nature human behaviour

SUPPLEMENTARY INFORMATION
https://doi.org/10.1038/s41562-019-0800-6

In the format provided by the authors and unedited.

Trait knowledge forms a common structure across social cognition

Ryan M. Stolier[1*], Eric Hehman[2] and Jonathan B. Freeman[3,4*]

[1]Department of Psychology, Columbia University, New York, NY, USA. [2]Department of Psychology, McGill University, Montreal, Québec, Canada. [3]Department of Psychology, New York University, New York, NY, USA. [4]Center for Neural Science, New York University, New York, NY, USA.
*e-mail: rms2262@columbia.edu; jon.freeman@nyu.edu

Supplementary Information

The entire analysis pipeline is provided and explained in Jupyter notebooks provided on our OSF page (https://osf.io/2uzsx/).

Supplementary Methods

Trait inference task instructions (all studies). These instructions were used for all tasks in which participants rated target and trait adjective stimuli in order to measure their inferences towards stimuli or their conceptual associations (included in Studies 1 – 6).

Conceptual association tasks. Specifically, participants were instructed, "In the following task, you will be presented with a series of adjective pairs. These are human personality traits. You will be asked to rate the likelihood that individuals with one of the traits possess the other trait."

Face trait task. We instructed participants, "In this task, we ask you to indicate how [TRAIT STIMULUS] a number of different people look. You will see a person's face, and are asked to judge their likely personality traits merely from their face. Importantly, go with your gut feeling. We all make snap judgments of others constantly, so feel free to report what you think about the person based on their face. Please respond quickly with your gut feeling. There are no right or wrong answers."

Familiar trait task. We instructed participants, "We are interested in personality impressions of different individuals. In this task, we ask about personality impressions of different well-known individuals, such as politicians, historical figures, and celebrities. While you may not know these individuals directly, we ask you to report how [TRAIT STIMULUS] each person is to the best of your knowledge and ability. Importantly, go with your gut feeling. We all hold snap personality impressions of others constantly, so feel free to report what you think about the person. Please respond quickly with your gut feeling. There are no right or wrong answers."

Group trait task. We instructed participants, "We are interested in the nature of stereotypes in the United States - not in studying whether people are prejudiced or not in any way, but in what common/well-known stereotypes are (these may or may not be stereotypes you yourself hold). Importantly, we are not interested in whether you endorse stereotypes or not, but instead we are interested in stereotypes that a typical American might hold. Please answer the following questions based on what you believe the stereotypes of a typical American are. In this task, we ask that you rate whether different groups of people are stereotyped as [TRAIT STIMULUS] by the typical American. Please base these ratings on what you think common stereotypes of these groups are. Please remember that stereotypes do not necessarily need to be accurate or inaccurate, negative, positive, or neutral - they just need to be widely held ideas about personality traits or behaviors in a group."

Valence task. Participants were instructed, "In this study, we would like to understand what you think about certain personality traits. You will be presented with a series of adjectives. We will have you rate each adjective on how negative or positive you believe the personality trait to be. For instance, how negative to positive is 'smart'? After reflecting on the trait word, please provide an honest response. There are no right or wrong answers." Next, participants were reminded, "In this task, you will view a series of adjectives. These are human personality traits.

Please rate the following traits based on how negative to positive they are." The matrix is provided in Supplementary Fig. 1.

Experimental study task instructions.

Conceptual association manipulation (Study 5). Upon entering the study, participants were given an overview of the study, "In this study, we are interested in how people think of others. For instance, who do you find to be kind or smart? You will complete two tasks. First, we would like to understand what you think about personality. We will have you learn about personality traits and provide your thoughts. Second, we would like to understand how you figure people out. You will be presented photos of faces and we are going to ask you a few questions about your impressions of them". After, participants read the faux scientific article (available via https://osf.io/2uzsx/). Next, participants were asked to summarize the article, "Now that you have read about human psychology and personality, we would like to hear your thoughts about the article. Please provide a summary of the article's main points, and provide a few of your thoughts about the article. In your response, provide at least a few sentences to both summarize and provide your thoughts, and remember your summary of the main points will be used as a check that you followed instructions and completed this study". Participants were debriefed to inform them of the fictitious article and its conclusions following the study.

