Empirical Research Paper

Personality and Social Psychology Bulletin
1–17

DOI: 10.1177/01461672241266651
journals.sagepub.com/home/pspb
Sage

Unconstrained Descriptions of Facebook Profile Pictures Support High-Dimensional Models of Impression Formation

Paul Connor[1], Gandalf Nicolas[2], Stephen Antonoplis[3], and Alex Koch[4]

Abstract

Dominant models of impression formation focus on two fundamental dimensions: a horizontal dimension of warmth/communion/trustworthiness and a vertical dimension of competence/agency/dominance. However, these models have typically been studied using theory-driven methods and stimuli of restricted complexity. We used a data-driven approach and naturalistic stimuli to explore the latent dimensions underlying >300,000 unconstrained linguistic descriptions of 1,000 Facebook profile pictures from 2,188 participants. Via traditional (Exploratory Factor Analysis) and modern (natural language dictionaries, semantic sentence embeddings) approaches, we observed impressions to form with regard to the horizontal and vertical dimensions and their respective facets of sociability/morality and ability/assertiveness, plus the key demographic variables of gender, age, and race. However, we also observed impressions to form along numerous further dimensions, including adventurousness, conservatism, fitness, non-conformity, and stylishness. These results serve to emphasize the importance of high-dimensional models of impression formation and help to clarify the content dimensions underlying unconstrained descriptions of individuals.

Keywords
impression formation, person perception, stereotyping, natural language processing

Received November 30, 2023; revision accepted June 12, 2024

If social psychology has central axioms, one is sure that people form and act upon impressions of others (Fiske & Taylor, 1991). Impression formation occurs constantly and spontaneously, and guides all aspects of our social lives, shaping whether, how, and why we interact with others. The nature of these impressions is therefore of fundamental interest to social psychologists, with ongoing questions including the number of dimensions underlying impression formation as well as their content, relationships, and relative priority (Abele et al., 2021).

Recent research applied natural language processing techniques to unconstrained descriptions of societal groups and found that people spontaneously described groups along a greater number of dimensions than captured by prior models (Nicolas et al., 2022). Extending on this, we used additional natural language processing techniques to model unconstrained descriptions of individuals (Facebook profile pictures). Overall, our results generalize high-dimensional impression formation (Nicolas et al., 2022) from labels of groups to photos of individuals, and help to elucidate both the need to consider and the nature of high-dimensional impression formation.

The Two Fundamental Dimensions

Dominant impression formation models posit that people evaluate others primarily along two fundamental dimensions: a horizontal dimension focused on social cooperation, variously labeled warmth, communion, or trustworthiness, and a vertical dimension focused on goal achievement, variously labeled competence, agency, or dominance (Abele et al., 2016; Fiske et al., 1999; Judd et al., 2005; Oosterhof & Todorov, 2008).[1] These dimensions are thought to provide the information needed most when interacting with others: their intentions toward us (horizontal), and their ability to carry out those intentions (vertical; Fiske et al., 2002). Impressions along these dimensions have been linked to

[1]Stevens Institute of Technology, Hoboken, NJ, USA
[2]Rutgers University, Piscataway, NJ, USA
[3]University of California, Riverside, USA
[4]The University of Chicago Booth School of Business, IL, USA

Corresponding Author:
Paul Connor, Department of Quantitative Social Science, Stevens Institute of Technology, 1 Castle Point Rd, Hoboken, NJ 07030, USA.
Email: pconnor@stevens.edu

neural activity (L. T. Harris & Fiske, 2006), affective states (Cuddy et al., 2007), and behavior (Cuddy et al., 2007), and are universally used to differentiate social groups (Cuddy et al., 2008; Fiske et al., 2007). Recently, an emerging consensus has proposed that these fundamental dimensions can each be subdivided into two distinct facets, the horizontal dimension into friendliness and morality (with morality the better predictor of overall evaluations and interpersonal outcomes; Brambilla et al., 2012; Ellemers, 2017), the vertical dimension into the facets of ability and assertiveness (Abele et al., 2016, 2021).

Theory-Driven Foundations

One contemporary critique of dominant two-dimensional models is that they were developed and tested in top-down, theory-driven ways, which may have constrained empirical results. For example, in first distinguishing the horizontal and vertical dimensions, Asch (1946) demonstrated that describing targets as "warm" or "cold" affected some trait perceptions (e.g., generous/ungenerous), but not others (e.g., strong/weak). Yet, Asch's (1946) decision to manipulate the traits of warmth/coldness was explicitly based on his intuition that "not all qualities have the same weight . . . some are felt to be basic, others secondary" (p. 262). Similarly, the choice of which trait perceptions to measure was based on ". . . an informal sense of what was fitting" (p. 262). Thus, although his work showed that people distinguish between the horizontal and vertical dimensions, Asch's methods could not rule out the possibility that other dimensions might be equally important in impression formation.

Extending on Asch's work, Rosenberg and colleagues (1968) asked participants to sort a large list of traits while endeavoring to ". . . put those traits which tend to go together . . . into the same category" (p. 285). Applying multidimensional scaling to traits' resulting similarity scores, the authors also found support for a two-dimensional horizontal/vertical model. Again, however, the selected list of traits was constrained by prior theory, with 39 of the 64 traits chosen ". . . to compare some of the results obtained in this study with the findings reported by Asch (1946, p. 284)".

Fiske and colleagues (1999) were also guided by strong theoretical priors regarding the importance of the horizontal and vertical dimensions in developing their Stereotype Content Model (SCM). Early studies chose rating scales with the explicit goal of measuring perceptions of these two dimensions: "participants rated these groups on 27 trait adjectives, reflecting positive and negative aspects of warmth and competence" (p. 479).

Other influential models have also treated the centrality of the horizontal and vertical dimensions as a theoretical foundation and sought to understand the relative priority of each dimension (Abele & Wojciszke, 2007, 2014), and how, when, and why perceptions on each dimension might function in a compensatory fashion (Yzerbyt et al., 2005, 2008). As a result, empirical work on these models has seldom considered alternative or higher dimensional models of impression formation.

Face perception researchers arrived at a similar two-dimensional model, proposing that individual faces are evaluated primarily in terms of trustworthiness and dominance (Oosterhof & Todorov, 2008). This work was more data-driven than the research described above, starting with a large sample of unconstrained descriptions of facial photographs. However, it also incorporated theory-driven methods capable of constraining empirical results. First, the initial unconstrained descriptions were sorted by researchers into 14 categories, a subjective process by which nearly 1/3 of the descriptions were omitted. Second, an extra category–dominance–was added due to its centrality in prior work (e.g., Wiggins, 1979). Third, ratings on these categories were submitted to Principal Components Analysis (PCA), which constrains dimensions to be uncorrelated (Widaman, 2018). Thus, although models of face perception incorporated data-driven methods, they too were developed in ways that may have allowed experimenters' prior expectations to shape results.

Beyond Two Dimensions: Data-Driven Methods and Complex Naturalistic Stimuli

Noting how rarely the dominant two-dimensional model of impression formation had been tested in a data-driven manner, Koch and colleagues (2016) proposed a novel approach to fill this empirical gap. Instead of asking participants to rate social groups on pre-selected traits, they asked participants to rate groups in terms of their perceived similarity and applied multidimensional scaling to the resulting similarity ratings. Surprisingly, although participants appeared to spontaneously differentiate groups according to the vertical dimension, they did not appear to do so along the horizontal dimension, instead, differentiating groups according to their perceived progressive or conservative beliefs.

Subsequent work by these authors clarified that impressions do arise spontaneously along the horizontal dimension but depend on individuals' idiosyncratic perceptions of self/other similarity in terms of agency and beliefs (Koch et al., 2020). Yet even after acknowledging the importance of the horizontal dimension, Koch and colleagues' Agency-Beliefs-Communion (ABC) model posed a challenge to dominant two-dimensional models. Not only did it suggest that perceptions of beliefs may be equally central to impression formation as the fundamental two dimensions, but it also suggested that this had been overlooked in prior work due to an over-reliance on theory-driven methods.

Other researchers have also used data-driven methods to look beyond two-dimensional models. In developing the Spontaneous Stereotype Content Model (SSCM), Nicolas and colleagues (2022) asked participants to spontaneously describe social groups and used Natural Language Processing (NLP) techniques to demonstrate that spontaneous impressions of groups occur along a large and diverse array of dimensions. The two fundamental dimensions and their facets were prominent, with groups spontaneously described most frequently in terms of sociability (e.g., "nice") and morality (e.g., "greedy"), followed by assertiveness (e.g., "hard-working"), and ability (e.g., "smart"). However, groups were also frequently described in terms of their progressive/conservative beliefs (e.g., "religious"), status (e.g., "rich"), emotions (e.g., "sad"), appearance (e.g., "fat"), deviance (e.g., "different"), health (e.g., "unhealthy"), geography (e.g., "foreign"), and demographics (e.g., "old"). Spontaneous descriptions also provided insight into the *representativeness* of dimensions for groups. For example, although doctors and nurses were rated similarly on warmth and competence via rating scales, doctors were more likely than nurses to be spontaneously described as competent, and nurses were more likely than doctors to be spontaneously described as warm. Incorporating this additional information helped improve predictions of global evaluations of groups over and above rating scales alone.

Face perception research has also suggested that incorporating data-driven methods, as well as increasingly complex stimuli, supports higher-dimensional models. Noting the dominant two-dimensional model of trustworthiness and dominance was developed with relatively controlled, homogeneous stimuli, Sutherland and colleagues (2013) tested its generalizability with a more complex set of face images varying in age, expression, pose, facial hair, and more. Via this approach, perceptions were best accounted for by a three-dimensional model incorporating trustworthiness, dominance, and youthfulness/attractiveness. In other work, Lin and colleagues (2021) used stimulus sampling techniques and ratings on a broad set of trait adjectives to demonstrate evidence for a four-dimensional model, incorporating trustworthiness, dominance, youthfulness, and masculinity/femininity. Other work has explored the generalizability of the trustworthy/dominance model across world regions and suggested that when more data-driven methods are used, results increasingly diverge from the dominant trustworthiness/dominance model in many regions (Jones et al., 2021; Sutherland et al., 2018; Todorov & Oh, 2021).

