Social Cognition, Vol. 43, No. 2, 2025, pp. 114–143
 https://doi.org/10.1521/soco.2025.43.2.114

Spontaneous Content of Impressions of Naturalistic Face Photographs

Gandalf Nicolas
Department of Psychology, Rutgers University

Stefan Uddenberg
Department of Psychology, University of Illinois

Alexander Todorov
Booth School of Business, University of Chicago

Across two studies (*N* = 4,526), we characterize a taxonomy of spontaneous face impressions by applying artificial intelligence text analyses to thousands of free-response descriptions of computer-generated faces. The taxonomy codes almost 100% of the impressions into Appearance (including Beauty), Sociability, Morality, Ability, Assertiveness, Emotion, Social Group, socioeconomic Status, Uniqueness, Family, Health, Occupation, Geographic origin, and political-religious Beliefs content. Results suggest that dimensions from low-dimensional models (e.g., Communion, Agency facets) are highly prevalent, but that alternative dimensions such as Uniqueness and Health are also prevalent. Most dimensions show high (positive) directions, and their correlational structure supports the clustering of low-dimensional models as separate from the expanded taxonomy dimensions. Finally, the taxonomy improves predictions of general evaluations of faces (how positive/negative the face is evaluated overall) and decision making in hypothetical scenarios (e.g., how much to prioritize a target for health care access or antidiscrimination protections).

Keywords: face perception, first impressions, computer vision, high-dimensional models, natural language processing

Faces are central to social relations (Todorov, 2017; Zebrowitz, 1990), serving as salient sources of impression formation (Sutherland & Young, 2022; Todorov, 2012). Psychologists have most often studied these face impressions by asking for judgment ratings along researcher-determined dimensions. However, these constrictive measures are unable to measure the nuances in linguistic content of impressions

Address correspondence to Gandalf Nicolas, Department of Psychology, Rutgers University – New Brunswick, 607 Tillet Hall, Piscataway, NJ 08854. E-mail: gandalf.nicolas@rutgers.edu

that come to mind upon seeing a face. In fact, recent efforts to content-code open-ended representations of social groups (Nicolas et al., 2022) and faces (Connor et al., 2024) have uncovered high-dimensional taxonomies that improve our understanding of social cognitions. Here, we use open-ended measures and artificial intelligence (AI) text analyses and stimuli to characterize a high-dimensional taxonomy of naturalistic face photographs, including its coverage of open-ended impressions, and the prevalence, direction, and predictive value of its dimensions.

IMPRESSION CONTENT

Dimensionality reduction of face ratings across multiple traits results in two relatively orthogonal evaluative dimensions: Trustworthiness/Communion and Dominance (related to Agency [Oosterhof & Todorov, 2008; Todorov & Oh, 2021]), even across 41 countries (B. C. Jones et al., 2021). Communion and Agency evaluations are also well established in other domains, such as stereotyping (Fiske et al., 2002, 2021). Face evaluations along these dimensions predict a myriad of socially relevant outcomes. For example, Communion impressions predict health care prioritization (Bagnis et al., 2020; Friehs et al., 2022), and Agency-related judgments predict election outcomes (Antonakis & Dalgas, 2009; Olivola & Todorov, 2010). Recent models, also relying primarily on scale ratings, have advanced additional dimensions of face impressions, such as youthfulness (Sutherland et al., 2013, 2016) and femininity (Lin et al., 2021).

Going beyond scales, studies using unconstrained open-ended measures along with content text analyses have shown further diversity of social perception dimensions. For example, the Spontaneous Stereotype Content Model (SSCM; Nicolas et al., 2022) proposed a comprehensive taxonomy of stereotypes associated with contemporary U.S. social categories (e.g., in terms of gender, race, age, occupation). Specifically, a large sample of Americans provided the "characteristics that came to mind" when thinking about each of the salient social groups, providing between one and six open-ended responses per target. The SSCM found 13 primary dimensions that Americans used to describe the characteristics they associated with social categories: Sociability and Morality (as facets of Communion), Ability and Assertiveness (as facets of Agency), socioeconomic Status, political-religious Beliefs (Koch et al., 2016, 2020), Appearance (including Beauty), Emotion, Deviance, Health, Occupation, Social Groups, and Geographic origin. Other content, such as family or culture-related (e.g., foods) stereotypes occurred less often. Not all these dimensions are equally important, however, with the SSCM showing that Communion and Agency facets were most prevalent in open-ended stereotypes, in line with lower-dimensional models. Nonetheless, the full taxonomy was needed to account for almost 90% of the open-ended responses. Furthermore, all the dimensions of the taxonomy showed value in predicting general prejudice and decision making across a variety of hypothetical scenarios. The SSCM taxonomy has also been found in stereotypes learned and reproduced by artificial intelligence (AI) language models (e.g., ChatGPT; OpenAI, n.d.), with implications for auditing and debiasing of these ever-more impactful technologies (Nicolas

& Caliskan, 2024b). Thus, in the closely related field of social category perceptions, high-dimensional and linguistic models have been shown to play relevant theoretical and practical roles as complements to lower-dimensional scale-based frameworks.

Most recently, the SSCM dimensions have been shown to also describe high-dimensional linguistic models of Facebook profile face pictures (Connor et al., 2024). In this study, participants saw Facebook photos and were asked to type three impressions they formed about the person in the picture. With some variability across methods, the impressions were well described by the SSCM taxonomy, but Status, Occupation, and Geography content was more infrequent than for stereotypes, among other smaller differences. Thus, high-dimensional models also seem to fit linguistic data of face impressions, at least based on social media stimuli.

A high-dimensional taxonomy trades parsimony for nuance. It is a good fit for open-ended responses, which show more variability than scale responses to researcher-selected items (Nicolas & Skinner, 2017; Nicolas et al., 2019). Additionally, it may cover impressions that go beyond the psychological traits that many (Oosterhof & Todorov, 2008), but not all (Vernon et al., 2014), previous low-dimensional models have focused on, such as descriptive appearance or social categories. Linguistic responses also provide unique insights, such as how prevalent different contents are in impressions. For example, perceivers may reliably respond to a scale asking how healthy versus unhealthy (i.e., direction) a face looks, but evaluations about Health may not be prevalent when asking for spontaneous impressions. Dimensional prevalence (also called representativeness) may relate to accessibility in memory (Higgins, 1996) and associations with physical characteristics. For these reasons, we focus on expanding our understanding of high-dimensional representations of face impression content, building on previous analyses of open-ended responses (Collova et al., 2019; Connor et al., 2024; Oosterhof & Todorov, 2008; Sutherland et al., 2018).

GENERALIZING A HIGH-DIMENSIONAL TAXONOMY TO AI-GENERATED FACES

Our first goal is to identify a high-dimensional taxonomy using a unique set of face stimuli: AI-generated face photographs. Most face impression research uses either photographs or 3-D digital faces. Photographs are realistic and ecologically valid, but their use is limited by small datasets and privacy concerns. Digital faces, on the other hand, can be generated in large numbers (Oosterhof & Todorov, 2008). However, the digital faces used so far have been unrealistic and very homogeneous.

Here, we instead use recently developed AI methods to create realistic and diverse digital faces (see Figure 1). This approach draws on models trained on large databases of naturalistic face photographs (Karras et al., 2020, 2021). The public is increasingly exposed to AI-generated faces using commercial models (e.g., ChatGPT; OpenAI, n.d.), and psychologists increasingly rely on AI for faster and less resource-intensive stimuli generation, underscoring the relevance of using such stimuli in our understanding of ecological face impressions. Our results can be

FIGURE 1. Randomly selected sample of 25 artificial faces used in Study 1.

used by other researchers to generate novel hyperrealistic face models of multiple dimensions (Todorov & Oh, 2021), further improving stimuli diversity. In addition, while most prior studies have used relatively few faces (typically, fewer than 100), this approach can generate a much larger stimulus set (Study 1, N = 3,022).

Most relevantly, we build upon the SSCM (Nicolas et al., 2022) and the recent high-dimensional model derived from Facebook profile pictures (Connor et al., 2024). In order to increase the robustness and generalizability of existing models, such as the SSCM, it is important to test different face datasets, given the potential role of stimulus set–specific factors affecting the results. For example, while Facebook (and likely other social media) profile pictures are stimuli that contemporary perceivers encounter often, they also may contain non–face specific information, such as filters and prominent backgrounds. Here, we wanted to test generalizability to a set of naturalistic faces that do not contain filters and with reduced background information, providing a closer look at content based primarily on facial information alone.

DESCRIBING THE PROPERTIES OF A HIGH-DIMENSIONAL TAXONOMY OF IMPRESSIONS

Based on previous high-dimensional models (e.g., Nicolas et al., 2022), we aim to characterize several properties of taxonomies derived from linguistic analyses of open-ended responses. The first property of interest is coverage, referring to how many of the participants' face impressions the taxonomy can account for. That is, we measure how many of the participants' impressions fall into the dimensions we identified. To be informative, a taxonomy aiming to describe content must be able to explain a large majority of impressions. Previous high-dimensional taxonomies have achieved >80% coverage of open-ended responses (Nicolas & Caliskan, 2024a; Nicolas et al., 2022).

The second property is content prevalence, referring to how often each dimension spontaneously arises in face impressions. Prevalence is a measure of primacy (Abele et al., 2021), with more important dimensions spontaneously coming to mind more often. Thus, prevalence provides information about which impressions may be more predictive of various outcomes (e.g., Nicolas & Caliskan, 2024a), as well as a means for researchers to choose smaller sets of impactful dimensions when a high-dimensional taxonomy is not desirable.

The third property is directionality, that is, how high or low a dimension is evaluated in a semantic differential. For example, Morality impressions may be more often about low (i.e., immorality) than high morality. Direction is the variable traditionally measured by psychological scales, and their role in predicting a variety of outcomes is well established, as reviewed above.

Fourth, we examine the correlational structure of impressions, providing connections to low-dimensional models. It is possible that dimensionality reduction of the higher-dimensional models aligns with existing low-dimensional models or provides an intermediate solution. We explore this possibility, providing a spectrum of dimensions that may describe face impressions.