Conceptual association through perception learning task (Study 6). For the learning phase, participants were instructed, "These instructions are very important, please read them carefully, as you will be tested for your ability to follow them. Psychologists have found that people perceive others' personality traits from their facial appearance. In one case, people are able to tell whether other people are cautious based on what they look like to various degrees. These judgments are not perfect or consistent; however, we figure them out none the less above average. In this task, we want to measure how well you are able to tell if people are cautious based upon their appearance. Remember, answers will never always be right or wrong, but just do the best you can. You will now begin the task. Please rate whether you believe the person shown is CAUTIOUS. Use the F key for 'LESS cautious' and the J key for 'MORE cautious'. You will be given feedback as to whether your answers are correct or incorrect.". Participants were debriefed to inform them of the fictitious manipulated relationship of traits following the study.

Data preparation and analysis.

Study 1.

Representational similarity analysis (RSA). All analyses were conducted with scientific and statistical libraries in Python. No subjects were removed from these data before analysis. To assess the correspondence of trait spaces across these many domains, we applied a quantitative method from systems neuroscience, RSA[1].

Each trait space may be represented as a matrix of all pair-wise similarities (e.g., correlations) between traits, or 'similarity matrix', as measured in each domain (e.g., correlation of face impressions across all trait-pairs; see Fig 2a; Supplementary Fig. 1). Each matrix may then be flattened into a vector (i.e., variable) of unique pair-wise trait similarities, by selecting values above the diagonal (thereby removing duplicate values on the opposing side of the diagonal given the similarity matrix is symmetrical, and also removing self-similar values along the diagonal). This similarity vector holds all unique information in its respective trait space similarity matrix (for an intuitive example, see[2]). Representation of each trait space matrix as a one-dimensional vector allows traditional univariate statistical methods to test the correspondence between trait space matrices. In the current research, we measure the correspondence of trait space matrices as the Spearman rank correlation between the two matrix vectors. Rank-ordering is preferable when comparing similarity matrices from different measures as it does not assume a linear relation[1]. Therefore, to conduct our analyses we computed similarity matrices per each unique trait space (conceptual, face trait impression, familiar person knowledge, group stereotype, and NEOPI trait spaces). Each similarity matrix was then converted to a vector, then values transformed into their rank position in the vector for submission to a Spearman correlation analysis to test significance.

Similarity matrices. Similarity matrices were computed for each unique trait space. Each similarity matrix was a symmetric matrix representing the pair-wise similarities between all 15 personality trait stimuli (Fig. 2a; Supplementary Fig. 1; see 'Personality trait stimuli' in the Methods). Excluding the conceptual trait similarity matrix, all similarity matrices were computed in the same way.

Data from each task (besides the conceptual trait task) were transformed into a format in which each trait is represented as a vector, in which its features are the level of that trait across different exemplars (per trait space, exemplars were, face: unique face stimuli, familiar person knowledge: unique familiar person stimuli, stereotype content: unique social group stimuli). For each social perceptual task, we calculated the average of each trait rating per unique stimulus to create these feature vectors per trait. Therefore, each dataset was a n (15; trait stimulus) × m (number of exemplars in that task) matrix, in which each value is the trait level of a given exemplar (e.g., 'friendly' vector in the face task is the 'friendly' rating of each face exemplar in that task). We then calculated the Pearson correlation between all trait vector pairs (Pearson correlation is used as the similarity measure to create each similarity matrix, whereas Spearman correlation is used to compare them[1]), providing the pair-wise similarity between traits as measured in each trait space matrix (a total of 105 possible unique pair-wise combinations of all trait stimuli; see Fig. 2). For the conceptual trait similarity matrix, we simply computed the mean similarity rating of each unique trait-pair, providing the full matrix.

