ MTurk workers (489 men, 442 women, 7 other; ranging in age from 18 to 77, $M = 35.74$, $SD = 10.86$) completed the study for $1. Applying the pre-registered exclusion criteria, 43 participants were excluded as they indicated that researchers should drop their data, 41 participants were excluded because their self-estimated ease to imagine the prescribed experience was three or lower. For the remaining 854 participants, two hypothesis-blind raters judged whether their open-ended memory of the target occupation was correct or not. The raters showed high agreement (98.6%, $\kappa = 0.95$, $p < .001$), the twelve cases where they did not agree were resolved by the first author. Based on this, another 131 participants were excluded for not remembering the correct target occupation. Finally, we z-standardized the evaluation of the target occupation itself and eliminated values over 1SD in the negative condition and below -1SD in the positive condition, yielding another eight exclusions. The final sample thus effectively consisted of 715 participants.

5.3.2. Results and discussion

As another precaution, we had preregistered to eliminate whole conditions for which the manipulation did not even have an effect of the target occupation itself. This was not the case, as all 28 condition produced significant effect ($ps < .003$), with a meta-analytic effect size of $d = 3.27$, 95% *CI* [2.98; 3.55]. Separately for the 28 target occupation condition we then calculated an effect size for each of the 149 other occupations, indicating the effect the manipulation had on the respective other occupations (e.g., "How do negative vs. positive experiences with actors affect how negative vs. positive are encounters with veterinarians expected to be?"). This yielded 149 *d*-scores for each of the 28 conditions (we excluded the evaluation of the target occupation as this would not constitute lateral attitude change). These *d*-scores were correlated with the three-dimensional distance to the target occupation in each of the 28 conditions. Lateral attitude change as a function of proximity would be reflected in a negative correlation as the contagious effect would be smaller (or a contrast effect), the greater the distance. The empirical correlations were all negative and with the exception of two (audiologists, massage therapist) significant, $p < .003$ (Fig. 5). Testing the distribution of these (*r*-to-*z*-transformed) correlations against zero yielded a significant result, $t(27) = 13.23$, $p < .001$, Cohen's $d = 2.50$, with an average (*z*-to-*r*-transformed) correlation of $r = -0.54$.

On an exploratory base, we also analyzed for the 28 conditions distance whether each dimension contributed uniquely to the effect. To this end, we calculated 28 regression analyses with three indicators of unidimensional distance as predictors and the *d* as criteria. We then tested the distribution of standardized regression weights across the 28 regressions against zero. Distance on the first dimension had the strongest impact on valence of expectations, $\overline{\beta} = -0.39$, $t(27) = 11.96$, $p < .001$, $d = 2.26$, but distance on the second, $\overline{\beta} = -0.28$, $t(27) = 11.17$, $p < .001$, $d = 2.11$, and third dimension, $\overline{\beta} = -0.21$, $t(27) = 6.90$, $p < .001$, $d = 1.30$, were equally significant, incremental predictors. We had speculated a priori whether contagious effects would be particularly pronounced for the most salient dimension, i.e. the dimension on which the respective focal group was either particularly high or low. The present data gave us a chance to at least explore this possibility by testing whether there was a curvilinear relation between the focal occupation's position on a given dimension and the weight this dimension had (i.e., the predictive power of distance on that dimension on the contagious effect). Supporting this notion, there was a curvilinear effect for the first, $\beta = -0.59$, $p < .001$ (linear: $\beta = 0.03$, $p = .872$), the second, $\beta = -0.61$, $p < .001$ (linear: $\beta = 0.10$, $p = .562$) and descriptively also for the third dimension, $\beta = -0.30$, $p = .078$ (linear: $\beta = -0.46$, $p = .011$; Fig. 6).