Finally, we aim to establish the predictive value of the taxonomy. We test whether the proposed dimensions improve predictions of general valence and decision making above established low-dimensional content. Specifically, the high-dimensional taxonomy not only expands the number of dimensions explained, but also includes the prevalence and direction properties, which may interact to improve predictability. For example, in studies of open-ended stereotypes in both humans and AI, direction along a specific dimension was more predictive of prejudice towards a group when prevalence of the dimension was also high in descriptions of the group (Nicolas & Caliskan, 2024b; Nicolas et al., 2022). Additionally, all main dimensions of the SSCM (whether using their direction, prevalence, or both) predicted decision making across a series of hypothetical scenarios, including what social groups to prioritize for vaccination programs, emotional counseling access, and antidiscrimination programs. These findings held when controlling for Communion and Agency. Thus, many underexplored impression dimensions (e.g., Deviance, Health) may help improve predictions of decisions in relevant real-world cases (Nicolas et al., 2022). Here, we test this possibility for a high-dimensional taxonomy of face impressions.

CURRENT STUDIES

We present two studies using state-of-the-art AI face stimulus generation and text analysis, characterizing a rich content taxonomy of spontaneous face impressions (see Supplement, available on the OSF site mentioned below, for a third study replicating findings with additional variations). We examine properties of the taxonomy, including coverage, prevalence, direction, correlational structure, and predictive value.

Our research partially builds on the SSCM (Nicolas et al., 2022) and the high-dimensional impressions taxonomy advanced by Connor and colleagues (2024). However, we examine whether the taxonomy generalizes to AI-generated face stimuli that contain no filters or other prominent nonfacial information. This may result, for example, in fewer impressions about political-religious Beliefs, which are often conveyed in social media data by filters. In addition, in Study 2, we attempt to reduce the influence of socially desirable information, such as smiles, to examine face impressions of more emotionally neutral stimuli. We include other robustness tests for this recent taxonomy, including examining the role of reliability and number of impressions provided in the task. Furthermore, while Connor and colleagues focused on identifying the dimensions of a high-dimensional taxonomy, here we elaborate on the dimensions' coverage, prevalence, direction, and correlational properties, using dictionary analyses. In addition, we provide novel examinations of the predictivity of this high-dimensional taxonomy of face impressions.

This report is partly data-driven, which provides the necessary structure for making theoretical progress. However, we have several hypotheses, in part based on expected convergence with previous high-dimensional social impressions models (Connor et al., 2024; Nicolas et al., 2022). First, we expect that our model will achieve appropriate coverage (over 80% of open-ended responses accounted for by identified dimensions).

Second, we expect that prominent models' dimensions will be most prevalent, given their established role in predicting behavior (Deska et al., 2020; Jaeger et al., 2020; Olivola & Todorov, 2010; Todorov et al., 2015). However, we hypothesize that many other dimensions found in parallel domains (e.g., stereotyping; Nicolas et al., 2022) will also be significantly prevalent, including socioeconomic Status, Uniqueness, Health, Social Groups, Emotion, and Occupation. Many of these dimensions have been previously studied in face research (A. L. Jones, 2018; Oh et al., 2020) but had not been incorporated into general models of impressions until the recent Connor and colleagues (2024) high-dimensional model.

Third, we expect that the average impression direction will be high rather than low. For most dimensions this means positive (vs. negative; Nicolas et al., 2021). Both the Pollyanna principle (Matlin & Stang, 1978) and positive self-presentation tendencies in naturalistic photographs (Wu et al., 2015) support this hypothesis. However, it contrasts with negative evaluations in other domains (e.g., stereotypes; Nicolas et al., 2022). We do expect relatively higher Ability and Sociability, and lower Morality and Health evaluations, given positive self-presentation and previous research on negativity biases for Morality and Health (Nicolas et al., 2022).

Fourth, we expect the correlational structure of the dimensions' prevalence to reflect low-dimensional theoretical models. Specifically, we expect Communion facets of Morality and Sociability to cluster together, and Agency-related dimensions (Ability, Assertiveness, Status) to cluster together (Abele et al., 2021). We also expect to find differentiation based on more internal versus more external traits.

Finally, we hypothesize that the high-dimensional taxonomy will improve predictions of general evaluations of faces and decision making in a variety of hypothetical scenarios, as compared to established lower-dimensional models (more specific preregistered hypotheses are discussed for Study 2).

Study 1 was approved by the Princeton University Institutional Review Board (IRB), and Study 2 was approved by the Rutgers University IRB. All studies, measures, manipulations, and data/participant exclusions are reported in the manuscript or the Supplement. Data and code are available at: https://osf.io/pmtxw/?view_only=a485688056a8495ba37082f38c690833

STUDY 1

In Study 1, we use dictionary analyses to characterize a taxonomy of face impression content, including the dimensions' coverage, prevalence, direction, and correlational structure.

Method

Study 1 was not preregistered.

Participants. Participants were 3,145 Amazon Mechanical Turk workers (after removing 13 participants for not following instructions, such as providing nonsense responses; exclusions did not affect conclusions). Participants' mean age was 38.18 years; most identified as women (53.8%; 45.2% men) and White (71.6%; 8.5% Asian; 8.5% Black; 5.2% multiracial; 4.2% Hispanic).

Power analyses for all studies were calculated using G*Power (Faul et al., 2009) and indicated > 99% power to detect a small (d = .2) effect in a paired samples t test, an approximation of the planned prevalence and direction pairwise comparisons. We note that computing power for the crossed-random effect models is complex. However, additional sensitivity analyses for power analysis using the *simr* R package showed that for small effect sizes and various variance specifications for mixed models ANOVA and pairwise comparison results consistently showed >80% power for our sample size.

Materials and Procedure. To generate stimuli faces we used StyleGAN2, a state-of-the-art generative adversarial network (Karras et al., 2020, 2021). This AI model was trained on the Flickr-Faces-HQ dataset, which includes 70,000 high-quality images (see Supplement for additional information). The generated faces do not depict real persons, yet they are photorealistic and retain the resolution and diversity of the training set (Peterson et al., 2022). Studies show that humans perform only slightly better than chance in discriminating AI-generated faces from real faces (although AI-generated faces may be seen as more trustworthy; Becker & Laycock, 2023; Boyd et al., 2023; Nightingale & Farid, 2022; Shen et al., 2021). Our use of these models is restricted to artificial stimuli generation and evaluation for purposes of research into often biased yet consequential face impressions.[1]

Each participant evaluated one of 3,022 distinct faces (some faces were seen, at random, more than once due to our sample size being slightly larger than the stimulus set). We presented one face per participant to avoid content elicitation based on comparison with other faces in the task. After providing consent, participants read: "In this study you will see a face, and you will be asked to provide the first 6 characteristics that you spontaneously think about when you see this person. These answers will be completely anonymous. We are interested in your immediate, gut reaction to the images. There are no right or wrong responses. Please use single words (and no more than two words, for example, an adjective + noun) for each of your answers." Then, participants saw the face stimulus and the instruction: "What are the first 6 characteristics that you spontaneously think about when you see this person?" along with six blank forms for typing responses. Finally, participants completed demographic questions. Table 1 provides demographic context for the stimuli, suggesting significant (perceived) demographic diversity.

Analysis Strategy. First, we spell-checked and preprocessed the text responses to remove capitalization, symbols, and inflections. Then, we used a dictionary approach to reduce the dimensionality of over 3,600 distinct responses and obtain an interpretable taxonomy.

Dictionary coding: Dictionaries were used to evaluate coverage, prevalence, and direction. These dictionaries are validated lists of words semantically associated with different dimensions of person evaluation (Nicolas et al., 2021). The 15 dictionary dimensions are: Sociability, Morality, Ability, Assertiveness, Emotions, Social Groups, Status, Beliefs, Uniqueness, Family, Health, Occupation, Geography, Appearance, and a dictionary grouping Other smaller content (plus we subset a Beauty dimension from the Appearance dictionary, given its relevance; see Table 2 for example words in each dictionary). To further validate the dictionaries, we show that they correlate with scale ratings of face impressions as expected (see Supplement).

Prevalence and coverage: Prevalence coding involved matching responses to the dictionaries (available at https://github.com/gandalfnicolas/SADCAT), such that the response was coded as either 0 or 1 for each dimension, depending on whether the response was about the dimension (i.e., it was in the dimension's

1. Other use cases should be evaluated based on evolving responsible practices in this emerging field.

TABLE 1. **Demographic Categorization Ratings of Stimuli**

	Study 1		Study 2	
	M	*SD*	*M*	*SD*
Male	46.25	24.8	43.84	24.68
Old	34.01	12.06	30.69	11.75
Gay	33.93	6.15	32.66	6.65
Asian	24.44	20.65	26.28	21.41
Black	11.69	12.64	11.95	12.27
Hispanic	34.84	13.55	35.06	13.82
Middle Eastern	32.61	15.03	32.65	15.04
Native	26.73	13.19	27.09	13.29
White	65.43	23.64	64.14	23.79

Note. Ratings ranged from 0 to 100 (for gender ranging from "feminine" to "masculine"; for the various racial categories ranging from, e.g., "Not Asian" to "Asian"; for age ranging from "young" to "old"). Scores derived from models, trained on the same set of artificial faces, predicting various demographic categorizations based on ratings from over 4,000 participants. Note that these scores reflect solely average categorizations and, as these are our participants' impressions, provide no information about "true" identity or attributes. Retrieved from Peterson et al.'s (2022) data on scale judgments of the same set of stimuli.

dictionary). For example, the words "friendly" and "unfriendly" are in the Sociability dictionary (Nicolas et al., 2021), so a response using either of these words would receive a score of 1 on Sociability. A response could be coded into multiple dictionaries if related to more than one dimension. Responses not included in dictionaries were coded as "No match," and this variable was used to quantify coverage. For prevalence analyses, we averaged each participant's responses for each dimension (resulting in a percentage of a participant's responses about the dimension per face). The predictor variable was a categorical indicator for each dimension, and the outcome was the response rate for the corresponding dimension.