In summary, recent research suggests that a full account of impression formation may require looking beyond the two fundamental dimensions and embracing higher-dimensional models. It also suggests that data-driven approaches, particularly leveraging perceivers' natural language about complex, naturalistic targets, may capture a higher number of impression formation dimensions than perceivers' ratings of targets on preselected scales. In the present work, we examined natural language descriptions of photos of individuals to explore the generalizability of high-dimensional impression formation (Nicolas et al., 2022) from groups to individuals and from group labels (i.e., concepts) to photos (i.e., visual information).

The Present Research

We studied unconstrained linguistic descriptions of Facebook profile pictures, which provide rich stimuli for impression formation research. In contrast to group labels (Koch et al., 2016; Nicolas et al., 2022), profile pictures display individuals, arguably the primary targets of impression formation. In contrast to previously used face stimuli (e.g., Lin et al., 2021; Oosterhof & Todorov, 2008), profile pictures provide a large amount of added complexity to perceivers, displaying a wide range of social cues beyond individuals' facial features capable of shaping the impressions people form of others. Such cues include affect, clothing, picture quality, body language and posture, geographical and behavioral contexts, and numerous other editing choices, such as cropping, color schemes, or the use of filters or overlays demonstrating support for social causes (Chapman & Coffé, 2016). Furthermore, Facebook profile pictures represent authentic naturalistic social stimuli that are frequently encountered (Davis, 2023), and which are (when public-facing) usable in academic research."

Despite these advantages, Facebook profile pictures have rarely been used in impression formation research. Instead, prior work has focused on identifying predictors of the kinds of pictures displayed (e.g., Chapman & Coffé, 2016; Hum et al., 2011) or assessing the accuracy with which information such as personality traits (e.g., Celli et al., 2014; Gosling et al., 2007), social class status (Becker et al., 2017), or educational attainment (Reiss & Tsvetkova, 2020) can be ascertained from profile pictures. By contrast, we focus solely on perceivers, and ask: What content dimensions underlie spontaneous impressions of profile pictures? As such, the sources (i.e., antecedent cues), accuracy, and functions (i.e., subsequent behaviors toward the subjects) of these spontaneous impressions are beyond the scope of the present research. We also focus on the U.S. context for comparability with the dimensions Nicolas and colleagues (2022) identified as underlying impressions of groups in U.S. society.

The progression of the article is as follows. First, we apply exploratory factor analysis (EFA) to ascertain the latent dimensions underlying descriptions of Facebook profile pictures. We then use cross-validation to assess the robustness of the emerging dimensions across different samples of targets and perceivers. Following this, we use natural language dictionaries to measure the relative prevalence of each dimension and inspect t-distributed stochastic nearest neighbor embedding plots to explore how each dimension informs the clustering of a visual map of intertarget similarities. Finally, we use semantic sentence embeddings to test each dimension's generalizability across impression formation contexts, comparing our EFA results with an alternate dimension reduction analysis based on descriptors' semantic similarities within large-scale text corpora. Taken together, our results generalized high-dimensional impression formation from labels of groups (Nicolas et al., 2022) to photos of individuals and provide further evidence of the need to consider, and the content dimensions involved in, high-dimensional impression formation.

Method

All study materials and data, code, and results are available on the Open Science Framework website (https://osf.io/ha3ge/). Our study had IRB approval.

Sampling Profile Pictures

We sampled 1,000 Facebook profile pictures of U.S. adults quasi-randomly by:

1. Randomly selecting a U.S. city name from a list of 62,439 cities.[2]
2. Entering the city name and home state into Facebook's search engine (e.g., "Gaines, Pennsylvania").
3. Selecting the first Facebook page in the search results with at least 300 "likes" from users.[3]
4. Selecting the user with the first publicly visible "like" on the selected page who (a) used a public-facing picture of themselves as their profile picture, (b) was the only person appearing in, or was unambiguously the focus of, the profile picture, (c) displayed publicly available information indicating residence in the United States (e.g., a U.S. location or workplace), and (d) used a profile picture displaying visible features associated with racial, age, and gender categories.[4]

Three independent raters subjectively coded the gender, age, and race/ethnicity of the profile picture subjects. Based on the lowest and highest estimates across three raters, there were between 622 and 625 women and 375 and 378 men categorizations ($\kappa_{Fleiss} = 0.96$), 269 to 356, 468 to 575, and 155 to 176 categorizations as people under 30 years (there were no children categorizations), 30 to 60 years, and above 60 years old ($\kappa_{Fleiss} = 0.62$), and 762 to 815 White, 74 to 88 Black, 51 to 130 Latino/a, 17 to 28 East Asian, and 16to 20 South Asian categorizations ($\kappa_{Fleiss} = 0.63$).

Participants

We recruited 2,188 U.S. adults from Prolific (age $M = 34.24$, $SD = 12.79$; 985 women, 1,057 men, 17 non-binary, 129 missing gender data; 214 Asian, 136 Black, 153 Hispanic, 1,482 White, 63 other race, 140 missing race data). Data were collected in two waves, with 1,148/1,040 responses collected in August/November 2020.[5] Due to the projects' exploratory/descriptive nature, we did not conduct power analyses.

Procedure

Obtaining Unconstrained Descriptions. Participants were shown 50 randomly sampled profile pictures one by one and instructed to "form an impression about the person in this Facebook profile picture." For each picture, participants generated three descriptors, using either single words or two words separated by a hyphen.[6] Profile pictures were described by an average of 109.35 participants each ($SD = 10.2$, range $= 74$-144).

Data Cleaning. We amended obvious spelling mistakes (e.g., replaced "energerti," with "energetic), and substituted responses inconsistent with instructions with obvious alternatives (e.g., "long hair" was replaced with "long-haired"). Responses without obvious alternatives were deleted (e.g., "bubry"), and responses containing multiple descriptors were split into their separate components (e.g., "old man, NFL fan" was split into "old," "man," and "NFL-fan"). Groups of close synonyms were manually aggregated into their most commonly used member (e.g., "goof," "goof-selfie, "goofball," "goofy," and "goofy-pose" were replaced with "goofy").[7] Following this, we were left with 311,370 applications of 6,993 unique descriptors. Notably, most descriptors were infrequent: 2,935 were used once, and 859 were used twice (see Supplementary Materials for examples of corrections, multi-element descriptors, synonyms, and common descriptors).

Results

Exploratory Factor Analysis

We assumed correlated dimensions, so decided against principal components analysis (PCA) in favor of exploratory factor analysis (EFA).[8] We treated profile pictures as cases, descriptors as variables, and counts the number of times each descriptor was applied to each profile picture as data points. Profile pictures' counts of descriptors varied considerably (range = 220-405), so we adjusted counts so each profile picture's descriptor count was equal.[9] We also log-transformed counts to better approximate the multivariate normal data assumed by EFA.[10] Analyses were performed in R using the psych package (Revelle, 2018).

With 1,000 targets it was not feasible to use all 6,993 descriptors in EFA. Past recommendations for rows-to-columns ratios in EFA have ranged from 20:1 (Hogarty et al., 2005) to 2:1 (Kline, 1979). This suggested that with 1,000 profile pictures, an upper limit of 500 descriptors might be included. To be as inclusive as possible, we used the 500 most common descriptors, which accounted for 85% of the data. Due to large discrepancies in frequency of use between different words, we relied on covariance matrices to preserve information about descriptors' relative frequency.[11] Sampling adequacy was confirmed via the Kaiser–Meyer–Olkin procedure, which estimates the proportion of variance in a data

Table 1. Top-Loading Descriptors for 14-Factor EFA Solution.

Factor number	1. Gender	2. Sociability[a]	3. Age[a]	4. Adventurousness	5. Ability[a]
Top 10 loadings	woman, 1.22	happy, 1.01	old, 1.39	adventurous, 0.84	smart, 0.65
	-man, -1.21	smiling, 0.75	-young, -1.05	outdoorsy, 0.71	professional, 0.61
	pretty, 0.7	friendly, 0.72	grandma, 0.41	traveler, 0.39	educated, 0.41
	mom, 0.61	-serious, -0.7	wise, 0.39	active, 0.37	successful, 0.32
	-handsome, -0.58	kind, 0.46	mature, 0.38	nature-lover, 0.37	rich, 0.28
	beautiful, 0.53	fun, 0.4	-cute, -0.37	fun, 0.25	business, 0.27
	-dad, -0.49	-unhappy, -0.39	retired, 0.31	natural, 0.24	-fun, -0.24
	-bearded, -0.32	-sad, -0.38	grandpa, 0.27	hiker, 0.21	ambitious, 0.2
	-funny, -0.28	nice, 0.37	-student, -0.25	relaxed, 0.16	-loud, -0.2
	grandma, 0.28	-mean, -0.37	dad, 0.24	explorer, 0.16	-insecure, -0.19
Cumulative Var.	0.06	0.12	0.17	0.2	0.22

Factor number	6. Non-Assertiveness	7. Morality[a]	8. Stylishness	9. Fitness[a]	10. Non-Conformity
Top 10 loadings	quiet, 0.41	caring, 0.68	stylish, 0.52	athletic, 0.62	weird, 0.39
	shy, 0.38	loving, 0.34	fashionable, 0.42	fit, 0.41	artistic, 0.31
	calm, 0.37	liberal, 0.28	confident, 0.36	strong, 0.35	creative, 0.29
	relaxed, 0.33	kind, 0.27	artistic, 0.29	active, 0.34	strange, 0.26
	sad, 0.3	compassionate, 0.26	trendy, 0.26	sporty, 0.33	-rich, -0.24
	tired, 0.27	activist, 0.25	cool, 0.25	healthy, 0.25	unique, 0.21
	introverted, 0.27	helpful, 0.24	vain, 0.23	hard-working, 0.19	odd, 0.21
	reserved, 0.26	-pretty, -0.21	hip, 0.23	determined, 0.19	quirky, 0.2
	lonely, 0.25	responsible, 0.2	rich, 0.21	energetic, 0.17	liberal, 0.2
	-proud, -0.24	supportive, 0.2	creative, 0.2	-overweight, -0.17	nerdy, 0.19
Cumulative Var.	0.24	0.27	0.29	0.31	0.32