Study 2. Study 2 applied an identical analysis as Study 1 (see above). The only difference was use of new and distinct items to represent the traits in each similarity matrix, for instance, 'likelihood to compliment others' and 'likelihood to agree with others' in place of 'kindness' in the matrices. Therefore, the only difference in the Study 2 analysis pipeline was the relabeling of each of the new NEOPI behavioral description items (e.g., 'likelihood to compliment others' and 'likelihood to agree with others') to their original trait terms (e.g., 'kindness'), so that the similarity matrices could share an identical form across domains. Similarity matrices are presented in Supplementary Fig. 2).

Study 3. The face and conceptual trait space matrices were prepared in the identical strategy of that used in Study 1, however within each participant. Furthermore, we calculated the group-average conceptual trait space matrix, in order to control for consensual trait associations and target any contribution of each participants' unique and subjective conceptual associations to their face impressions. Next, a dataset was prepared to be submitted to a multilevel mixed-effects model. In this multilevel dataset, data are cross-classified between subject and trait-pair, in which each row is a trait-pair. There are four variable columns with data per each row (specific to a subject and trait-pair): subjective face trait impression correlation, subjective conceptual association, group-average conceptual association, and valence similarity (via Study 1 data for control). Analyses were performed as a linear mixed-effects model with the lmer package in R ('lme4', 'lmerTest'), applying an additional set of algorithms to assist convergence ('brms'). All variables were *z*-normalized within participants to assist model convergence. Random slopes and intercepts were allowed for all predictor variables.

Study 4. In Study 4, we ask whether the amount to which each perceiver associates two trait concepts relates to the correlation between those trait inferences towards faces, familiar people, and groups. That is, we intended to test whether perceivers with weaker/stronger conceptual trait associations also show more weakly/strongly correlated inferences. To do so, within each perceiver, we calculated two variables: their conceptual and perceptual (face, person, or group) trait associations. To estimate their perceptual trait association, we calculated the Pearson correlation coefficient between both trait evaluations of the target stimuli within each participant (between the vectors of their inferences of all target stimuli on each of the two traits they were assigned). To estimate their conceptual trait associations, we averaged the two conceptual trait item responses. Therefore a single dataset was created including data from participants across all trait-pair combinations. Lastly, to test our hypothesis, we calculated the Spearman correlation between participant perceptual trait and conceptual trait associations (Spearman correlation used so as to not assume a strictly linear relationship between distances in the two spaces)[1]. Analyses were conducted across trait-pair terms, to assess the tendency of conceptual trait associations to relate to inference correlations, across trait-pairs in general.

Study 5. Participants who did not demonstrate a condition-consistent conceptual association were omitted for analyses (e.g., participants omitted if in the positive association direction condition they reported a neutral (4) or negative (< 4) association of their trait-pair, and vice versa for the negative association direction condition). In order to study how conceptual associations (e.g., negative vs. positive associations of ‘friendliness’ and ‘intellectualism’) of participants impact face impression correlations (e.g., lower or higher correlation of ‘friendly’ and ‘intellectual’ face impressions), we created a dataset where the participants’ subjective ratings were nested within participant, along the one trait they rated faces upon from their assigned trait pair (e.g., ‘friendly’ ratings for a participant assigned to both ‘friendly’ and ‘intellectual’). For each participant, the dependent variable was the average rating of each face (from Study 1 data) along the other trait from the participant’s trait pair they did not rate faces along (e.g., ‘intellectual’). This allowed us to estimate the relationship between each participant’s subjective perception of faces along one trait from their assigned trait-pair with the appearance of the faces along the other trait from the pair. This was done to reduce transparency and suspicion of the research goals (e.g., that we were interested in the association of the two traits they read about in the faux article, in the context of faces). Only faces were used in this study given potential limitations of the manipulation, suspicion, and social desirability in responses in the context of familiar person impressions and group stereotypes. Face impression tasks also benefit from relative unawareness from participants that they make such inferences and of how they do so[3]. Given participants rated faces along only one of the two traits to which they were assigned, to measure the face trait impression relation of the trait-pairs within each subject, the multilevel model predicted the appearance of faces along one trait (which participants did not judge; face rating data via Study 1; both studies used the same face stimuli) with participants’ subjective ratings of faces along the second trait in their assigned trait-pair. For instance, if a participant was assigned to ‘friendly’-‘intellectual’, and only rated faces on ‘intellectual’, we estimated their ‘friendly’-‘intellectual’ face impression association by predicting the average ‘friendly’ ratings of those face stimuli as measured in Study 1 with the participant’s ‘intellectual’ ratings of the face stimuli from the current study. In order to test impact of conceptual association direction condition, their assigned between-subjects condition was included as a contrast coded variable (-