Direction: Responses were assigned a value of –1 if they were low on the dimension, 0 if neutral, and +1 if high. For example, "weird" scores high on the Uniqueness dictionary, while "typical" scores low; "friendly" scores high on Sociability, "unfriendly" scores low. For the Beliefs dimension, the endpoints are arbitrary, with high direction indicating conservatism/religiousness (e.g., "traditional"), while low direction indicates liberalism/secularism (e.g., "progressive"). Given lack of clarity about appropriate endpoints, direction indicators are not available for some dimensions (e.g., Occupations). For analyses, direction across each participant's responses for a face was averaged.

Correlational structure: Next, we examined correlations in dimensions' prevalence using text embeddings. Text embeddings are numerical-vector representations of text, encoding information about semantic relations between words. For example, words such as "quick" and "fast" have more similar embeddings than "quick" and "hair." The embeddings are retrieved from an AI model that learns semantic relationships by training on large text corpora (e.g., the common crawl, a vast sample of World Wide Web content; see https://commoncrawl.org/). Here, we use the SBERT embeddings model (Reimers & Gurevych, 2019). SBERT is a recent model that accommodates context and multiword responses and returns 768-dimensional vectors for each response.

To obtain these correlations we first obtained highly prototypical dictionary words for each dimension (Nicolas et al., 2021). For example, the words "friendly" and "sociable" are highly prototypical of the Sociability dimension (see online repository for list). Thus, we obtained the text embeddings of multiple prototypical words per dimension and averaged them, resulting in an embedding for each dimension. After this, we correlated each response's embedding to each dimension's embedding, resulting in a score of how semantically related responses are to the different dimensions. Finally, we computed a correlation matrix and hierarchical clustering between these variables.

Software. All analyses were run using the R 4.3 (R Core Team, 2024) *lme4* (Bates et al., 2015) and *lmerTest* (Kuznetsova et al., 2017) packages. We use estimated marginal means and two-tailed comparisons with Tukey corrections using the package *emmeans* (Lenth, 2016). We used maximally converging linear mixed models (participants and stimulus as random factors).

Results and Discussion

The taxonomy dimensions and their top nonoverlapping words (in our data), as derived from the dictionaries, are presented in Table 2.

Coverage. As hypothesized, the dimensions measured by the dictionaries had very high coverage: The 15 dictionaries accounted for over 96% of the total responses. The remaining unaccounted-for responses tended to be largely idiosyncratic or involved grammatical mistakes or nonsense responses. For comparison, looking at only the two (Communion & Agency; Oosterhof & Todorov, 2008) or three (adding Beauty; Sutherland et al., 2018) most established content dimensions, responses accounted for were only 47.5–54.6% of the total. This suggests that the taxonomy can be used to explain the vast majority of face impressions of AI-generated photos.

Prevalence. We proceed to present the prevalence of the taxonomy dimensions. Prevalence provides information about the primacy of content, as more prevalent dimensions will be encountered more often in impressions and may thus have a larger effect on downstream consequences (Abele et al., 2021). In addition, studies needing to focus on a smaller number of measures may use more prevalent dimensions to achieve maximal coverage.

Tests comparing the various dimensions' response proportions (prevalence) were significant, $F(16, 53448) = 1{,}077.7$, $p < .001$, $\eta^2 = 0.244$, 95% CI = [.238, .250]; see Table 3 and Figure 2. We found that Appearance and Sociability contents were

TABLE 2. Nonoverlapping Top Words and Frequencies for the Dimensions, Study 1

Appearance	**freq.**	**Sociability**	**freq.**	**Emotion**	**freq.**
pretty	263	friendly	577	happy	939
cute	226	nice	399	content	90
smile	181	sweet	123	sad	70
handsome	102	fun	120	loving	51
Ability	**freq.**	**Morality**	**freq.**	**Social Groups**	**freq.**
smart	491	honest	125	young	452
intelligent	190	sincere	43	old	148
bright	56	trustworthy	40	male	127
educated	54	genuine	37	female	118
Assertiveness	**freq.**	**Status**	**freq.**	**Uniqueness**	**freq.**
confident	95	rich	38	interesting	37
determined	40	successful	31	average	36
energetic	39	wealthy	28	normal	27
active	28	poor	19	unique	9
Health	**freq.**	**Geography**	**freq.**	**Beliefs**	**freq.**
healthy	51	foreign	16	conservative	18
drunk	5	outsider	7	liberal	13
lame	5	suburban	7	religious	9
stressed	5	Irish	4	questioning	8
Occupation	**freq.**	**Family**	**freq.**	**Other**	**freq.**
professional	67	child	52	artistic	8
hard working	26	mother	23	artsy	3
business	12	motherly	21	scientific	3
teacher	12	baby	17	romantic	2

the most common. Emotion, Ability, Morality, Social Group (e.g., gender, race), and Assertiveness (including dominance-related content) impressions followed, and then Beauty (i.e., attractiveness, a subcategory of Appearance). Largely, these patterns align with our expectation that dimensions from low-dimensional models would be highly prevalent, although some visually salient dimensions were also highly prevalent. Additionally, we found significant prevalence of content dimensions proposed by previous high-dimensional models (Nicolas et al., 2022). However, some differences arise: for example, while impressions of social media photos, which include filters, resulted in relatively high prevalence of the political-religious Beliefs dimension, our plain photos of faces showed limited prevalence of this dimension. Potentially, perceivers form impressions of others' Beliefs with

TABLE 3. **Dimension Means and Significance of Pairwise Comparisons, Dictionary Prevalence Coding**

Study 1			Study 2		
Dimension	**Mean**	***SE***	**Dimension**	**Mean**	***SE***
Appearance	0.275	0.004	Appearance	0.268	0.002
Sociability	0.236	0.004	Sociability	0.199	0.002
Emotion	0.149	0.003	Assertiveness	0.133[a]	0.002
Ability	.119[a]	0.003	Social Groups	0.132[a]	0.002
Morality	.117[a]	0.003	Emotion	0.128[a]	0.002
Social Groups	.112[a]	0.002	Ability	0.123[a]	0.002
Assertiveness	.084[b]	0.002	Morality	0.098	0.002
Beauty	.077[b]	0.002	Beauty	0.076[b]	0.002
No match	0.034	0.001	No match	0.065[b]	0.002
Status	.030[c]	0.001	Status	0.034[c]	0.002
Other	.027[c,d]	0.001	Uniqueness	0.027[c,d]	0.002
Uniqueness	.024[d,e]	0.001	Occupation	0.026[c,d]	0.002
Family	.021[e,f]	0.001	Health	0.020[d,e]	0.002
Health	.020[e,f]	0.001	Other	0.019[d,e]	0.002
Occupation	.019[e,f,g]	0.001	Beliefs	0.017[d,e]	0.002
Geography	.018[f,g]	0.001	Geography	0.012[e]	0.002
Beliefs	.015[g]	0.001	Family	0.011[e]	0.002

Note. Mean values are proportions of responses (ranging from 0 to 1). Values in a column sharing a superscript are not significantly different from each other, $p > .05$. The Beauty dictionary is a subset of the Appearance dictionary but included separately due to its relevance to existing models. The "Other" dimension includes content such as artistic, academic, and philosophical words that fell into smaller dictionaries (Nicolas et al., 2022). Values add up to more than 1 due to the inclusion of the overlapping Beauty dictionary, and the fact that responses may be coded into multiple dimensions.

very low frequency from faces alone (vs. social media photos or social group stereotypes; Koch et al., 2016, 2020).

Direction. The direction of impressions is the variable most often measured in the literature, usually in the form of scales. Direction indicates where in a semantic differential impressions fall, and it has been shown to be predictive of a multitude of outcomes (Todorov et al., 2015). Average patterns of direction across dimensions shed light on the nature of impressions (Abele et al., 2021), information integration (Nicolas & Fiske, 2023), and biases in AI (Nicolas & Caliskan, 2024b), among others.

There were significant differences in dimensional direction, $F(8, 11{,}104) = 199.12$, $p < .001$, $\eta^2 = 0.13$, 95% CI = [.11, .14] (see Figure 3). In line with our general positivity prediction, most evaluations were about high directions (e.g., faces were mostly

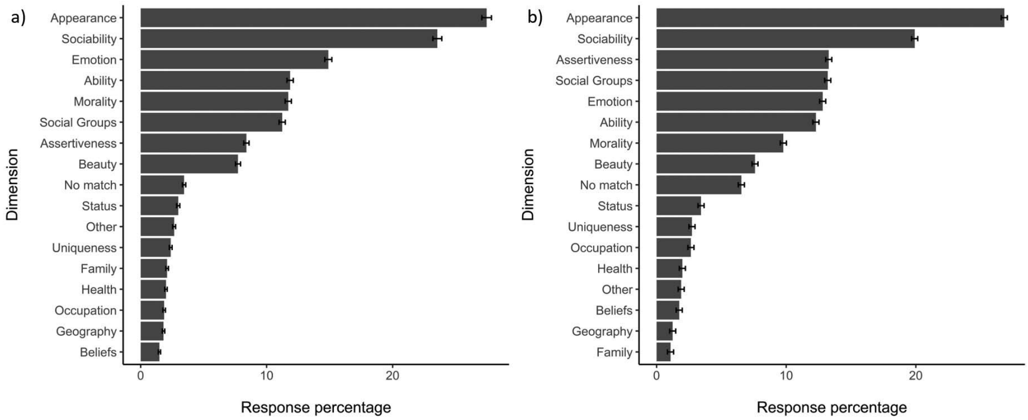

FIGURE 2. Prevalence of dimensions of face impressions: panel a) Study 1, panel b) Study 2. Error bars indicate ±1 standard errors. Dictionary coding of responses is shown as percentages for each dimension. Due to its relevance, Beauty is shown separately, but it is also counted within the overarching Appearance dimension.

evaluated as attractive and sociable). However, Health (e.g., "tired") and Beliefs (e.g., "liberal") impressions tended to be low, partially supporting our hypothesis about specific dimensions' direction.