Factor number	11. Conservatism	12. Attractiveness	13. Race	14. Weight
Top 10 loadings	conservative, 0.41	pretty, 0.27	white, 0.47	overweight, 0.3
	country, 0.37	-grandma, -0.22	-black, -0.41	loud, 0.25
	-fun, -0.25	-mom, -0.21	blonde, 0.26	-sweet, -0.25
	southern, 0.24	fake, 0.2	-strong, -0.22	-young, -0.21
	republican, 0.23	insecure, 0.2	-proud, -0.22	mom, 0.2
	-funny, -0.22	sexy, 0.19	vain, 0.18	-cute, -0.18
	hard-working, 0.2	vain, 0.19	fake, 0.17	-grandma, -0.18
	patriotic, 0.19	selfish, 0.19	selfish, 0.16	-artistic, -0.16
	-outgoing, -0.19	beautiful, 0.18	-hard-working, -0.15	-wise, -0.16
	proud, 0.18	filter, 0.18	insecure, 0.15	-retired, -0.15
Cumulative Var.	0.34	0.35	0.37	0.38

Note. Cumulative Var. = Cumulative variance explained.
[a]For ease of interpretation, loadings are reversed for factors whose strongest loading were predominantly negative (sociability, age, ability, morality, and fitness).

matrix that may be common variance. Our data yielded a KMO score of .75, above the standard minimum acceptable level of .50. To determine the optimal number of factors, we performed a parallel analysis (Horn, 1965), which suggested a 14-factor solution.[12] We ran an EFA extracting 14 factors using oblimin oblique rotation (Clarkson & Jennrich, 1988).

Results (Table 1) revealed 13 interpretable factors (one factor was not clearly interpretable, more on this below). Factors 1, 3, and 13 captured the key demographic variables of gender, age, and race (Black vs White), respectively. Factors 2 and 7 captured dual facets of the horizontal dimension, sociability ("happy," "smiling") and morality ("caring," "loving"), while Factors 5 and 6, respectively, captured dual facets of the vertical dimension, ability ("smart," "professional"), and assertiveness ("quiet," "shy"; we labeled Factor 6 non-assertiveness due to the strongest loading descriptors representing the inverse of assertiveness). Factors 8 (stylishness; "stylish," "fashionable"), 9 (fitness; "athletic," "fit"), and 12 (attractiveness; "pretty," "grandma") each captured distinct aspects of physical appearance. Factors 4 (adventurousness; "adventurous," "outdoorsy"), 10 (non-conformity; "weird," "artistic"), and

11 (conservatism; "conservative," "country") each captured individuals' proclivities for either exploring beyond conventional norms and boundaries (e.g., explorers or artists) or conforming to traditional social conventions (e.g., conservatives). This orientation toward either exploration and change or conformity and preservation has been argued to represent the key underlying principle in social perceptions of the dimension of conservative/progressive beliefs (Koch et al., 2016). By contrast, Factor 14 (labeled weight) was less conceptually coherent, displaying relatively weak loadings, and high-loading descriptors–"overweight," "loud," and "sweet"–suggesting a construct capturing both descriptive content concerning targets' body size and evaluative content concerning targets' character.

Collectively, the EFA solution accounted for 38% of the data's variance. Although this is low, we estimated that the top 500 descriptors had an average split half-reliability of 0.57. This suggests substantial unreliability in the columns as measures of the overall applicability of each descriptor to each profile picture,[13] and a relatively low upper limit on the factor solution's ability to account for variance in the data. Of the explained variance, approximately ⅓ (12.9% of total variance) was accounted for by the fundamental two dimensions' facets sociability, ability, non-assertiveness, and morality, and another ⅓ (12.4% of total variance) by the demographic factors of gender, age, and race. The final third was divided roughly equally between adventurousness, non-conformity, and conservatism (5.8% of total variance), and stylishness, fitness, and attractiveness (5.5% of total variance)

Factor Scores. To further explore the factor solution, we calculated factor scores using sum scores with a loading cut-off of 0.2. Figure 1 displays AI-generated recreations of the highest-scoring profile picture on each factor as well as factor score distributions. Table 2 reports correlations between factor scores, which were small to moderate with a few exceptions (e.g., age and attractiveness, consistent with Sutherland et al., 2013).

Cross-Validation. Next, we used cross-validation to assess the robustness of each factor. We divided data into its August and November waves, and randomly split profile pictures into two groups of 500: Groups A and B. We then compared the EFA factor structure in the August wave's responses to Group A to the factor structure in the November wave's responses to Group B. With 500 profile pictures in each dataset, EFAs relied on the 250 most frequent descriptors, which accounted for 75% of the data. Parallel analysis recommended a 10-factor solution, so we extracted 10 factors from each dataset using the procedures described above. Inspection of results (Table 3) revealed factors in each solution representing gender, sociability, age, adventurousness, stylishness, non-assertiveness, ability, morality, non-conformity, and race, all of which were present in our original 14-factor solution. Fitness was partially subsumed within non-conformity ("strong" was among the top-loading descriptors, negatively loaded), and attractiveness was incorporated within both age ("cute" and "pretty" were top-loading descriptors) and stylishness ("pretty" was a top-loading descriptor). Conservatism and weight were not clearly represented.

To test whether the respective solutions were compatible with identical underlying factor structures, we computed congruence coefficients between the ten factors in the August/Group A solution and the 10 factors in the Procrustes-rotated November/Group B solution (McCrae et al., 1996). We then used a bootstrapping procedure (Chan et al., 1999) to compute 95% confidence intervals for observed congruence coefficients under identical latent factor structures (see Supplementary Materials for details). As shown in Table 3, each observed congruence coefficient fell within the range expected under identical underlying factor structures. This suggests (at least for the first 9 factors, as congruence was relatively weak for the 10th factor, race) robustness across different targets and participants.

Discussion: EFA. Our EFA illuminated thirteen interpretable factors, including the four facets of the fundamental two dimensions: sociability, morality, ability, and assertiveness (named non-assertiveness in our results), three key demographic variables of gender, age, and race, three distinct aspects of physical appearance: fitness, stylishness, and attractiveness, and three distinct aspects of individuals' orientation toward exploring beyond or adhering to conventional physical and social boundaries: adventurousness, non-conformity, and conservatism.

These results are largely consistent with those of Nicolas and colleagues (2022). Despite being based on descriptions of group labels rather than individuals, and developed via different analytical tools, their SSCM encompasses 11/13 of these dimensions, including sociability, morality, assertiveness (the inverse of our non-assertiveness factor), ability, deviance (similar to our non-conformity factor), health (similar to our fitness factor), social groups (which encompasses our age, gender, and race factors), appearance, and beliefs (similar to our conservatism factor). Our EFA results diverge from the SSCM only with regard to the presence of stylishness and adventurousness as stand-alone dimensions in our results and to the extra dimensions of status ("wealthy," "high-status"), occupation (e.g., "lawyer"), and geography ("foreign," "Mexican") in the SSCM's taxonomy.[14] Thus, although our results suggest that the precise nature of the dimensions encompassed by spontaneous responses likely shifts between contexts, there nonetheless appears to be considerable consistency in the high-dimensional structure of impression formation with regard to descriptions of both individuals and groups.

Inspection of the factor scores produced noteworthy observations. Top-scoring profile pictures on the morality