6. General discussion

Throughout a series of studies, we have empirically established a data-driven model of occupational stereotypes and its downstream consequences on processes of categorization and lateral attitude change. It is particularly remarkable that two models that started from different sets of stimuli (both exhaustive lists that cover the whole population of job titles in the respective context) in two different national contexts largely converged (with some minor nuance) in suggesting that people predominantly discriminate between occupations based on agency and progressiveness (to a lesser extent sociability, respectively communion). The first dimension does not seem surprising as status is a fundamental human motive (Anderson, Hildreth, & Howland, 2015), and hierarchical organization is ubiquitous in the professional world (Magee & Galinsky, 2008). The second dimension, however, is somewhat less expected. People seem to be particularly alert to differences in presumed political orientations of occupational groups. Choosing a profession is one of the most consequential choices we can make to express ourselves and our values. Some people choose to put their labor in the upholding and protection of the status quo (e.g., firefighter, police officers), whereas others express their openness to change in artistic vanguard (e.g., musicians, artists, dancers). These fundamental values of conservation vs. change have recently been discussed in connection to the core human strategies of exploitation (of the status quo) and exploration as two opposite strategies to effectively deal with the environment (Koch, Imhoff, et al., 2016).

In summary, without constraining the design, the stimuli or participants in any way, the current research suggests that stereotype dimensions about occupations align surprisingly well with stereotypes about social groups (Koch, Imhoff, et al., 2016). Other research traditions have often taken this for granted and just imposed dimensions presumed to be central for social group perceptions (Fiske et al., 2002) to the domain of occupations (Fiske & Dupree, 2014). The current research provides initial evidence for the notion that this may indeed be adequate, but for different dimensions than warmth and competence.

Furthermore, we did not only observe two largely identical descriptions of how people saw others based on their occupations, we also observed that the characteristics associated with the job are indeed used to infer person characteristics (Study 2), further corroborating the relevance of these similarities. The last two studies provided insight in how the position of stimuli on the model influences basic cognitive processes. First, people lump together seemingly similar groups as superordinate categories of occupations high or low on the respective dimensions. Second, proximity in the three-dimensional space predicts whether valence spills over from one occupation to another with neighbors being more affected than occupations in greater distance. Negative attitudes towards artists are thus likely to affect one's attitude towards dancers but not towards midwifes (or even in the opposite direction). All these studies point to the relevance of the dimensional model and they stand independent of how we interpret the dimensions that run through the space.

Stereotypes are a fascinating topic of research on ideas and mental schemas. Part of their fascination, however, is derived from the lingering suspicion that these stereotypes guide our interpretation and behavior in more realistic settings than psychology studies. Our research paves the way for more theory-driven confirmatory tests of such behavior across a variety of settings. For two national contexts, we have ratings on the most relevant stereotype dimensions and inter-occupational similarity readily available for other researchers to adopt for fascinating future research questions: Is occupational similarity a

predictor of successful matches in the mating market? Does a match between personality and occupational stereotype predict job satisfaction? All raw data is available on our OSF project page (https://osf.io/4rmnv/?view_only=fee3f4f700b54af3ab8ee45e299507d0) and might prove useful to elucidate these issues in the future.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jesp.2018.04.001.

Open practices

All studies in this article earned Open Materials and Open Data badges for transparent practices. Materials, data, and supplemental analyses are available at https://osf.io/4rmnv/. The final experiment (Study 4c) earned a Preregistered badge for preregistering hypotheses and data analysis plan prior to data collection. The preregistration is available at https://aspredicted.org/8eg66.pdf. All preregistered analyses for Study 4c were carried out as planned and reported in the article, no other preregistrations were made for studies reported in this article.

Acknowledgments

This research was supported by a University of Cologne Advanced Postdoc Grant to Roland Imhoff. We are greatly indebted to Ron Dotsch – the idea for this research was born over a joint lunch half way between Nijmegen and Cologne. Ron, you will be greatly missed in academia! We would also like to thank Maike Hermann, Anna Monschau and Felix Zimmer for help with Study 4. All materials, data and code are available under https://osf.io/4rmnv.