Correlational Structure. Examining the correlations between impressions facilitates connection with lower-dimensional models. We provide a clustering solution to reduce dimensionality, exploring whether higher-order patterns in the taxonomy fall into meaningful clusters.

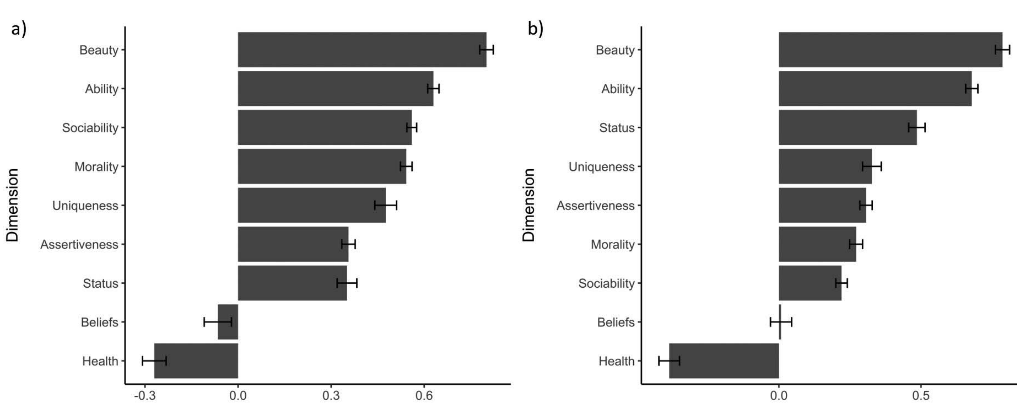

FIGURE 3. Direction of dimensions of face impressions: panel a) Study 1, panel b) Study 2. Error bars indicate ±1 standard errors. Due to its relevance, Beauty is shown separately, but it is also counted within the overarching Appearance dimension. Direction is computed from dictionaries only.

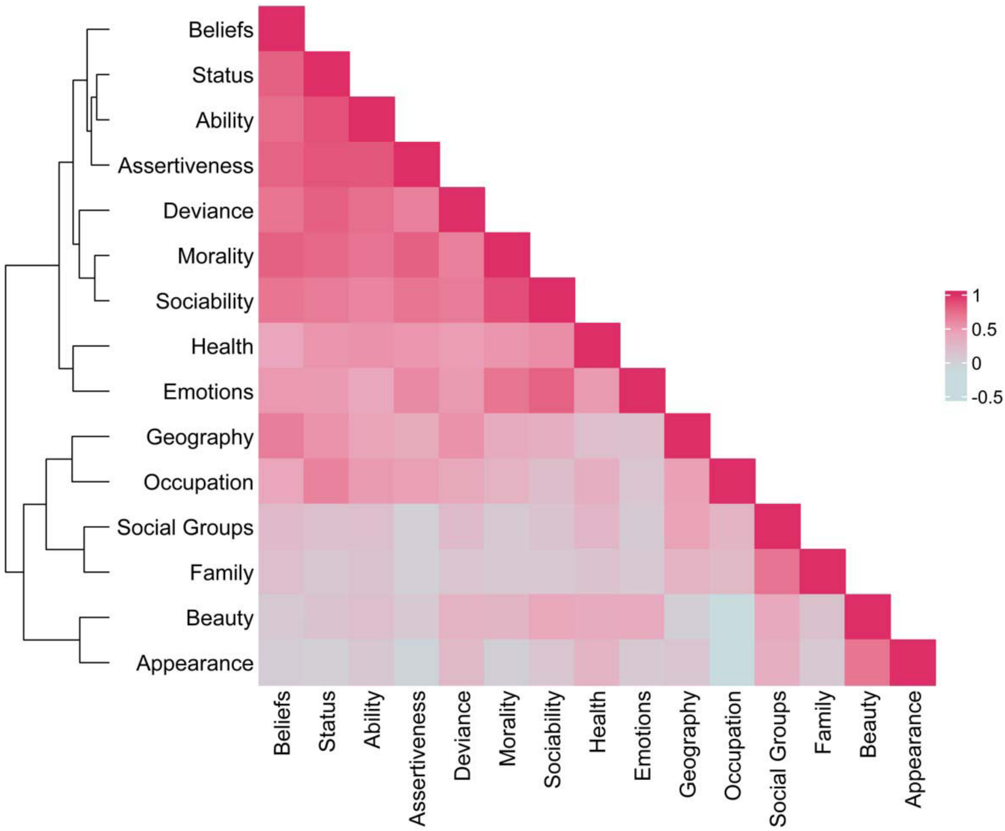

FIGURE 4. Word embedding prevalence scores correlations and hierarchical cluster, Study 1. Color indicates correlations between the word embeddings prevalence scores for the dimensions. Hierarchical cluster was conducted on the correlation matrix, showing a data-driven grouping of related dimensions, based on their tendency to occur together in face impressions.

Clustering based on text-embeddings prevalence correlations aligned with an overall internal versus external distinction of impressions, with group membership and other mixed categories overlapping with both internal and external features. The correlation patterns also support the facet structure of Communion (Morality + Sociability) and Agency (Ability + Assertiveness, here including the highly related Status dimension; Fiske et al., 2002), in line with low-dimensional models (see Figure 4).

Note that this correlational structure refers to the extent to which, when one response is about a dimension, it is also about another dimension (i.e., prevalence). Also note that while correlations largely reflect semantic relatedness, they may also reflect social biases present in natural language (e.g., the word "gay" as a general negative term; Luccioni & Viviano, 2021; Nicolas & Skinner, 2012).

STUDY 2

Study 1 characterized a taxonomy of first impressions of AI-generated faces. The content largely aligned with a previous high-dimensional taxonomy of face impressions (Connor et al., 2024), suggesting generalizability to more focused facial stimuli (e.g., without filters). We also described the coverage, prevalence, direction, and correlational structure of the taxonomy. However, Study 1 had some limitations. For example, while presenting only one face per participant has the advantage of minimizing contrast or assimilation effects due to rating a face among other faces, it prevents computation of reliability measures. To balance strengths and limitations, Study 2 asked for impressions of multiple faces, rated twice per participant. Additionally, faces in Study 1 may be representative of face photographs, but may be particularly susceptible to positive self-presentation (e.g., smiles). In Study 2 we addressed this issue by including only neutral faces. Study 2 also allowed participants to provide as many impressions as preferred (up to 10), to examine robustness to instruction variation. Previous research has used responses from 1 to 10 and found largely congruent results (Connor et al., 2024; Nicolas et al., 2022). In general, we expected these changes to replicate the taxonomy, with reduced, but still highly prevalent, Sociability content.

Moreover, Study 2 presents evidence for the taxonomy's predictive value. That is, here we provide evidence that the proposed dimensions improve predictions of relevant psychological outcomes. First, we examine the full taxonomy's improvements in prediction of global valence evaluations of faces as compared to traditional low-dimensional models. Global evaluations relate to real-world outcomes for social targets (e.g., Wallace et al., 2005) and are often used to measure the relevance of impression dimensions (e.g., Goodwin et al., 2014).

Second, in addition to broad predictions of global evaluations, higher-dimensional taxonomies may provide information about evaluations and decision making in specific contexts, where low-dimensional models may be too broad or not cover the relevant dimensions (see Nicolas et al., 2022). In fact, while improvement in a broad measure of valence may be moderate given already high correlations between the low-dimensional model's impressions (particularly along Communion) and general valence (e.g., Todorov, 2008; Todorov et al., 2008), additional benefits may be particularly relevant for more specific outcomes that have high dimension-outcome congruency with alternative dimensions (see Ajzen & Fishbein, 1977). Thus, we determined whether the proposed taxonomy improves predictions of decision making across specific but socially relevant contexts.

Using hypothetical scenarios where participants rated how much they would prioritize a person (based on their facial appearance) for health care services or discrimination protections (among others), we test whether the various dimensions of the taxonomy capture meaningful impressions that perceivers use in social decisions, above and beyond Communion, Agency, and Attractiveness. These tests are meant to show the relevance of specific dimensions for specific contexts, rather than testing the effect of the whole taxonomy for a global outcome (as in the global

TABLE 4. **Decision-Making Scenarios and Predictors**

Scenario Name	Scenario Item ("How much would you prioritize this person for programs . . .")	Hypothesized Predictors
Health care	. . . increasing access to health care	Health; Appearance
Counseling	. . . making emotional counseling more available to them	Emotions; Uniqueness
Inclusion	. . . aimed at ensuring they feel included in their community or workplace	Uniqueness
Immigration	. . . aimed at ensuring they are not unfairly stopped by immigration officials	Geography; Appearance
Hiring	. . . aimed at detecting/preventing discrimination based on LinkedIn profiles	Appearance
Face recognition	. . . aimed at detecting/preventing discrimination in facial recognition technologies	Social Groups
Polarization	. . . aimed at educating them about political polarization	Beliefs
Income	. . . aimed at providing them with additional sources of income	Status

evaluations models described above). Our predictions about which dimensions will be most relevant for a given scenario (shown in Table 4) are based on the principles of compatibility, which suggest a stronger attitude–behavior link when the attitude measured more closely matches the behavior predicted (e.g., principles of compatibility; Ajzen & Fishbein, 1977). Similarly, we hypothesize specific stereotype dimensions to be predictive, above and beyond low-dimensional models, if they refer to context similar to the decision-making scenario (e.g., Health dimension and health care decisions; Beliefs dimension and political polarization). However, other dimensions in the expanded taxonomy may also predict the decision-making scenarios, which we tested in exploratory analyses using the full taxonomy to predict decision making. In general, our argument is that perceivers may use a wider variety of impression dimensions to inform their decisions in specific, yet relevant, contexts (Todorov, 2009).