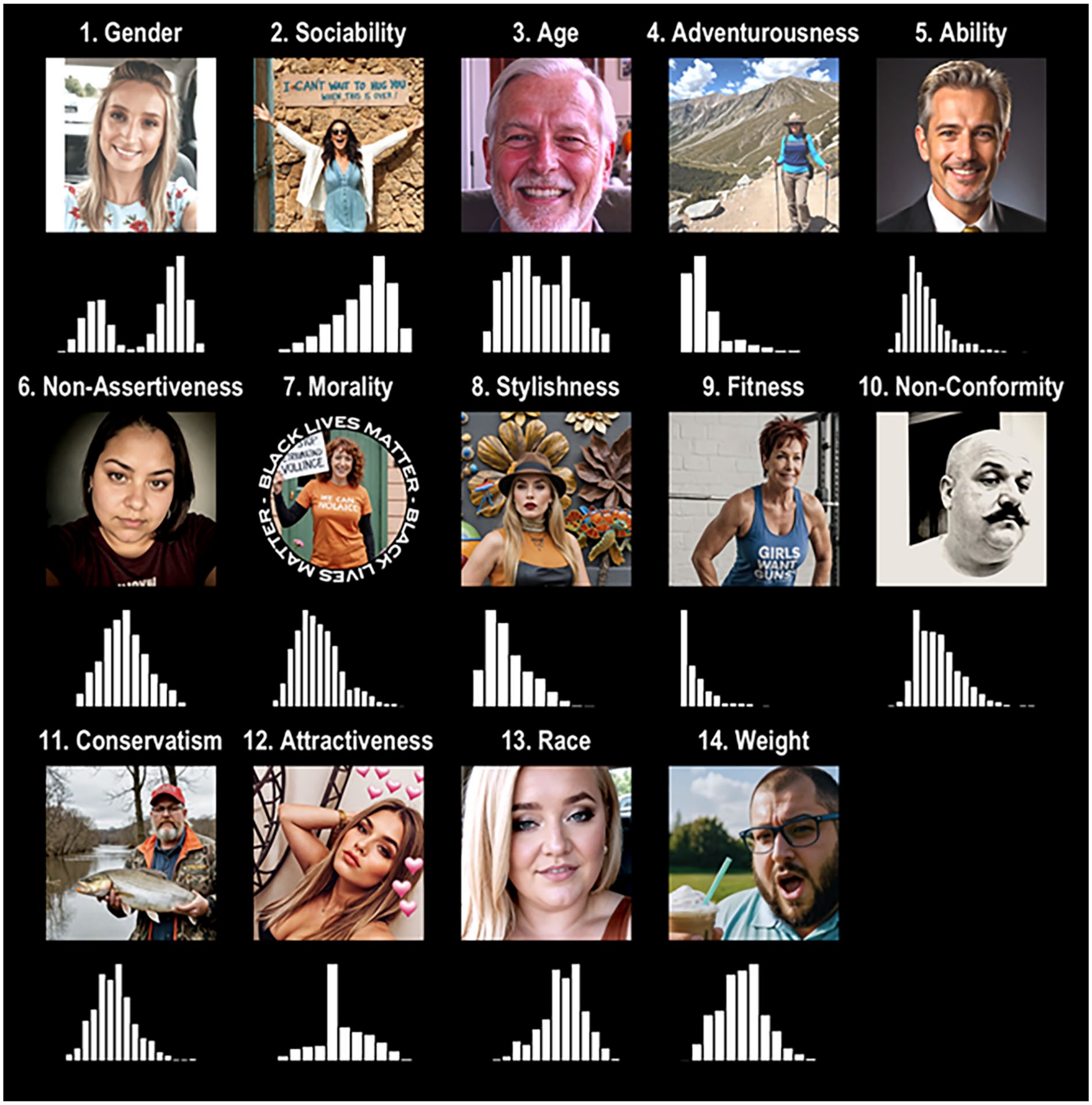

Figure 1. AI-Generated Recreations of the Highest-Scoring Profile Picture on Each Factor, and Factor Score Distributions.
Note. To protect individuals' privacy, we used AI text-to-image generation and photo editing software to recreate close likenesses to the top-scoring profile pictures on each factor.

factor tended to display graphics or slogans communicating beliefs (often support for the Black Lives Matter movement or COVID-19 responses), consistent with the notion that perceptions of communion/morality stem from perceived similarities in beliefs (Koch et al., 2020). By contrast, explicit signals of beliefs were not ubiquitous among profile pictures described as conservative, suggesting participants inferred conservatism from other kinds of cues (e.g., the presence of U.S. flags and activities such as hunting and fishing).

Also notable was the low correlation ($r = -.03$) between ability and non-assertiveness scores. This is consistent with prior work that found relatively weak correlations between competence and dominance in face perceptions (Sutherland et al., 2016), divergent relationships between facial competence and dominance and overall perceived valence (Oliveira et al., 2020), and that accepting (vs. rejecting) criticism increases perceived competence but decreases perceived dominance (Methner et al., 2020). This poses a puzzle for impression

Table 2. Correlations Between EFA Factor Scores.

Factor	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.
1. Gender													
2. Sociability	0.21												
3. Age	−0.37	−0.12											
4. Adventurousness	−0.13	0.30	0.00										
5. Ability	−0.07	−0.13	0.09	−0.29									
6. Non-Assertiveness	0.06	−0.18	0.12	−0.16	−0.03								
7. Morality	−0.02	0.27	0.38	−0.02	0.04	0.16							
8. Stylishness	0.08	−0.07	−0.43	−0.08	0.09	−0.18	−0.41						
9. Fitness	−0.26	−0.04	−0.02	0.47	−0.11	−0.29	−0.10	−0.04					
10. Non-Conformity	0.00	−0.18	−0.13	−0.11	−0.26	0.10	−0.01	0.18	−0.17				
11. Conservatism	−0.08	−0.47	0.24	−0.17	0.30	0.06	0.04	−0.19	0.14	−0.19			
12. Attractiveness	0.08	−0.01	−0.70	−0.02	0.00	−0.18	−0.53	0.47	0.01	0.08	−0.14		
13. Race	0.26	0.12	0.03	−0.02	−0.04	0.28	0.01	−0.07	−0.31	0.06	−0.08	−0.05	
14. Weight	−0.36	−0.25	0.62	−0.05	−0.01	−0.06	0.08	−0.27	0.01	−0.01	0.18	−0.38	−0.09

Table 3. Cross-Validation of 10-Factor Solutions From Reduced Datasets With Distinct Participants and Profile Pictures.

Factor	Gender	Sociability	Age	Adventurousness	Stylish	Non-assertiveness	Ability	Morality	Non-conformity	Race
August/Group A 10-factor solution										
Top loading descriptors	-man	happy	-old	-adventurous	confident	quiet	smart	-caring	artistic	-white
	woman	smiling	young	-outdoorsy	-weird	shy	professional	-loving	-strong	black
	pretty	friendly	-grandma	-active	stylish	calm	-fun	-kind	-hard-working	funny
	mom	-serious	cute	-nature-lover	fashionable	relaxed	educated	-sweet	creative	fun
	beautiful	fun	-mom	-traveler	pretty	tired	-funny	-warm	fashionable	cool
	-handsome	kind	pretty	-athletic	vain	sad	successful	-compassionate	cool	-selfish
November/Group B 10-factor solution										
Top loading descriptors	woman	happy	-old	-adventurous	confident	shy	smart	-caring	-strong	-pretty
	-man	friendly	young	-outdoorsy	stylish	sad	professional	-kind	stylish	-conservative
	pretty	-serious	-grandma	-active	vain	quiet	educated	-loving	-hard-working	black
	mom	smiling	-mature	-nature-lover	-weird	calm	-fun	-sweet	fun	-country
	-handsome	fun	cute	-traveler	rich	-proud	-loud	professional	-conservative	-cute
	-dad	nice	-wise	-athletic	pretty	relaxed	serious	white	artistic	-white
November/Group B Procrustes-rotated 10-factor solution										
Top loading descriptors	-woman	happy	-old	-adventurous	stylish	shy	-smart	-caring	weird	pretty
	man	friendly	young	-outdoorsy	confident	calm	-professional	-kind	funny	conservative
	-pretty	-serious	-grandma	-active	vain	sad	-educated	-loving	-hard-working	beautiful
	-mom	smiling	-mature	-nature-lover	rich	quiet	fun	white	-athletic	cute
	handsome	fun	-wise	-traveler	fashionable	-proud	loud	-sweet	-strong	country
	dad	kind	cute	-natural	fake	relaxed	-successful	professional	fun	-black
Congruence[a]	0.95	0.95	0.96	0.92	0.89	0.9	0.89	0.87	0.81	0.58
95% CIs	[0.93, 0.98]	[0.93, 0.96]	[0.94, 0.97]	[0.87, 0.95]	[0.76, 0.9]	[0.74, 0.88]	[0.74, 0.91]	[0.58, 0.89]	[0.65, 0.87]	[0.42, 0.85]

Note. —Denotes a negative loading descriptor.
[a]Congruence refers to the congruence between the initial August/Group A 10-factor solution and the November/Group B Procrustes-rotated 10-factor solution. [b] Confidence intervals were calculated from congruences between the initial August/Group A 10-factor solution and 1000 Procrusted-rotated 10-factor solutions computed from bootstrapped re-samples of the same data.

formation researchers, as it suggests that despite these factors theoretically representing the dual facets of the vertical dimension, the profile pictures perceived as being high or low in ability and assertiveness appear to display little overlap.

Finally, our cross-validation procedure suggested that the majority of these factors are robust across different samples of profile pictures and participants. Four factors were either subsumed within alternate factors (attractiveness into age and stylishness, fitness into non-conformity), exhibited low congruence across the subsets (race), or simply failed to emerge within the reduced factor solutions (conservatism). However, this may have occurred due to cross-validation EFAs using 250 descriptors rather than 500, and extracting 10 factors rather than 14.

Dimension Prevalence: Natural Language Dictionaries

To assess the relative prevalence of each dimension we used natural language dictionaries. The SADCAT R package

Table 4. Relative Prevalence and Most Common Words From SADCAT Dictionaries in Our Data.

Dictionary	Prevalence	Size	Most used	Most used (high)	Most used (low)	Prop. high
Sociability	76,304	1,266	nice, friendly	nice, friendly	shy, quiet	0.75
Appearance	50,897	1,910	pretty, white			
Emotions	49,988	1,209	happy, calm			
Social groups	36,159	139	old, young			
Morality	35,747	2,557	kind, good	kind, good	vain, selfish	0.67
Assertiveness	31,166	836	confident, calm	confident, strong	quiet, insecure	0.77
Ability	29,070	1,066	smart, good	smart, good	simple, silly	0.78
Age[a]	18,338	104	old, young	old, mature	young, immature	0.6
Beauty	17,726	272	pretty, beautiful	pretty, beautiful	loud, creepy	0.86
Occupation	14,101	2,052	professional, worker			
Gender[a]	12,608	381	woman, man	woman, girl	man, masculine	0.6
Other	11,081	1,561	mom, grandma			
Deviance	10,674	173	funny, weird	funny, weird	normal, average	0.77
Body properties	10,487	405	fat, athletic			
Adventurousness[a]	8,633	38	adventurous, outdoorsy	adventurous, outdoorsy	safe, cautious	0.95
Status	7,863	600	strong, rich	strong, rich	humble, poor	0.75
Body part	7,658	379	fat, stern			
Family	7,267	225	mom, grandma			
Relative	7,267	225	mom, grandma			
Politics	6,913	384	conservative, patriotic	conservative, traditional	liberal, open	0.46
Skin	6,432	77	white, blonde			
Health	6,335	1,529	fit, drinker	fit, healthy	drinker, crazy	0.22
Stylishness[a]	6,330	75	cool, stylish	cool, stylish	unfashionable, dated	0.99
Clothing	4,794	607	stylish, trendy			
Body covering	4,576	237	brunette, bald			
Geography	3,572	966	country, American			
Art	2,798	405	artistic, musician			
Inhabitant	1,915	670	American, hispanic			
Country	1,796	315	country, tourist			
Religion	870	809	Christian, spiritual	conservative, traditional	liberal, open	0.46
Lacks knowledge	618	8	unsure, human			
Beliefs (other)	568	128	materialistic, attitude			
Stem	563	780	gray, dynamic			
Humanities	290	83	romantic, classic			
Insults	204	43	blue, nasty			
Fortune	44	30	unfortunate, lucky			

[a]Age, Gender, Adventurousness, and Stylishness are custom dictionaries created with SADCAT's semi-automated dictionary creation functionality.