References

Abele, A. E., & Petzold, P. (1998). Pragmatic use of categorical information in impression formation. *Journal of Personality and Social Psychology, 75*, 347–358.
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.
Alves, H., Koch, A. S., & 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.
Anderson, C., Hildreth, J. A., & Howland, L. (2015). Is the desire for status a fundamental human motive? A review of the empirical literature. *Psychological Bulletin, 141*, 574–601.
Biernat, M., & Vescio, T. K. (1993). Categorization and stereotyping: Effects of group context on memory and social judgment. *Journal of Experimental Social Psychology, 29*, 166–202.
Brunswik, E. (1955). Representative design and probabilistic theory in a functional psychology. *Psychological Review, 62*, 193–217.
Deng, C. P., Armstrong, P. I., & Rounds, J. (2007). The fit of Holland's RIASEC model to US occupations. *Journal of Vocational Behavior, 71*, 1–22.
Dunham, Y., Stepanova, E. V., Dotsch, R., & Todorov, A. (2015). The development of race-based perceptual categorization: Skin color dominates early category judgments. *Developmental Science, 18*, 469–483.
Eagly, A. H., Wood, W., & Diekman, A. B. (2000). Social role theory of sex differences and similarities: A current appraisal. In T. Eckes, & H. M. Trautner (Eds.). *The developmental social psychology of gender* (pp. 123–174). Mahwah, NJ: Lawrence Erlbaum.
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.
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, 111*, 13593–13597.
Glaser, T., Dickel, N., Liersch, B., Rees, J., Süssenbach, P., & Bohner, G. (2015). Lateral attitude change. *Personality and Social Psychology Review, 19*, 257–276.
Glick, P., Zion, C., & Nelson, C. (1988). What mediates sex discrimination in hiring decisions? *Journal of Personality and Social Psychology, 55*, 178–186.
Harwood, J., Paolini, S., Joyce, N., Rubin, M., & Arroyo, A. (2011). Secondary transfer effects from imagined contact: Group similarity affects the generalization gradient. *British Journal of Social Psychology, 50*, 180–189.
Hayes, A. F., & Krippendorff, K. (2007). Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures, 1*, 77–89.
Heck, D. W., Arnold, N. R., & Arnold, D. (2018). TreeBUGS: An R package for hierarchical multinomial-processing-tree modeling. *Behavior Research Methods, 50*(1), 264–284.
Holland, J. L. (1959). A theory of vocational choice. *Journal of Counseling Psychology, 6*, 35–45.
Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities and work environments.* Odessa, FL, US: Psychological Assessment Resources.
Hout, M. C., & Goldinger, S. (2016). SpAM is convenient but also satisfying. *Journal of Experimental Psychology: General, 145*, 383–387.
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.
Hout, M. C., Papesh, M. H., & Goldinger, S. D. (2013). Multidimensional scaling. *Wiley Interdisciplinary Reviews: Cognitive Science, 4*, 93–103.
Imhoff, R., & Burder, M. (2014). Speaking (un-)truth to power: Conspiracy Mentality as a generalised political attitude. *European Journal of Personality, 28*, 25–43.
Imhoff, R., & Koch, A. (2017). How orthogonal are the big two of social perception? On the curvilinear relation between agency and communion. *Perspectives on Psychological Science, 12*, 122–137.
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.
Jaworska, N., & Chupetlovska-Anastasova, A. (2009). A review of multidimensional scaling (MDS) and its utility in various psychological domains. *Tutorial in Quantitative Methods for Psychology, 5*, 1–10.
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.
Klauer, K. C., & Wegener, I. (1998). Unraveling social categorization in the "Who said what?" paradigm. *Journal of Personality and Social Psychology, 75*, 1155–1178.
Koch, A., Alves, H., Krüger, T., & Unkelbach, C. (2016). A general valence asymmetry in similarity: Good is more alike than bad. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*, 1171–1192.