Method

Unless noted, Study 2 used the same methods as Study 1. Study 2 was preregistered (https://aspredicted.org/QWG_4CH).

Participants. Participants were 1,381 Amazon Mechanical Turk workers (after exclusion of 28 participants for responding too quickly and having negative reliability; exclusions did not affect conclusions). Participants' mean age was 43.14 years; most identified as men (49.5%; 48.2% women) and White (68.9%; 15.9% Black; 7.8% Asian; 5% multiracial; 8.4% Hispanic).

Materials and Procedure. We obtained 210 neutral stimuli faces from a fine-tuned version of Study 1's model. The fine-tuning used faces from multiple sources (e.g., the FFHQ dataset; Karras et al., 2021) that were coded as "neutral in appearance or minimally expressive" by human coders and automated methods (Albohn et al., 2019). As in Study 1, these stimuli achieved acceptable quality levels.

Each participant rated three faces, twice across two blocks. Faces were presented one at a time, in random order (each face seen by an average of 20 participants). Participants saw the same instructions as in Study 1, but were informed they would see each face twice, and were asked to "Please type as many characteristics as you think about," followed by 10 text boxes.

After the open-ended task, participants saw the same faces and rated each on direction along multiple dimensions, using a 1 (*not at all*) to 5 (*a lot*) scale. We asked: "To what extent do you view this person as . . . ," followed by items (in parentheses: dimension name): "sociable" (Sociability), "trustworthy" (Morality), "intelligent" (Ability), "confident" (Assertiveness), "conservative" (Beliefs), "wealthy" (Status), "healthy" (Health), "experiencing a positive emotion" (Emotion), "unique/different from most people" (Uniqueness), "American" (Geography), "physically attractive" (Beauty/Appearance), and "having recognizable features" (Appearance). After rating each face twice, in separate blocks, they saw them again to rate general valence in a 1 (*very negatively*) to 5 (*very positively*) scale ("In general, how do you feel about this person?").

Subsequently, participants saw each face once more to rate them across decision-making scenarios, using a 1 (*none at all*) to 5 (*a great deal*) scale. We asked: "Suppose you are in a decision-making position for the scenarios below. How much would you prioritize this person for programs . . . ," followed by multiple scenarios presented in Table 4. This table also includes the specific dimensions of the taxonomy that we expected to predict each scenario. We aimed to include at least one scenario for each of the more novel dimensions, to understand their predictive value. Finally, participants completed demographic questions.

Analysis Strategy. Study 2 includes all the analyses described in Study 1 (aggregation was based on responses provided per participant). To compute reliability, we ran mixed models with the variable of interest as outcome, an intercept, and three random factors: participant, stimulus, and participant-by-stimulus. Reliability is the sum of all the intraclass correlation coefficients (ICCs), except the residual.

To predict general valence we used the *glmnet* package (Tay et al., 2023) for regularized regressions, comparing the predictive value of the low-dimensional models with the extended taxonomy's. Regularized regression reduces overfitting when there are many predictor dimensions (provided R^2s already account for number of predictors). We selected the lambda that minimized mean cross-validated error for each model. As alternative analyses, we provide traditional χ^2 model comparisons with significance testing, as well as Akaike information criterion (AIC) values. A decrease in AIC of at least 2 points tends to indicate an improvement in predictivity, controlling for overfitting (Burnham & Anderson, 2004).

To predict decision-making scenarios, we used linear mixed models (participant and stimulus as random intercepts), with the hypothesized dimensions' prevalence or scale values (direction) as predictors of their corresponding decision rating. We note that we had no specific hypotheses about whether prevalence, direction, or both would be most predictive for each scenario. Instead, evidence that at least one of these variables for a dimension predicted outcomes would satisfy our goal of showing they can provide relevant information. We controlled for Morality and Sociability (Communion facets), Ability and Assertiveness (Agency facets), and Attractiveness scales, since our goal was to show the unique value of the high-dimensional taxonomy dimensions, above established dimensions. As an alternative exploratory analysis, we also ran regularized regressions using all dimensions as predictors, as above, for each decision-making outcome, to see if dimensions beyond those hypothesized also added predictive value.

Results and Discussion

Main results replicating Study 1 are presented in Figures 3 and 4, and Table 3, while further details and additional analyses are in the Supplement. As shown, despite differences in methods (e.g., neutrality of faces), both studies' results are largely congruent. Small differences include a lower rate of Sociability content (but maintaining its ranking) and higher rate of Assertiveness impressions in Study 2, potentially driven by the more neutral faces.[2]

Reliability. To understand how responses to the same face in the first and second block related per participant, we provide reliability measures. First, on average, participants provided 5.75 (SD = 2.34) responses per face, and this number correlated at r = .85 between blocks. For prevalence variables, test–retest reliability was fair to good (Cicchetti, 1994) across dimensions (M = .62, range: .481–.718; see Supplement for full table). Reliability was higher for direction (M = .85, range: .750–.935), suggesting that given that a dimension comes to mind, it is evaluated more consistently. This is in line with reliability for the traditional scales, which also measure direction (M = .824, range: .771–.918). This difference between prevalence and direction was expected as what content comes to mind across two blocks in the same experiment is more susceptible to order effects and interpretation of instructions (e.g., participants' interpretation about whether the same responses are expected across two open-ended replicates may vary).

Prediction of Global Evaluations of Faces. Next, we were interested in testing whether the high-order taxonomy predicts general positive–negative impressions of faces in a way that justifies the loss of parsimony of lower-dimensional models. First, a baseline regularized model using scales for facets of Communion, Agency, and Attractiveness as predictors resulted in an R^2 = .525. A model including all the

2. In previous research using social groups, Assertiveness prevalence did not vary as a function of number of responses (Fiske et al., 2021), but future research should further explore change-over-responses to examine this pattern in face impressions.

scale items as predictors, covering additional dimensions such as Health, Emotion, Beliefs, Status, Uniqueness, Geography (foreignness), and Appearance, increased variance explained (R^2 = .554). Adding prevalence metrics for all dimensions further improved predictiveness, R^2 = .566. Because most dimensions are evaluative (i.e., they include a valenced component), increases are moderate when looking at general valence prediction. For example, a model with just Communion as a predictor already accounts for 57.4% of the variance on global evaluations. Adding the well-established second dimension of Agency improves variance explained to only 57.9% (an R^2 change of .005). Based on recent guidelines (Funder & Ozer, 2019), the R^2 change of .03–.04 that we find here is considered a "medium" effect. Thus, the improvement in prediction of global evaluation from any one evaluative dimension is likely to be small, but even such small changes have been shown to be meaningful, for example in models adding Agency and Attractiveness to a Communion-only model (e.g., Oosterhof & Todorov, 2008). See also Funder and Ozer (2019) for further discussion of the potential cumulative psychological impact of smaller effect sizes.

An alternative approach to regularized regression is using model comparisons with p values and/or AICs. A model including all dimensions' scales (marginal R^2 = .559) was significantly better than a model with just Communion, Agency, and Attractiveness (marginal R^2 = .523), χ^2 = 218.93, p < .001, AIC = 8272.1 vs. 8477.1. Further adding prevalence variables also improved the all-scales model (marginal R^2 = .567), χ^2 = 83.27, p < .001, AIC = 8208.9 vs. 8272.1. AIC differences as small as 2 are often considered sufficient evidence of improvement in predictivity when controlling for overfitting (Burnham & Anderson, 2004). Our AIC difference between the high-dimensional model (prevalence + direction) and the low-dimensional model is 268.2 points.

These analyses suggest that even for a broad variable such as global evaluations, the expanded taxonomy improves variance explained.

Prediction of Decision Making. As shown in Table 5, the dimensions from the expanded taxonomy provide value in predictions of decision making about faces, even when controlling for Communion, Agency, and Attractiveness. Specifically, all the dimensions from the taxonomy predicted decision making across a range of practically relevant scenarios. The Social Groups dimension predicting the face recognition scenario was the only unsupported preregistered hypothesis. However, the closely related Geography dimension was predictive, p < .001.

While our argument is that the extended taxonomy's greater diversity of dimensions allows for better matching of impressions and outcomes, thus improving predictions, it is possible that other dimensions that we did not hypothesize to be relevant for a specific decision-making scenario do indeed add predictive value. In an exploratory analysis, we use the same regularized regression approach as for global evaluations, but for each of the decision-making outcomes. Results, shown in Table 6, again suggest that using the extended taxonomy, both direction and prevalence indicators, improved variance explained over the baseline Agency

TABLE 5. Decision-Making Models Using Hypothesized Dimensions

Scenario	Predicted dimension	*B*	*p*
Health care	Health Prevalence	0.047	< .001
	Health Direction (healthy)	–0.026	.162
	Appearance Prevalence	0.015	.407
	Appearance Direction (physically attractive)	0.039	.024
	Appearance Direction (having recognizable features)	0.074	< .001
Counseling	Emotion Prevalence	0.101	< .001
	Emotion Direction (experiencing a positive emotion)	–0.047	.018
	Uniqueness Prevalence	0.014	.307
	Uniqueness Direction (unique/different from most people)	0.093	< .001
Inclusion	Uniqueness Prevalence	0.012	.358
	Uniqueness Direction (unique/different from most people)	0.137	< .001
Immigration	Geography Prevalence	0.027	.038
	Geography Direction (American)	–0.17	< .001
	Appearance Prevalence	–0.006	.73
	Appearance Direction (physically attractive)	0.106	< .001
	Appearance Direction (having recognizable features)	0.05	.003
Hiring	Appearance Prevalence	–0.001	.966
	Appearance Direction (physically attractive)	0.078	< .001
	Appearance Direction (having recognizable features)	0.065	< .001
Face recognition	Social Groups Prevalence	0.0003	.848
Polarization	Beliefs Prevalence	0.028	.028
	Beliefs Direction (conservative)	0.061	< .001
Income	Status Prevalence	–0.017	.199
	Status Direction (wealthy)	–0.142	< .001

Note. Direction results shown for scale predictors (item in parentheses). Controlling for Sociability and Morality (Communion), Ability and Assertiveness (Agency), and Attractiveness.