(Nicolas, 2022; Nicolas et al., 2021) provides dictionaries–lists of words centered on particular concepts–encompassing a wide range of stereotype content, including personality, affect, beliefs, social role, appearance, and more,[15] as well as 10 "high" and "low" sub-dictionaries representing opposite poles of select constructs' spectrums (e.g., high morality = "good"; low morality = "selfish"). SADCAT also allows users to create custom dictionaries using a semi-automated procedure that expands lists of seed words into larger dictionaries based on semantic relationships. We used this functionality to create additional dictionaries for the concepts of gender, age, adventurousness, and stylishness, plus high and low sub-dictionaries for each. For each custom dictionary, we used the top two loading terms of its associated EFA dimension as seed words (e.g., for gender "woman" and "man" served as seed words; see Supplementary Materials). Table 4 reports the number of descriptors from each dictionary in our data.

Discussion: Natural Language Dictionaries. As Table 4 shows, sociability was the most commonly used dictionary (primarily high sociability, e.g., "nice,"), followed by appearance (e.g., "pretty"), emotions (e.g., "happy"), social groups (e.g., "old"), morality (primarily high morality, e.g., "kind"), and assertiveness (primarily high assertiveness, e.g., "confident"), and ability (primarily high ability, e.g., "smart").

The relative predominance of the sociability and morality dictionaries is consistent with arguments regarding the

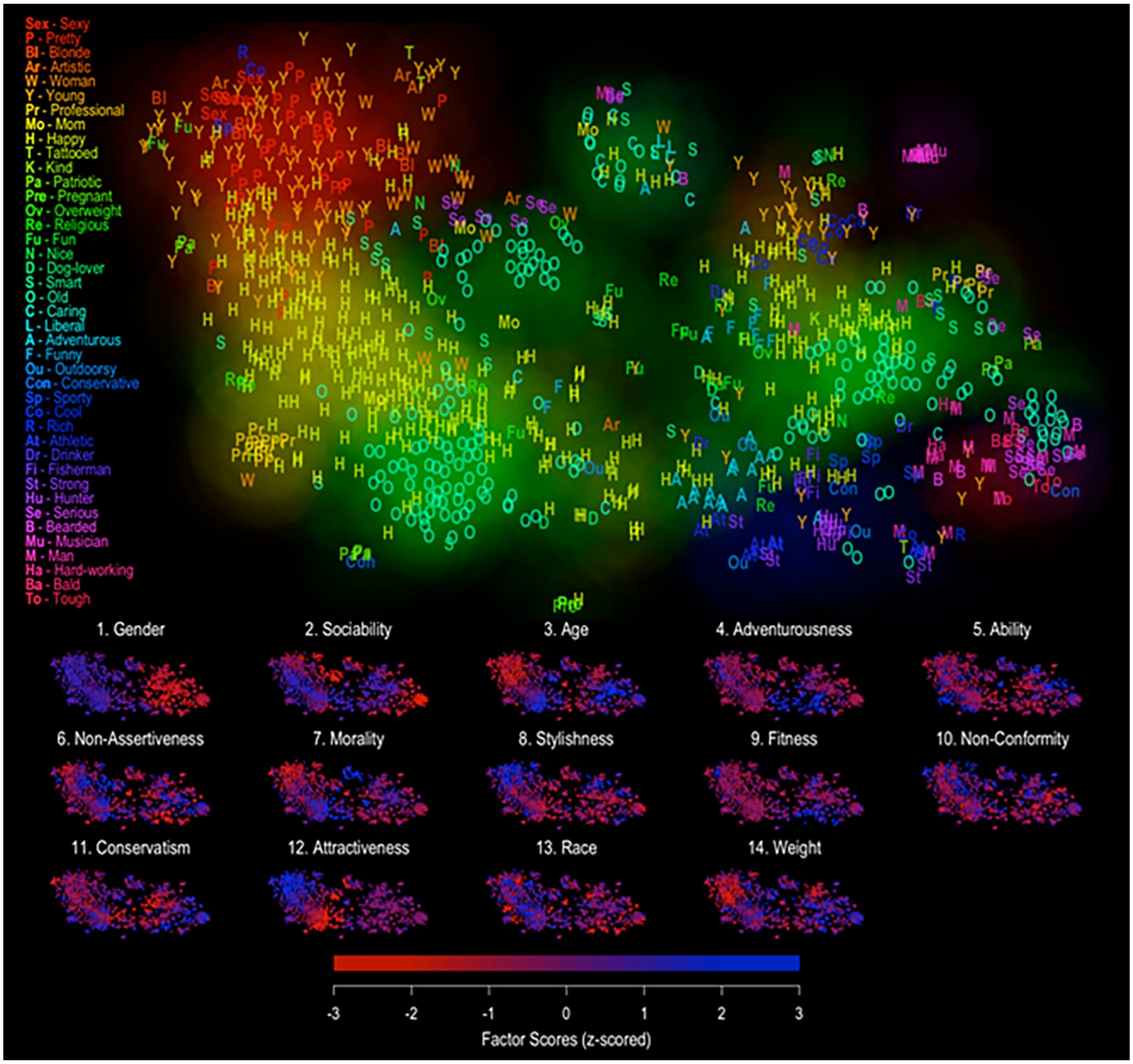

Figure 2. t-SNE Plot.
Note. t-SNE plot based on profile pictures' similarities with regard to scores on the 500 most frequent descriptors. Each profile picture is represented by its most frequent descriptor, abbreviated. Lower panels display the plot colored according to scores on each EFA dimension.

primacy of the horizontal dimension in impression formation (e.g., Abele & Bruckmüller, 2013; Brambilla et al., 2012; Wojciszke & Abele, 2008). It also aligns with Nicolas and colleagues' (2022) results regarding spontaneous descriptions of social groups. However, our results also differed in important ways from their findings, with appearance, emotions, and social groups dictionaries used relatively more frequently in descriptions of profile pictures compared with descriptions of groups, and status used less frequently. This suggests that just as the content of high-dimensional models of impression formation may shift from context to context, so too does the frequency with which particular dimensions are employed.

Dimension Diagnosticity: t-SNE

To explore the extent to which each dimension represented an important axis of inter-individual differentiation, we used t-distributed stochastic neighbor embedding (t-SNE), a data visualization technique for representing high-dimensional data in two or three non-linear dimensions (Van der Maaten & Hinton, 2008). Using t-SNE, we mapped the profile pictures in two dimensions according to their relative similarities with regard to the 500 most common descriptors. In the resulting solution (Figure 2), each profile picture is situated relatively closer to profile pictures producing relatively similar descriptions, and represented by a color-coded abbreviation of its most common descriptor (e.g., profile pictures described as "happy" more than any other descriptor are represented by the letter H). To assess the relationship between the t-SNE configuration and the EFA dimensions, the lower panels of Figure 2 color-code each point according to scores on each factor.

Discussion: t-SNE. Our t-SNE solution appeared sensitive to each EFA factor, albeit to varying degrees. As Figure 2 shows, the gender dimension was associated with the clearest clustering, with profile pictures perceived as women and men arranged in distinct groups on the left and right of the plot, respectively. By contrast, perceived race was associated with less clear clustering, with high- and low-scoring targets on the race dimension relatively dispersed across the plot. This suggests that targets perceived as women and men were described differently from each other overall compared with targets perceived as Black or White.

Unsurprisingly, the factors explaining most variance in our EFA–gender, sociability, and age–all served as influential organizing principles within the plot, with high- and low-scoring targets on each factor clearly separated. Most other factors were similarly associated with clear clustering, although some–non-conformity and race–were clustered less clearly, with high- and low-scoring targets relatively integrated throughout the solution. This suggests that although these dimensions may represent identifiable axes of intertarget variability, they are relatively less diagnostic of perceived global differences between target individuals.

Also notable was that a number of factors were associated with relatively more or less defined clustering, depending on the target gender. For example, virtually all the high- and low-scoring targets on the attractiveness dimension were described as women, a trend also seen for the age and morality dimensions. By contrast, conservatism appeared to produce clearer clustering occurring among men. This suggests differences in the relative representativeness (Nicolas, 2022) of these factors for different genders, and corroborates that the nature of the dimensions underlying impression formation may depend not just on the kind of stimuli perceived (i.e., groups vs. individuals) but also on the perceived social categories of targets (e.g., their gender, see Oh et al., 2020; Sutherland et al., 2015; their age, see Collova et al., 2019; Twele & Mondloch, 2022; or their culture, see Sutherland et al., 2018).

Dimension Generality: Semantic Sentence Embeddings

Finally, to assess the generalizability of each dimension to a broader source of data, we used semantic sentence embeddings. Semantic embeddings are based on the observation that semantically similar words (e.g., "happy" and "joyful") tend to be used in similar sentences (e.g., "It was a ___ moment when my child was born"), and so generally occur in close proximity with similar sets of other words (Z. S. Harris, 1954; Lenci, 2018). Via machine learning models trained on large-scale text corpora, semantic embeddings capture these idiosyncratic patterns of co-occurrence within numerical vectors that can be compared with provide a measure of the relative semantic similarity between different texts.

Nicolas and colleagues (2022) used semantic embeddings to investigate the dimensional structure of spontaneous stereotypes of social groups. We followed Nicolas and colleagues' procedures, but with one minor difference: rather than relying on embeddings for single words, we obtained embeddings for full sentences via the *all-mpnet-base-v2* pretrained SBERT model (Reimers & Gurevych, 2019). Theoretically, this enabled us to obtain semantic similarity data more tailored to the specific context of formation impressions of people, compared with word embeddings (see Supplemental Materials for more details). To calculate sentence embeddings, we placed each descriptor within a sentence structure explicitly describing a person (e.g., "happy" became "this person is happy"). To ensure sentences were grammatically correct we altered their structures slightly when necessary (e.g., "makeup" became "this person is wearing makeup").