Koch, A., Imhoff, R., Dotsch, R., Unkelbach, C., & Alves, H. (2016). The ABC of stereotypes about groups: Agency/socioeconomic success, conservative–progressive beliefs, and communion. *Journal of Personality and Social Psychology, 110*, 675–709.
Koch, A., Kervyn, N., Kervyn, M., & Imhoff, R. (2018). Studying the cognitive map of the U.S. states: Ideology and prosperity stereotypes predict interstate prejudice. *Social Psychological and Personality Science*. http://dx.doi.org/10.1177/1948550617715070 (in press).
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.
Koenig, A. M., Eagly, A. H., Mitchell, A. A., & Ristikari, T. (2011). Are leader stereotypes masculine? A meta-analysis of three research paradigms. *Psychological Bulletin, 137*, 616–643.
Lammers, J., Koch, A., Conway, P., & Brandt, M. J. (2017). The political domain appears simpler to the politically extreme than to political moderates. *Social Psychological and Personality Science, 8*, 612–622.
Magee, J. C., & Galinsky, A. D. (2008). 8 social hierarchy: The self-reinforcing nature of power and status. *The Academy of Management Annals, 2*, 351–398.
Peabody, D. (1967). Trait inferences: Evaluative and descriptive aspects. *Journal of Personality and Social Psychology, 7*, 1–18.
Pettigrew, T. F. (1997). Generalized intergroup contact effects on prejudice. *Personality and Social Psychology Bulletin, 23*, 173–185.
Pratto, F., Stallworth, L. M., Sidanius, J., & Siers, B. (1997). The gender gap in occupational role attainment: A social dominance approach. *Journal of Personality and Social Psychology, 72*, 37–53.
Rotter, J. B., & Stein, D. K. (1971). Public attitudes toward the trustworthiness, competence, and altruism of twenty selected occupations. *Journal of Applied Social Psychology, 1*, 334–343.
Rounds, J., Armstrong, P. I., Liao, H. Y., Lewis, P., & Rivkin, D. (2008). *Second generation occupational interest profiles for the O* NET system: Summary*. The National Center for O* NET Development.
Rutjens, B. T., & Heine, S. J. (2016). The immoral landscape? Scientists are associated with violations of morality. *PLoS ONE, 11*, e0152798.
Schwartz, S. H., & Bilsky, W. (1987). Toward a universal psychological structure of human values. *Journal of Personality and Social Psychology, 53*, 550–562.
Schwarz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., & Simons, A. (1991). Ease of retrieval as information: Another look at the availability heuristic. *Journal of Personality and Social Psychology, 61*, 195–202.
Sidanius, J., Liu, J. H., Pratto, F., & Shaw, J. S. (1994). Social dominance orientation, hierarchy attenuators and hierarchy-enhancers: Social dominance theory and the criminal justice system. *Journal of Applied Social Psychology, 24*, 338–366.
Stangor, C., Lynch, L., Duan, C., & Glas, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology, 62*, 207–218.
Tausch, N., Hewstone, M., Kenworthy, J. B., Psaltis, C., Schmid, K., Popan, J. R., ... Hughes, J. (2010). Secondary transfer effects of intergroup contact: Alternative accounts and underlying processes. *Journal of Personality and Social Psychology, 99*, 282.
Taylor, S. E., Fiske, S. T., Etcoff, N. L., & Ruderman, A. J. (1978). Categorical and contextual bases of person memory and stereotyping. *Journal of Personality and Social Psychology, 36*, 778–793.
Verosky, S. C., & Todorov, A. (2010). Generalization of affective learning about faces to perceptually similar faces. *Psychological Science, 21*, 779–785.
Walther, E. (2002). Guilty by mere association: Evaluative conditioning and the spreading attitude effect. *Journal of Personality and Social Psychology, 82*, 919.
Wilbourn, M. P., & Kee, D. W. (2010). Henry the nurse is a doctor too: Implicitly examining children's gender stereotypes for male and female occupational roles. *Sex Roles, 62*, 670–683.
Wish, M., Deutsch, M., & Biener, L. (1970). Differences in conceptual structures of nations: An exploratory study. *Journal of Personality and Social Psychology, 16*, 361.
Young, F. W., Takane, Y., & Lewyckyj, R. (1978). ALSCAL: A nonmetric multidimensional scaling program with several individual-differences options. *Behavior Research Methods, 10*, 451–453.