+ Communion + Attractiveness model. An alternative using *p* values and AIC showed congruent results (see Supplement).

GENERAL DISCUSSION

We characterized a high-dimensional taxonomy of spontaneous face impressions across two studies and a large set of realistic and diverse AI-generated face photographs. This taxonomy includes the dimensions shown in Figure 5. We delineated the taxonomy's properties by establishing its coverage of spontaneous impressions, the prevalence and direction of various dimensions of content, dimensional intercorrelations, and predictive value. Furthermore, we tested the robustness of our findings using a variety of methods.

TABLE 6. Exploratory Decision-Making Regularized Models Showing Predictive Value of the Extended Taxonomy

Scenario	M0: Baseline R^2	M1: Extended (direction only) R^2	M2: Extended (direction + prevalence) R^2	M2 vs. M0 R^2 Difference	M2 vs. M0 R^2 Ratio
Health care	0.089	0.112	0.117	0.03	1.33
Counseling	0.059	0.082	0.090	0.03	1.53
Inclusion	0.104	0.137	0.142	0.04	1.37
Immigration	0.063	0.129	0.132	0.07	2.08
Hiring	0.083	0.112	0.118	0.04	1.42
Face recognition	0.081	0.108	0.110	0.03	1.36
Polarization	0.055	0.073	0.073	0.02	1.34
Income	0.083	0.103	0.111	0.03	1.34

Note. Baseline model includes only Agency, Communion, and Attractiveness scale (direction) ratings. M1 adds all dimensional scale ratings, and M2 also adds prevalence scores. Difference scores are $R^2M2–R^2M0$, Ratio scores are R^2M2 / R^2M0. All R^2s account for overfitting since they are retrieved from regularized models.

Previous research using open-ended face impressions had used social media stimuli including filters and sometimes prominent backgrounds (Connor et al., 2024). Here, we used AI-generated images that focused much more on facial information. First, we found that the same set of dimensions explained >95% of impressions provided by our participants (i.e., coverage). For comparison, traditional content from low-dimensional models (Oosterhof & Todorov, 2008; Sutherland et al., 2018), such as Communion and Agency, accounted for approximately half of impressions. Thus, the identified high-dimensional model is a good fit for organizing the nuanced content that arises in spontaneous impressions of faces. High-dimensional models may also be a good fit for naturalistic images, which show more diversity in demographics and emotion than more controlled and standardized images (Todorov & Oh, 2021). Gaps in our understanding of the content of impressions can lead us to ignore relevant evaluations that may affect behavior.

Second, prevalence results point at differences in the primacy (Abele et al., 2021) of various dimensions in spontaneous face impressions, as some dimensions come to mind more often than others. In line with low-dimensional models, Morality and Assertiveness were among the most prevalent dimensions. On the other hand, in contrast with group stereotypes (Koch et al., 2016, 2020), Beliefs was one of the least prevalent dimensions. This finding also contrasts with the previous high-dimensional taxonomy of face impressions, potentially due to Beliefs impressions in this previous research being driven higher by nonfacial information or politically themed filters applied to social media photos (Connor et al., 2024). The taxonomy also helps differentiate dimensions that have often been examined in conjunction in the past (e.g., Sociability and Morality as facets of Communion), but that show a clear dissociation in face impressions (e.g., Sociability is more prevalent than Morality content). Thus, prevalence data align and confirm the primacy

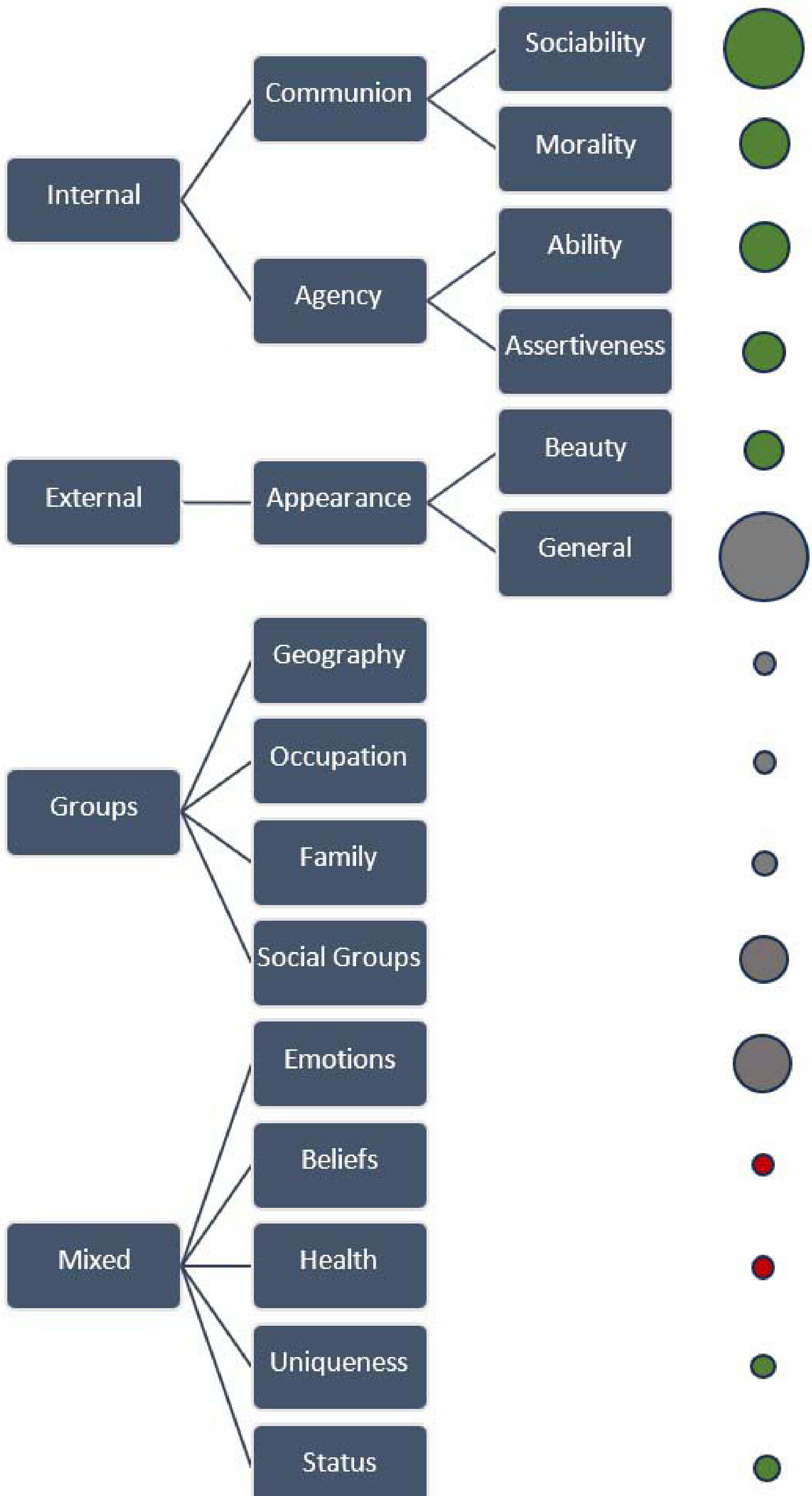

FIGURE 5. High-dimensional spontaneous face impressions taxonomy and properties. Proposed hierarchical organization based on a synthesis of the studies across methods and the existing literature. Size of circles indicates prevalence (based on Study 1 dictionary analysis) and color indicates average direction (green = high; red = low; grey = NA).

of Communion and Agency, albeit with differences in their facets. However, this also reveals that a complete understanding of the content of impressions requires a higher-dimensional taxonomy, including content related to uniqueness, health, emotions, social groups, geography, status, beliefs, and appearance.

Third, our results provide insight into the directionality of face impressions. Overall, the direction of impressions was high (which, for most dimensions, means positive in valence), potentially reflecting a general tendency of positivity towards relatively neutral stimuli (Matlin & Stang, 1978). This differs from negativity biases in other tasks (e.g., information integration; Nicolas & Fiske, 2023) or domains (e.g., stereotypes; Nicolas et al., 2022). The distinction may lie in face impressions not depending on retrieval of negatively biased information, such as for stereotypes, or perhaps because face photographs do not typically exhibit negative expressions (Wu et al., 2015).

Fourth, dimensional correlations support a separation of external features and internal traits in impressions. These correlations also largely support the facet structure of Communion and Agency (Abele et al., 2021), with high correlations among Status, Ability, and Assertiveness and between Sociability and Morality.

Finally, the expanded taxonomy's dimensions show predictive value. Compared to the low-dimensional model's dimensions, the full taxonomy moderately improved predictions of global valence evaluations of the faces. The full taxonomy also improved predictions of decision making both for the specific hypothesized dimensions and in models using the full extended taxonomy. In fact, results suggest a greater impact of the extended taxonomy on these decision-making scenarios than on global evaluations. Specifically, although R^2 *differences* between baseline and full-taxonomy models are similar across outcomes, given the already high correlations between low-dimensional models' dimensions and general valence (e.g., Todorov, 2008), the R^2 *ratios* for the decision-making outcomes are much larger (up to more than doubling explained variance) than for global evaluations (an ~8% increase). In other words, dimensions beyond Communion, Agency, and Attractiveness matter for global evaluations, but are particularly valuable for decision making on specific but relevant contexts. Thus, the prevalence of these additional dimensions in face impressions may point to their function in navigating the diversity of real-world decisions that perceivers encounter. Furthermore, both prevalence and direction indicators provided unique information, underscoring the value of measuring both properties of impressions (Nicolas et al., 2022). This suggests that the taxonomy improves our understanding of the consequences of impressions for both general and context-specific outcomes.

In Figure 5 we present a synthesis of our results, showing our proposed hierarchical organization of the dimensions identified. This organization connects our taxonomy to existing low-dimensional models of content (Lin et al., 2021; Oosterhof & Todorov, 2008; Sutherland et al., 2013) and identifies distinctions (e.g., internal vs. external attributions) that may explain discrepancies and facilitate theoretical integration.