We then applied K-Means cluster analysis to the matrix of cosine similarities between the sentence embeddings for the most common 500 descriptors and inspected a number of indices regarding the appropriate number of clusters provided by the NBclust R package (Charrad et al., 2014). We ultimately decided on a model with 42 clusters as one optimal solution (see Supplementary Materials for full clusters). Table 5 reports the most central 5 descriptors in each cluster (centrality = cosine similarity with cluster centroid), with clusters arranged according to the corresponding EFA factor.

Discussion: Semantic Sentence Embeddings. The clusters emerging via semantic sentence embeddings were largely consistent with our EFA results, but in many cases added nuance by separating distinct facets of dimensions. For example, seven clusters represented different aspects of sociability, such as positive and negative affect (clusters 1, 25, 31, & 33), and friendliness and unfriendliness (clusters 3, 6, & 19), and six clusters represented aspects of ability, including wittiness (cluster 7), obliviousness (cluster 16), industriousness (cluster 17), intelligence (clusters 26 and 37), and higher education (cluster 30). Notably,: status (cluster 29), an SSCM dimension that did not emerge in our previous results also appeared to be represented. A small number of clusters also represented relatively novel content, including relaxation (cluster 12), concentration (cluster 39), and some idiosyncratic aspects of appearance (see Table 5).

General Discussion

Recent work has sought to deepen our understanding of the underlying dimensions of impression formation by using

Table 5. Semantic Sentence Embedding Clusters Arranged According to Aligned EFA Factors.

Factor	Gender		Sociability							
Cluster	14.	20.	1.	3.	6.	15	19.	25.	31.	33.
Central descriptors	man masculine gentleman adult tall	mom married dad family-oriented divorced	happy pleased joyful smiling glad	amiable nice kind agreeable likable	outgoing social gregarious pleasant chatty	egotistical insecure pretentious arrogant narcissistic	rude mean obnoxious gross ugly	upset unhappy sad moody emotional	casual laid-back carefree spontaneous modest	lively vibrant energetic enthusiastic spirited
Factor	Age	Adventurousness	Ability						Morality	
Cluster	8.	18.	7.	16.	17.	26.	30.	37.	10.	40.
Central descriptors	old grandpa grandma middle-age gray	explorer adventurous traveler tourist hiker	witty smart clever funny silly	unaware unsure uncertain free-spirited different	hard- working professional working- class experienced worker	capable knowledgeable accomplished sophisticated diligent	student college interested curious inquisitive	stupid dumb ignorant mad naive	caring considerate sympathetic understanding compassionate	trustworthy loyal reliable dependable untrustworthy
Factor	Assertiveness				Stylishness		Fitness	Non-Conformity		Attractiveness
Cluster	11.	24.	27.	38.	2.	22.	9.	5.	21.	4.
Central descriptors	confident assertive prideful optimistic proud	shy quiet lonely loner timid	intimidating bossy intense dangerous stubborn	cautious worried scared concerned anxious	nice-picture photogenic model posing feminine	stylish fashionable well-dressed glamorous fancy	active athletic sporty fit rugged	strange weird odd amusing quirky	fine normal indifferent plain okay	lovely beautiful cute handsome good-looking
Factor	Conservatism			Race	Weight	Extras				
Cluster	35.	36.	41.	13.	32.	12	28.	29.	39.	
Central descriptors	christian religious gay lesbian guy	republican democrat conservative political liberal	military veteran patriotic basic	caucasian hispanic latino white american	fat obese heavy healthy unhealthy	relaxed comfortable tired peaceful warm	great awesome amazing good best	poor materialistic rich middle-class addict	busy attentive patient trying focused	

Note. Three omitted clusters referred to hair (Most central descriptors = “dark-haired,” “short-haired,” “curly”), glasses (“sunglasses,” “glasses,” “dark”) or the color red (“red” “redhead” “redneck”). Centrality = cosine similarity with cluster centroid.

increasingly data-driven methods (e.g., Jones et al., 2021; Koch et al., 2016; Lin et al., 2021; Nicolas et al., 2022) and incorporating more complex, naturalistic stimuli (Sutherland et al., 2013). We sought to build on these efforts by studying unconstrained linguistic descriptions of an authentic and highly complex naturalistic social stimulus set: Facebook profile pictures. Overall, responses confirmed the importance and centrality of the "fundamental two" horizontal and vertical dimensions and their respective facets of sociability/morality and ability/assertiveness (Abele et al., 2016, 2021). However, we also observed numerous further dimensions to play a role in participants' spontaneous responses, with descriptions also informed by the demographic variables of gender, age, and race, as well as multiple distinct aspects of physical appearance in profile pictures' attractiveness, stylishness, and fitness, and multiple distinct aspects of individuals' tendencies to explore beyond or conform to social conventions in profile pictures' perceived adventurousness, non-conformity, and conservatism.

Despite emerging via different stimuli and different methods, these results are largely consistent with the Spontaneous Stereotype Content Model (SSCM; Nicolas et al., 2022). The present work therefore conceptually replicates most of the SSCM's proposed dimensions and demonstrates their generalizability to impressions of individuals. However, as noted earlier, there were multiple discrepancies between our results and the SSCM. Two dimensions emerged in our data–adventurousness and stylishness–which were not differentiated into separate dimensions within the SSCM but were rather subsumed respectively as facets of assertiveness and appearance. And multiple SSCM dimensions–status, occupation, geography–failed to emerge consistently in our results. In addition, our natural language dictionary analysis suggested that the concepts of appearance, emotions, and social groups are relatively more frequently used in descriptions of profile pictures, whereas the concept of status is relatively more frequent in descriptions of groups.

Taken together, these results suggest that although there may be considerable stability in the high-dimensional structure of impression formation across contexts, the specific latent dimensions emphasized depend on the nature of the stimuli being evaluated. Given our field's long-standing reliance on simpler models and theory-driven methods, we believe that it will be important for further work to incorporate additional data and data sources, as well as further analytical approaches, to provide a clearer picture of the relative stability/lability of these high-dimensional structures across contexts, and to build a more complete understanding of where and when each dimension is and is not likely to be important for impression formation. It will also be important to examine the generality of these high-dimensional models to different cultural contexts, given that the present work and the SSCM have been largely based on predominantly White U.S.-based samples.

Another aspect of the SSCM supported by our results is the idea that groups/individuals differ in important ways in how likely they are to spontaneously elicit judgments related to those dimensions (i.e., dimensional representativeness). Our t-SNE plot suggested that the attractiveness, age, and morality dimensions were more frequently used to describe women, while the conservatism dimension was more frequently used to describe men. Future research may also explore the way the dimensional structure of impression formation depends not only on the kind of stimuli being evaluated (i.e., groups vs. individuals) but on the social categorization of targets.

It will also be important for future research to consider how impressions along each of these dimensions inform behaviors. In-keeping with traditional views of social perceivers as "cognitive-misers" (Taylor, 1981) and/or "motivated tacticians" (Fiske, 2004) who attend only to information relevant to the successful navigation of social environments, we assume that the dimensions emerging in our data provide important information to perceivers regarding targets' underlying character, motivations, and behavior. However, although Nicolas and colleagues (2022) demonstrated that perceivers rely on additional dimensions beyond warmth and competence to guide consequential decisions (e.g., health stereotypes incrementally predicted decisions to prioritize groups for Covid-19 vaccine eligibility), we still know little about the relative importance of each of these dimensions, or how and why impressions of individuals' adventurousness, stylishness, fitness, non-conformity, and conservatism guide behaviors across different individuals and contexts.

It should also be noted that by focusing on unconstrained descriptions, we are likely examining the end results of a complex interplay between distinct cognitive processes, with a number of the emerging content dimensions in our data appear resembling relatively discrete social categories (e.g., gender and race), and others resembling traits perceived as varying along a continuum (e.g., ability and sociability). However, while we expect that each of our observed dimensions results to some extent from both categorical and continuous differentiation processes (e.g., a target labeled "old" has likely been both categorized as such and placed somewhere along a continuous age continuum; Atwood et al., 2024), we consider it outside the scope of the present project and data to make any inferences regarding underlying mechanisms. Ultimately, in the present project, we are more concerned with the breadth and content of the resulting dimensions than the exact processes by which they came about.

A number of limitations should be noted. First, we modeled only impressions of U.S. residents by other U.S. residents. Further research needs to test whether high-dimensional impression formation also applies to other cultural contexts (especially those that are not Western, educated, industrialized, rich, and democratic; Muthukrishna et al., 2020), and how the emerging dimensions differ across cultural contexts.

Relatedly, Facebook profile pictures provide a rich but idiosyncratic variety of stimuli (White et al., 2017), and so could produce results lacking in generalizability to other impression formation contexts. In particular, it is likely that individuals' use of (and alertness toward others' use of) impression management strategies is heightened in the context of social media compared with other contexts (Roulin & Levashina, 2016). It is promising that these dimensions that our EFA differentiated also emerged in clusters determined by semantic similarities derived from large-scale text corpora and that many of the emerging dimensions have also been observed in prior work using group labels (Koch et al., 2016; Nicolas et al., 2022) or facial photographs (Lin et al., 2021; Sutherland et al., 2013). However, it remains possible that the more novel dimensions emerging in our data–adventurousness and stylishness–might be especially salient in the context of Facebook profile pictures, and may not be similarly emphasized in other contexts.

A final limitation is that there is arguably no perfect method for analyzing free-response text data. Here, we have relied heavily on EFA and found it to provide a useful and interpretable means of dimension reduction as well as results that generalize well across other methodological approaches. However, EFA also has a number of drawbacks with regard to free-response text data. First, it is highly data-hungry. Even with 1000 profile pictures, it was only feasible for us to include the most prevalent 500 descriptors in analyses, and by doing so we omitted 15% of our data. Moreover, even with over 300,000 total descriptions gathered, the average reliability of the descriptor columns remained low, constraining our factor solution's ability to account for the variance in the data.

Second, because EFA relies on covariation between descriptors to detect latent dimensions, it risks conflating conceptually distinct but correlated concepts. For example, the gender factor incorporated high loadings on descriptors of attractiveness (e.g., "pretty," "handsome"), the conservatism factor incorporated high loadings on descriptors of light-heartedness and humor ("fun," "funny"), and the sociability factor combined descriptors of personality ("friendly") and displayed affect ("happy"). Such conflations are less of a risk in cluster analyses based on semantic sentence embeddings, where similarities are based purely on the semantic use of texts. However, there may also be occasions when semantically distinct descriptors do in fact indicate a single latent dimension. For example, the descriptors "conservative," "country," and "southern" were all top-loading descriptors on the conservatism factor. Although each of these words refers to distinct concepts (political ideology, rurality, and geographical region, respectively), they nonetheless plausibly form a coherent impression formation dimension, by describing a perceived archetype of North American cultural and political conservatism. The same can be said for "smart," "educated," "professional," and "successful." These descriptors also mean different things, but (in our view) do not seem incoherent as top-loading descriptors of the ability factor. This suggests that covariation-based and semantic methods may have complementary strengths and weaknesses, and may both be benefcial for researchers in this space.

Since Asch's (1946) seminal work, social psychologists have produced a large body of knowledge focused on the fundamental two horizontal and vertical dimensions of human impression formation (Abele & Wojciszke, 2014; Judd et al., 2005), and for good reason. The present work confirms yet again that these fundamental two dimensions (and their respective facets of sociability/morality and ability/assertiveness) are central to the way people perceive and describe others. However, the present work also shows that there is much more to impression formation than just two dimensions and that our spontaneous impressions of others reflect more diverse and complex mental models than two dimensions allow. We believe that this additional content deserves greater attention, and perhaps especially so in the case of impression formation content related to progressive/conservative beliefs, given recent trends around affective political polarization (Iyengar et al., 2019). We therefore hope that the present research can serve to guide and inspire further work into how our complex, high-dimensional impressions of others give shape to our social lives.

Acknowledgments

We would like to thank Madalin Facino, Mriganka Singh, and Seungrim Yoo for their assistance with this project.

Author Contributions

P.C. and A.K. designed and implemented the study; P.C. conducted analyses and drafted the manuscript; and A.K., G.N., and S.A. provided editing and feedback on analyses and manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: P. C. was supported by the National Science Foundation (SPRF 2104594), S.A. was supported by grants from the National Institute on Aging (R01-AG018436, R01-AG067622, & R01-AG064006), and the Claude D. Pepper Older Americans Independence Center (OAIC) at Northwestern University (P30AG059988).

ORCID iDs

Paul Connor https://orcid.org/0000-0002-4995-4679
Gandalf Nicolas https://orcid.org/0000-0001-8215-1758
Alex Koch https://orcid.org/0000-0002-6267-8066

Data Availability Statement

All data, stimuli, and code used in the present project are available at https://osf.io/ha3ge/.

Supplemental Material

Supplemental material is available online with this article.

Notes

1. A recent adversarial collaboration on ongoing controversies about the Big Two dimensions suggested to labels them “horizontal” and “vertical,” to avoid making the impression that one pair of labels of the Big Two is more correct than another pair of labels (Abele et al., 2021).
2. The original list was obtained from https://github.com/grammakov/USA-cities-and-states.
3. Official city pages and pages with < 300 likes did not tend to generate sufficient engagement to provide usable profile pictures.
4. These stipulations were implemented to avoid profile pictures providing insufficient information for perceivers, such as photos taken from far away.
5. The second wave was added due to concerns that data for infrequent words in the first wave was highly noisy.
6. In Supplementary Materials we report Factor Analysis results based solely on the first descriptors provided by participants for each profile picture; these are highly similar to the results obtained from the full dataset.
7. Although this step departed from a purely data-driven process, it had little effect on results. In Supplementary Materials we report Factor Analyses based on data without synonym aggregation which are highly similar to those from synonym-collapsed data.
8. Given debate about whether PCA or EFA is more appropriate to summarize data (Jones et al., 2021; Todorov & Oh, 2021), we ran PCAs to test the robustness of our EFA results and examine the overlap between low-dimensional PCA results and standard models of impression formation. Overall, high-dimensional PCA solutions bear strong resemblance to EFA models, while low-dimensional solutions tended to conflate conceptually distinct dimensions (e.g., gender and sociability; see Supplementary Materials).
9. Specifically, each cell was multiplied by the row sum divided by 405 (the maximum row sum), making the sum of each row equal 405.
10. We added one to each count before taking the natural log to handle zeros. See Supplementary Materials for further information about the rationale behind this.
11. For count variables, descriptors used more frequently have more variance and, thus, more capacity for covariance. This heightened capacity for covariance results in higher item loadings in an EFA when the covariance matrix of descriptors is used, but not when the correlation matrix is used, because the correlation matrix standardizes variances. Thus, if we want to preserve information about descriptors’ relative frequency, we should use the covariance matrix instead of the correlation matrix.
12. Parallel analysis assesses whether successive factors explain more variance than would be expected to be explained by parallel factors extracted from random data.
13. Descriptor counts will tend to be more reliable when (a) descriptors are frequently used, (b) inter-rater agreement on whether descriptors apply to profile pictures is high, and (c) inter-rater agreement on the preferred term to describe an attribute is high.
14. Our participants used descriptors related to these dimensions, but they were either relatively infrequent (e.g., “foreign,” “lawyer,” and “high-class” were the 501st, 827th, and 1,340th most frequent descriptors, respectively), or subsumed by other dimensions.
15. We did not use the warmth, competence, or beliefs SADCAT dictionaries due to their high overlap with the dictionaries capturing sociability and morality, ability and assertiveness, and politics and beliefs (other), respectively.

References

Abele, A. E., & Bruckmüller, S. (2013). The Big Two of agency and communion in language and communication. In J. Forgas, O. Vincze, & J. Laszlo (Eds.), *Social cognition and communication* (pp. 173–184). Psychology Press.

Abele, A. E., Ellemers, N., Fiske, S. T., Koch, A., & Yzerbyt, V. (2021). Navigating the social world: Toward an integrated framework for evaluating self, individuals, and groups. *Psychological Review*, *128*(2), 290–314. https://doi.org/10.1037/rev0000262

Abele, A. E., Hauke, N., Peters, K., Louvet, E., Szymkow, A., & Duan, Y. (2016). Facets of the fundamental content dimensions: Agency with competence and assertiveness—Communion with warmth and morality. *Frontiers in Psychology*, *7*, Article 1810. https://doi.org/10.3389/fpsyg.2016.01810

Abele, A. E., & Wojciszke, B. (2007). Agency and communion from the perspective of self versus others. *Journal of Personality and Social Psychology*, *93*(5), 751.

Abele, A. E., & Wojciszke, B. (2014). Communal and agentic content in social cognition: A dual perspective model. In J. M. Olson & M. Zanna (Eds.), *Advances in experimental social psychology* (Vol. 50, pp. 195–255). Academic Press.

Asch, S. E. (1946). Forming impressions of personality. *The Journal of Abnormal and Social Psychology*, *41*(3), 258–290. https://doi.org/10.1037/h0055756

Atwood, S., Gibson, D. J., Briones Ramírez, S., & Olson, K. R. (2024). Flexibility in continuous judgments of gender/sex and race. Journal of *Experimental Psychology*: General. Advance online publication. https://doi.org/10.1037/xge0001512

Becker, J. C., Kraus, M. W., & Rheinschmidt-Same, M. (2017). Cultural expressions of social class and their implications for group-related beliefs and behaviors. *Journal of Social Issues*, *73*(1), 158–174. https://doi.org/10.1111/josi.12209

Brambilla, M., Sacchi, S., Rusconi, P., Cherubini, P., & Yzerbyt, V. Y. (2012). You want to give a good impression? Be honest! Moral traits dominate group impression formation. *British Journal of Social Psychology*, *51*(1), 149–166. https://doi.org/10.1111/j.2044-8309.2010.02011.x

Celli, F., Bruni, E., & Lepri, B. (2014, November). Automatic personality and interaction style recognition from Facebook profile pictures. In *Proceedings of the 22nd ACM International Conference on Multimedia* (pp. 1101–1104). Association for Computing Machinery. https://doi.org/10.1145/2647868.2654977

Chan, W., Ho, R. M., Leung, K., Chan, D. K.-S., & Yung, Y.-F. (1999). An alternative method for evaluating congruence coefficients with Procrustes rotation: A bootstrap procedure. *Psychological Methods*, *4*(4), 378–402. https://doi.org/10.1037/1082-989X.4.4.378

Chapman, H., & Coffé, H. (2016). Changing Facebook profile pictures as part of a campaign: Who does it and why? *Journal of Youth Studies*, *19*(4), 483–500. https://doi.org/10.1080/13676261.2015.1083962

Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2014). NbClust: An R package for determining the relevant number of clusters in a data set. *Journal of Statistical Software*, *61*(6), 1–36. https://doi.org/10.18637/jss.v061.i06

Clarkson, D. B., & Jennrich, R. I. (1988). Quartic rotation criteria and algorithms. *Psychometrika*, *53*(2), 251–259. https://doi.org/10.1007/BF02294136

Collova, J. R., Sutherland, C. A. M., & Rhodes, G. (2019). Testing the functional basis of first impressions: Dimensions for children's faces are not the same as for adults' faces. *Journal of Personality and Social Psychology*, *117*(5), 900–924. https://doi.org/10.1037/pspa0000167

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2007). The BIAS map: Behaviors from intergroup affect and stereotypes. *Journal of Personality and Social Psychology*, *92*(4), 631–648. https://doi.org/10.1037/0022-3514.92.4.631

Cuddy, A. J. C., Fiske, S. T., & Glick, P. (2008). Warmth and competence as universal dimensions of social perception: The stereotype content model and the BIAS map. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (*Vol. 40*, pp. 61–149). Elsevier Academic Press. https://doi.org/10.1016/S0065-2601(07)00002-0

Davis, K. (2023, August 22). *Facebook vs. Twitter. How do they stack up in 2023*. https://www.websiteplanet.com/blog/facebook-vs-twitter-stack/

Ellemers, N. (2017). *Morality and the regulation of social behavior: Groups as moral anchors*. Psychology Press.