The current taxonomy, given how comprehensively it accounts for perceivers' impressions (coverage), may provide a high-dimensional framework that researchers seeking more parsimonious solutions could work from. For example, researchers may use only the most prevalent dimensions, use overarching dimensions (vs. facets), or use only those dimensions expected to be most predictive in specific contexts. We provide information about these different kinds of "primacy" (Bai et al., 2024) to allow for flexibility in balancing nuance and comprehensiveness with parsimony and resources. That is, we expect this taxonomy to provide flexibility in modeling face impressions along a low-dimensional to high-dimensional spectrum, given the desired balance of parsimony and generality versus nuance and specificity. Given the centrality of faces to social interaction, understanding the content of face impressions has numerous implications. Here, we show initial evidence of the role of the expanded taxonomy in evaluation and decision making (in line with previous evidence in the stereotype domain; Nicolas et al., 2022). Future research may show the role of the expanded taxonomy in a much larger set of outcomes, as has been done for established dimensions (e.g., criminal justice, voting behavior; Bagnis et al., 2020; Olivola et al., 2014; Todorov et al., 2015).

STRENGTHS, LIMITATIONS, AND FUTURE DIRECTIONS

We expect our findings to serve as an initial baseline of the characteristics of a high-dimensional taxonomy of face impressions (Connor et al., 2024), but we also expect that the patterns presented here may vary based on context, culture, individual differences, and other factors. This should provide a hypothesis-generating opportunity to examine the dynamic nature of face impressions along both direction and prevalence, and a more comprehensive set of dimensions. For example, future studies may measure differences in the extent to which participants vary along the taxonomy dimensions or their beliefs about how these dimensions correlate (Stolier et al., 2018). In addition, while we found minimal differences by varying one type of context (i.e., whether faces were evaluated in isolation or sequentially along other faces), other types of contextual information (e.g., dress or environmental information; Cesario, 2022; Hester & Hehman, 2023; Oh et al., 2020) may alter the content of face impressions.

Here, we aimed to characterize content without dimensionality-reduction methods that rely on interstimulus variability (e.g., factor analysis). When applied to data across participants, the latter methods may provide misleading conclusions when impressions are not shared equally across participants. For example, whether different participants believe a face is beautiful or moral may depend on the perceiver's social identities (Albohn et al., 2024, 2025; Koch et al., 2020; Martinez et al., 2020). Thus, aggregating across participants can mask stable individual differences, resulting in an "averaging out" of evaluations. Because our methods operate directly on all participants' responses, they are not affected by this limitation. Notwithstanding these notes, previous research using principal components analysis or factor analysis on similarly organized data still found a largely congruent high-dimensional model of face impressions (Connor et al., 2024).

An issue, present to some degree in all face databases, is the inability to completely control for emotional expressions (but see Albohn et al., 2025). This may induce biases in descriptive analyses, which may be heavily influenced by predominant expressions in the stimuli. To illustrate, photographs may be biased towards positive self-presentation (e.g., smiles), leading, for example, to more Sociability impressions than expected. However, Study 2's neutral faces largely replicated the taxonomy. Sociability remained among the most prevalent dimensions (however, direction for Sociability and Morality was lower in neutral faces; other differences in prevalence, such as slightly lower Morality and higher Assertiveness occurred for neutral faces). Furthermore, multiple real-world contexts may also reflect self-presentational tendencies, including face impressions in social media, CVs, and politician photographs, among others, with relevant implications (e.g., biases in hiring). In fact, naturalistic photographs have several advantages compared to standardized stimuli, including potentially higher ecological validity (Burton et al., 2011). AI-generated images in particular are increasingly used by the general public using commercial AI and by psychologists looking to create high-quality stimuli with few resources. As such, our findings help further establish the generalizability of high-dimensional psychological models (e.g., SSCM) and their ability to explain impressions of influential real-world and research facial stimuli. Nonetheless, future research should continue to examine the generalizability of the proposed high-dimensional taxonomy to additional stimulus types in acknowledgment of unique features that social media and AI-generated faces may show (e.g., potentially higher attractiveness of AI-generated faces; Nyce, 2023).

In general, we advocate a multimethod approach to balance methodological strengths and limitations. For example, dictionary coding has limitations, such that they are all-or-nothing codes, but words relate to content to different degrees. In the Supplement, we explore alternative methods (e.g., text embeddings similarities), with highly congruent results.

CONCLUSION

Faces are a significant source of information in social interactions. Here, we describe the coverage, content prevalence, directionality, dimensional intercorrelations, and predictive value of a taxonomy of spontaneous impressions of naturalistic face photographs. We used cutting-edge interdisciplinary methods from AI language models and computer vision to provide an account of impressions of a highly diverse and realistic set of human faces. A taxonomy of face impressions that is more comprehensive than current low-dimensional models is a foundation for further theoretical development and practical applications in an area central to human behavior.

Funding statement. This work was supported by the Richard N. Rosett Faculty Fellowship at the University of Chicago Booth School of Business.

Data and code availability. https://osf.io/pmtxw/?view_only=a485688056a8495ba37082f38c690833

REFERENCES

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

Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, *84*(5), 888–918. https://doi.org/10.1037/0033-2909.84.5.888

Albohn, D. N., Brandenburg, J. C., & Adams, R. B. (2019). Perceiving emotion in the "neutral" face: A powerful mechanism of person perception. In U. Hess & S. Hareli (Eds.), *The social nature of emotion expression: What emotions can tell us about the world* (pp. 25–47). Springer International Publishing. https://doi.org/10.1007/978-3-030-32968-6_3

Albohn, D. N., Martinez, J. E., & Todorov, A. (2024). Determinants of shared and idiosyncratic contributions to judgments of faces. *Journal of Experimental Psychology: Human Perception and Performance*, *50*(11), 1117–1130. https://doi.org/10.1037/xhp0001239

Albohn, D. N., Uddenberg, S., & Todorov, A. (2025). Individualized models of social judgments and context-dependent representations. *Scientific Reports*, *15*, 4208. https://doi.org/10.1038/s41598-025-86056-1

Antonakis, J., & Dalgas, O. (2009). Predicting elections: Child's play! *Science (New York, N.Y.)*, *323*(5918), 1183. https://doi.org/10.1126/science.1167748

Bagnis, A., Caffo, E., Cipolli, C., De Palma, A., Farina, G., & Mattarozzi, K. (2020). Judging health care priority in emergency situations: Patient facial appearance matters. *Social Science & Medicine (1982 260*, 113180. https://doi.org/10.1016/j.socscimed.2020.113180

Bai, X., Nicolas, G., & Fiske, S. T. (2024). Social stereotypes: Content and process. In D. Carlston, K. Johnson, & K. Hugenberg (Eds.), *The Oxford handbook of social cognition* (2nd ed.; pp. 442–470). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780197763414.013.16

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. https://doi.org/10.18637/jss.v067.i01

Becker, C., & Laycock, R. (2023). Embracing deepfakes and AI-generated images in neuroscience research. *European Journal of Neuroscience*, *58*(3), 2657–2661. https://doi.org/10.1111/ejn.16052

Boyd, A., Tinsley, P., Bowyer, K., & Czajka, A. (2023). The value of AI guidance in human examination of synthetically-generated faces. *Proceedings of the Thirty-Seventh AAAI Conference on Artificial Intelligence and Thirty-Fifth Conference on Innovative Applications of Artificial Intelligence and Thirteenth Symposium on Educational Advances in Artificial Intelligence*, *37*, 5930–5938. Association for the Advancement of Artificial Intelligence. https://doi.org/10.1609/aaai.v37i5.25734

Burnham, K. P., & Anderson, D. R. (2004). Multimodel inference: Understanding AIC and BIC in model selection. *Sociological Methods & Research*, *33*(2), 261–304. https://doi.org/10.1177/0049124104268644

Burton, A. M., Jenkins, R., & Schweinberger, S. R. (2011). Mental representations of familiar faces. *British Journal of Psychology (London, England: 1953)*, *102*(4), 943–958. https://doi.org/10.1111/j.2044-8295.2011.02039.x

Cesario, J. (2022). What can experimental studies of bias tell us about real-world group disparities? *Behavioral and Brain Sciences*, *45*, e66. https://doi.org/10.1017/S0140525X21000017

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*(4), 284–290. https://doi.org/10.1037/1040-3590.6.4.284

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

Connor, P., Nicolas, G., Antonoplis, S., & Koch, A. (2024). Unconstrained descriptions of Facebook profile pictures support high-dimensional models of impression formation. *Personality and Social Psychology Bulletin*. Advance online publication. https://doi.org/10.1177/01461672241266651

Deska, J. C., Kunstman, J. W., Bernstein, M. J., Ogungbadero, T., & Hugenberg, K. (2020). Black racial phenotypicality shapes social pain and support judgments. *Journal of Experimental Social Psychology*, *90*, 103998. https://doi.org/10.1016/j.jesp.2020.103998

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878–902. https://doi.org/10.1037/0022-3514.82.6.878

Fiske, S. T., Nicolas, G., & Bai, X. (2021). The stereotype content model: How we make sense of individuals and groups. In P. A. M. Van Lange, E. T. Higgins, & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (3rd ed.; pp. 392–410). Guilford Press.