Fiske, S. T. (2004). Intent and ordinary bias: Unintended thought and social motivation create casual prejudice. *Social Justice Research*, *17*, 117–127. https://doi.org/10.1023/B:SORE.0000027405.94966.23

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

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*, 878–902. https://doi.org/10.1037/0022-3514.82.6.878

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition*. Mcgraw-Hill Book Company.

Fiske, S. T., Xu, J., Cuddy, A. C., & Glick, P. (1999). (Dis) respecting versus (dis) liking: Status and interdependence predict ambivalent stereotypes of competence and warmth. *Journal of Social Issues*, *55*(3), 473–489. https://doi.org/10.1111/0022-4537.00128

Gosling, S. D., Gaddis, S., & Vazire, S. (2007). *Personality impressions based on Facebook profiles*. International Conference on Web and Social Media, 7, 1–4.

Harris, L. T., & Fiske, S. T. (2006). Dehumanizing the lowest of the low: Neuroimaging responses to extreme out-groups. *Psychological Science*, *17*(10), 847–853. https://doi.org/10.1111/j.1467-9280.2006.01793.x

Harris, Z. S. (1954). Distributional structure. *Word*, *10*(2–3), 146–162. https://doi.org/10.1080/00437956.1954.11659520

Hogarty, K. Y., Hines, C. V., Kromrey, J. D., Ferron, J. M., & Mumford, K. R. (2005). The quality of factor solutions in exploratory factor analysis: The influence of sample size, communality, and overdetermination. *Educational and Psychological Measurement*, *65*(2), 202–226. https://doi.org/10.1177/0013164404267287

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 32, 179–185. https://doi.org/10.1007/BF02289447

Hum, N. J., Chamberlin, P. E., Hambright, B. L., Portwood, A. C., Schat, A. C., & Bevan, J. L. (2011). A picture is worth a thousand words: A content analysis of Facebook profile photographs. *Computers in Human Behavior*, *27*(5), 1828–1833. https://doi.org/10.1016/j.chb.2011.04.003

Iyengar, S., Lelkes, Y., Levendusky, M., Malhotra, N., & Westwood, S. J. (2019). The origins and consequences of affective polarization in the United States. *Annual Review of Political Science*, *22*(1), 129–146. https://doi.org/10.1146/annurev-polisci-051117-073034

Jones, B. C., DeBruine, L. M., Flake, J. K., Liuzza, M. T., Antfolk, J., Arinze, N. C., . . . Sirota, M. (2021). To which world regions does the valence–dominance model of social perception apply? *Nature Human Behaviour*, *5*(1), 159–169. https://doi.org/10.1038/s41562-020-01007

Judd, C. M., James-Hawkins, L., Yzerbyt, V., & Kashima, Y. (2005). Fundamental dimensions of social judgment: Understanding the relations between judgments of competence and warmth. *Journal of Personality and Social Psychology*, *89*, 899–913. https://doi.org/10.1037/0022-3514.89.6.899

Kline, P. (1979). *Psychometrics and psychology*. Academic Press.

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*(5), 675–709. https://doi.org/10.1037/pspa0000046

Koch, A., Imhoff, R., Unkelbach, C., Nicolas, G., Fiske, S., Terache, J., Carrier, A., & Yzerbyt, V. (2020). Groups' warmth is a personal matter: Understanding consensus on stereotype dimensions reconciles adversarial models of social evaluation. *Journal of Experimental Social Psychology*, *89*, Article 103995. https://doi.org/10.1016/j.jesp.2020.103995

Lenci, A. (2018). Distributional models of word meaning. *Annual Review of Linguistics*, *4*, 151–171. https://doi.org/10.1146/annurev-linguistics-030514-125254

Lin, C., Keles, U., & Adolphs, R. (2021). Four dimensions characterize attributions from faces using a representative set of English trait words. *Nature Communications*, *12*(1), 1–15. https://doi.org/10.1038/s41467-021-25500-y

McCrae, R. R., Zonderman, A. B., Costa, P. T., Bond, M. H., Jr., & Paunonen, S. V. (1996). Evaluating replicability of factors in the Revised NEO Personality Inventory: Confirmatory factor analysis versus Procrustes rotation. *Journal of Personality and Social Psychology*, *70*(3), 552–566. https://doi.org/10.1037/0022-3514.70.3.552

Methner, N., Bruckmüller, S., & Steffens, M. C. (2020). Can accepting criticism be an effective impression management strategy for public figures? A comparison with denials and a counterattack.

Basic and Applied Social Psychology, *42*(4), 254–275. https://doi.org/10.1080/01973533.2020.1754824

Muthukrishna, M., Bell, A. V., Henrich, J., Curtin, C. M., Gedranovich, A., McInerney, J., & Thue, B. (2020). Beyond Western, Educated, Industrial, Rich, and Democratic (WEIRD) psychology: Measuring and mapping scales of cultural and psychological distance. *Psychological Science*, *31*(6), 678–701.

Nicolas, G. (2022). *SADCAT: Dictionary creation with stereotype content dictionaries*(R package version 0.1.0). https://github.com/gandalfnicolas/

Nicolas, G., Bai, X., & Fiske, S. T. (2021). Comprehensive stereotype content dictionaries using a semi-automated method. *European Journal of Social Psychology*, *51*(1), 178–196. https://doi.org/10.1002/ejsp.2724

Nicolas, G., Bai, X., & Fiske, S. T. (2022). A spontaneous stereotype content model: Taxonomy, properties, and prediction. *Journal of Personality and Social Psychology*, *123*(6), 1243–1263. https://doi.org/10.1037/pspa0000312

Oh, D., Dotsch, R., Porter, J., & Todorov, A. (2020). Gender biases in impressions from faces: Empirical studies and computational models. *Journal of Experimental Psychology: General*, *149*(2), 323–342. https://doi.org/10.1037/xge0000638

Oliveira, M., Garcia-Marques, T., Garcia-Marques, L., & Dotsch, R. (2020). Good to bad or bad to bad? What is the relationship between valence and the trait content of the Big Two? *European Journal of Social Psychology*, *50*(2), 463–483. https://doi.org/10.1002/ejsp.2618

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences*, *105*(32), 11087–11092. https://doi.org/10.1073/pnas.0805664105

Reimers, N., & Gurevych, I. (2019). Sentence-Bert: Sentence embeddings using Siamese BERT-networks. *arXiv preprint arXiv:1908.10084*. https://doi.org/10.48550/arXiv.1908.10084

Reiss, M. V., & Tsvetkova, M. (2020). Perceiving education from Facebook profile pictures. *New Media & Society*, *22*(3), 550–570. https://doi.org/10.1177/1461444819868678

Revelle, W. (2018). *psych: Procedures for Personality and Psychological Research, Northwestern University, Evanston, Illinois, USA*. https://cran.r-project.org/web/packages/psych/

Rosenberg, S., Nelson, C., & Vivekananthan, P. S. (1968). A multidimensional approach to the structure of personality impressions. *Journal of Personality and Social Psychology*, *9*(4), 283–294. https://doi.org/10.1037/h0026086

Roulin, N., & Levashina, J. (2016). Impression management and social media profiles. In R. Landers & G. Schmidt (Eds.), *Social media in employee selection and recruitment: Theory, practice, and current challenges* (pp. 223–248). Springer.

Sutherland, C. A., Liu, X., Zhang, L., Chu, Y., Oldmeadow, J. A., & Young, A. W. (2018). Facial first impressions across culture: Data-driven modeling of Chinese and British perceivers' unconstrained facial impressions. *Personality and Social Psychology Bulletin*, *44*(4), 521–537. https://doi.org/10.1177/0146167217744194

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*(1), 105–118. https://doi.org/10.1016/j.cognition.2012.12.001

Sutherland, C. A., Oldmeadow, J. A., & Young, A. W. (2016). Integrating social and facial models of person perception: Converging and diverging dimensions. *Cognition*, *157*, 257–267. https://doi.org/10.1016/j.cognition.2016.09.006

Sutherland, C. A., Young, A. W., Mootz, C. A., & Oldmeadow, J. A. (2015). Face gender and stereotypicality influence facial trait evaluation: Counter-stereotypical female faces are negatively evaluated. *British Journal of Psychology*, *106*(2), 186–208. https://doi.org/10.1111/bjop.12085

Taylor, S. E. (1981). A categorization approach to stereotyping. In D. L. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior* (pp. 88–114). Lawrence Erlbaum.

Todorov, A., & Oh, D. (2021). The structure and perceptual basis of social judgments from faces. In B. Gawronski (Ed.), *Advances in experimental social psychology* (pp. 189–245). Elsevier Academic Press. https://doi.org/10.1016/bs.aesp.2020.11.004

Twele, A. C., & Mondloch, C. J. (2022). The dimensions underlying first impressions of older adult faces are similar, but not identical, for young and older adult perceivers. *British Journal of Psychology*, *113*(4), 1009–1032. https://doi.org/10.1111/bjop.12568

Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of Machine Learning Research*, *9*(11), 2579–2605.

White, D., Sutherland, C. A., & Burton, A. L. (2017). Choosing face: The curse of self in profile image selection. *Cognitive Research: Principles and Implications*, *2*(1), 1–9. https://doi.org/10.1186/s41235-017-0058-3

Widaman, K. F. (2018). On common factor and principal component representations of data: Implications for theory and for confirmatory replications. Structural Equation Modeling: *A Multidisciplinary Journal*, *25*(6), 829–847.

Wiggins, J. S. (1979). A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of Personality and Social Psychology*, *37*(3), 395.

Wojciszke, B., & Abele, A. E. (2008). The primacy of communion over agency and its reversals in evaluations. *European Journal of Social Psychology*, *38*(7), 1139–1147. https://doi.org/10.1002/ejsp.549

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. https://doi.org/10.1177/0146167208318602

Yzerbyt, V. Y., Provost, V., & Corneille, O. (2005). Not competent but warm . . . really? Compensatory stereotypes in the French-speaking world. *Group Processes & Intergroup Relations*, *8*(3), 291–308. https://doi.org/10.1177/1368430205053944