Friehs, M.-T., Kotzur, P. F., Kraus, C., Schemmerling, M., Herzig, J. A., Stanciu, A., Dilly, S., Hellert, L., Hübner, D., Rückwardt, A., Ulizcay, V., Christ, O., Brambilla, M., De keersmaecker, J., Durante, F., Gale, J., Grigoryev, D., Igou, E. R., Javakhishvili, N., . . . Yzerbyt, V. (2022). Warmth and competence perceptions of key protagonists are associated with containment measures during the COVID-19 pandemic: Evidence from 35 countries. *Scientific Reports*, *12*(1), Article 1. https://doi.org/10.1038/s41598-022-25228-9

Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science*, *2*(2), 156–168. https://doi.org/10.1177/2515245919847202

Goodwin, G. P., Piazza, J., & Rozin, P. (2014). Moral character predominates in person perception and evaluation. *Journal of Personality and Social Psychology, 106*(1), 148–168. https://doi.org/10.1037/a0034726

Hester, N., & Hehman, E. (2023). Dress is a fundamental component of person perception. *Personality and Social Psychology Review*, *27*(4), 414–433. https://doi.org/10.1177/10888683231157961

Higgins, E. T. (1996). Knowledge activation: Accessibility, applicability, and salience. In E. T. Higgins & A. W. Kruglanski (Eds.), *Social psychology: Handbook of basic principles* (pp. 133–168). Guilford Press.

Jaeger, B., Todorov, A. T., Evans, A. M., & van Beest, I. (2020). Can we reduce facial biases? Persistent effects of facial trustworthiness on sentencing decisions. *Journal of Experimental Social Psychology*, *90*, 104004. https://doi.org/10.1016/j.jesp.2020.104004

Jones, A. L. (2018). The influence of shape and colour cue classes on facial health perception. *Evolution and Human Behavior*, *39*(1), 19–29. https://doi.org/10.1016/j.evolhumbehav.2017.09.005

Jones, B. C., DeBruine, L. M., Flake, J. K., Liuzza, M. T., Antfolk, J., Arinze, N. C., Ndukaihe, I. L. G., Bloxsom, N. G., Lewis, S. C., Foroni, F., Willis, M. L., Cubillas, C. P., Vadillo, M. A., Turiegano, E., Gilead, M., Simchon, A., Saribay, S. A., Owsley, N. C., Jang, C., . . . Coles, N. A. (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/s41562-020-01007-2

Karras, T., Laine, S., & Aila, T. (2021). A style-based generator architecture for generative adversarial networks. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, *43*(12), 4217–4228. https://doi.org/10.1109/TPAMI.2020.2970919

Karras, T., Laine, S., Aittala, M., Hellsten, J., Lehtinen, J., & Aila, T. (2020). Analyzing and improving the image quality of StyleGAN. *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA* (pp. 8107–8116). IEEE. https://doi.org/10.1109/CVPR42600.2020.00813

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,* 103995. https://doi.org/10.1016/j.jesp.2020.103995

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Lenth, R. V. (2016). Least-squares means: The R package lsmeans. *Journal of Statistical Software*, *69*(1), 1–33. https://doi.org/10.18637/jss.v069.i01

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), Article 1. https://doi.org/10.1038/s41467-021-25500-y

Luccioni, A., & Viviano, J. (2021). What's in the box? An analysis of undesirable content in the common crawl corpus. In C. Zong, F. Xia, W. Li, & R. Navigli (Eds.), *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (Volume 2: Short Papers)* (pp. 182–189). Association for Computational Linguistics. https://doi.org/10.18653/v1/2021.acl-short.24

Martinez, J. E., Funk, F., & Todorov, A. (2020). Quantifying idiosyncratic and shared contributions to judgment. *Behavior Research Methods*, *52*, 1428–1444. https://doi.org/10.3758/s13428-019-01323-0

Matlin, M. W., & Stang, D. J. (1978). *The Pollyanna principle: Selectivity in language, memory, and thought*. Schenkman Pub. Co.

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

Nicolas, G., & Caliskan, A. (2024a). *A taxonomy of stereotype content in large language models*. arXiv. https://doi.org/10.48550/arXiv.2408.00162

Nicolas, G., & Caliskan, A. (2024b). Directionality and representativeness are differentiable components of stereotypes in large language models. *PNAS Nexus, 3*(11), pgae493. https://doi.org/10.1093/pnasnexus/pgae493

Nicolas, G., & Fiske, S. T. (2023). Valence biases and emergence in the stereotype content of intersecting social categories. *Journal of Experimental Psychology: General*, *152*(9), 2520–2543. https://doi.org/10.1037/xge0001416

Nicolas, G., & Skinner, A. L. (2012). "That's so gay!" Priming the general negative usage of the word *gay* increases implicit anti-gay bias. *Journal of Social Psychology*, *152*(5), 654–658. https://doi.org/10.1080/00224545.2012.661803

Nicolas, G., & Skinner, A. L. (2017). Constructing race: How people categorize others and themselves in racial terms. In H. Cohen & C. Lefebvre (Eds.), *Handbook of categorization in cognitive science* (2nd ed.; pp. 607–635). Elsevier Academic Press. https://doi.org/10.1016/B978-0-08-101107-2.00025-7

Nicolas, G., Skinner, A. L., & Dickter, C. L. (2019). Other than the sum: Hispanic and Middle Eastern categorizations of Black–White mixed-race faces. *Social Psychological and Personality Science*, *10*(4), 532–541. https://doi.org/10.1177/1948550618769591

Nightingale, S. J., & Farid, H. (2022). AI-synthesized faces are indistinguishable from real faces and more trustworthy. *Proceedings of the National Academy of Sciences of the United States of America*, *119*(8), e2120481119. https://doi.org/10.1073/pnas.2120481119

Nyce, C. M. (2023, October 24). AI has a hotness problem. *The Atlantic*. https://www.theatlantic.com/technology/archive/2023/10/ai-image-generation-hot-people/675750/

Oh, D., Shafir, E., & Todorov, A. (2020). Economic status cues from clothes affect perceived competence from faces. *Nature Human Behaviour*, *4*(3), 287–293. https://doi.org/10.1038/s41562-019-0782-4

Olivola, C. Y., Funk, F., & Todorov, A. (2014). Social attributions from faces bias human choices. *Trends in Cognitive Sciences*, *18*, 566–570. https://doi.org/10.1016/j.tics.2014.09.007

Olivola, C. Y., & Todorov, A. (2010). Elected in 100 milliseconds: Appearance-based trait inferences and voting. *Journal of Nonverbal Behavior*, *34*(2), 83–110. https://doi.org/10.1007/s10919-009-0082-1

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

OpenAI. (n.d.). *Introducing ChatGPT*. Retrieved November 8, 2023, from https://openai.com/blog/chatgpt

Peterson, J. C., Uddenberg, S., Griffiths, T. L., Todorov, A., & Suchow, J. W. (2022). Deep models of superficial face judgments. *Proceedings of the National Academy of Sciences*, *119*(17), e2115228119. https://doi.org/10.1073/pnas.2115228119

R Core Team. (2024). *R: A language and environment for statistical computing* [Computer software].

Reimers, N., & Gurevych, I. (2019). Sentence-BERT: Sentence Embeddings using Siamese BERT-networks. In K. Inui, J. Jiang, V. Ng, & X. Wan (Eds.), *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)* (pp. 3982–3992). Association for Computational Linguistics. https://doi.org/10.18653/v1/D19-1410

Shen, B., RichardWebster, B., O'Toole, A., Bowyer, K., & Scheirer, W. J. (2021). A study of the human perception of synthetic faces [Paper presentation]. *2021 16th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2021), Jodhpur, India*. https://doi.org/10.1109/FG52635.2021.9667066

Stolier, R. M., Hehman, E., Keller, M. D., Walker, M., & Freeman, J. B. (2018). The conceptual structure of face impressions. *Proceedings of the National Academy of Sciences*, *115*(37), 9210–9215. https://doi.org/10.1073/pnas.1807222115

Sutherland, C. A. M., 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 & Social Psychology Bulletin*, *44*(4), 521–537. https://doi.org/10.1177/0146167217744194

Sutherland, C. A. M., 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. M., 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. M., & Young, A. W. (2022). Understanding trait impressions from faces. *British Journal of Psychology (London, England: 1953)*, *113*(4), 1056–1078. https://doi.org/10.1111/bjop.12583

Tay, J. K., Narasimhan, B., & Hastie, T. (2023). Elastic net regularization paths for all generalized linear models. *Journal of Statistical Software*, *106*(1), 1–31. https://doi.org/10.18637/jss.v106.i01

Todorov, A. (2008). Evaluating faces on trustworthiness: An extension of systems for recognition of emotions signaling approach/avoidance behaviors. *Annals of the New York Academy of Sciences*, *1124*(1), 208–224. https://doi.org/10.1196/annals.1440.012

Todorov, A. (2009). On the richness and limitations of dimensional models of social perception. *Behavioral and Brain Sciences*, *32*, 402–403. https://doi.org/10.1017/S0140525X09991014

Todorov, A. (2012). The social perception of faces. In S. T. Fiske & C. N. Macrae (Eds.), *The SAGE handbook of social cognition* (pp. 96–114). SAGE Publications.

Todorov, A. (2017). *Face value: The irresistible influence of first impressions*. Princeton University Press.

Todorov, A., & Oh, D. (2021). The structure and perceptual basis of social judgments from faces. *Advances in Experimental*

Social Psychology, *63*, 189–245. https://doi.org/10.1016/bs.aesp.2020.11.004

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

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. (2008). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences*, *12*, 455–460. https://doi.org/10.1016/j.tics.2008.10.001

Vernon, R. J. W., Sutherland, C. A. M., Young, A. W., & Hartley, T. (2014). Modeling first impressions from highly variable facial images. *Proceedings of the National Academy of Sciences*, *111*(32), E3353–E3361. https://doi.org/10.1073/pnas.1409860111

Wallace, D. S., Paulson, R. M., Lord, C. G., & Bond Jr., C. F. (2005). Which behaviors do attitudes predict? Meta-analyzing the effects of social pressure and perceived difficulty. *Review of General Psychology*, *9*(3), 214–227. https://doi.org/10.1037/1089-2680.9.3.214

Wu, Y.-C. J., Chang, W.-H., & Yuan, C.-H. (2015). Do Facebook profile pictures reflect user's personality? *Computing for Human Learning, Behaviour and Collaboration in the Social and Mobile Networks Era*, *51*, 880–889. https://doi.org/10.1016/j.chb.2014.11.014

Zebrowitz, L. A. (1990). *Social perception*. Thomson Brooks/Cole Publishing Co.