1 for ‘Negative’, 1 for ‘Positive’). Thereby, this dataset allows us to test whether participants in the positive association direction condition show higher trait inference correlations than participants in the low correlation condition. To perform this analysis, we used a multilevel mixed effects model to regress (Study 1) average ratings of the faces on (Study 5) participant subjective ratings, condition, and their interaction (analysis performed via the lmer package in R; ‘lme4’, ‘lmerTest’). Participant ratings were group-centered. Intercepts were random but slopes were fixed.

***Study** 7*. The NEOPI trait space matrix was prepared from an open dataset (see below). Facet vectors of trait scores from many participants were Pearson correlated, measuring the similarity of actual personality traits as the correlation of these personality traits (measuring whether individuals lower/higher in one trait are lower/higher in other traits). The NEOPI trait space matrix is provided in Fig. 5.

NEOPI dataset.

All NEOPI data used to create the NEOPI trait similarity matrix (Study 7) was obtained from a publicly available dataset from prior published research[4]. Below is a summary of methods from this prior research, as well as criteria for the subset of this data utilized in the current research.

Participants. To measure the similarity structure of personality traits in the general population, we obtained a personality measurement dataset available via the Open Science Framework (OSF). In this data, a large body of participants (initial n = 334,161) completed the 300-item NEO personality test[4] (retrieved from https://osf.io/tbmh5) via a public website (http://www.personal.psu.edu/~j5j/IPIP/). In accordance with previous validity standards (publicly available by the author at http://ipip.ori.org[5]), participant responses were filtered for duplication, insufficient attentiveness, missing responses, and weak internal consistency (final n = 307,313; M_{age} = 25.2 years, SD_{age} = 10.0 years; 185,149 Female, 122,164 Male; race/ethnicity data unavailable).

Stimuli. Participants completed the 300-item NEOPI, used to measure the 30 facet subscales of the five-factor model[4]. This included a total of 300 items, with 10 items pertaining to each of the 30 total subscales of the NEOPI. An important limitation in studying the overlap of trait spaces is that semantic similarity between trait adjectives used in each task may contribute to trait inference correlation structure[6]. The strength of using the NEOPI to measure actual personality structure is its use of a wide range of self-descriptions to measure each personality trait. Rather than asking participants if they perceive themselves as each adjective (used in the social perception tasks, e.g., ‘trustworthy’), participants rated themselves on several self-descriptions that correspond to the personality construct in question (e.g., they indicate the degree to which they, “Believe that others have good intentions” or “Suspect hidden motives in others”). This mitigates the possibility that a similarity in NEOPI trait space is similar to social cognitive trait spaces merely due to participants answering semantically related items similarly.

Protocol. Participants completed the 300-item NEOPI, used to measure the 30 facet subscales of the five-factor model[4]. Participants were first given instructions, “The following pages contain phrases describing people's behaviors. Please use the rating scale next to each phrase to describe how accurately each statement describes you. Describe yourself as you generally are now, not as you wish to be in the future. Describe yourself as you honestly see yourself, in relation to other people you know of the same sex as you are, and roughly your same age. So that you can describe yourself in an honest manner, your responses will be kept in absolute confidence. Please read each statement carefully, and then click the circle that corresponds to the accuracy of the statement", followed by general protocol instructions and informed consent. Following, participants completed the 300-items in randomized order (1 - "Very inaccurate" – 5 - "Very accurate" Likert scales), in five block sets of 60 items each.

Supplementary Results

Confirmation of a common trait space. A qualitative observation of prior work[7] is that trait space models are similarly structured across social perceptual domains. This is an important assumption of the present work, as if social perceptual trait spaces reflect the conceptual trait space, they should share structure across domains. We tested this assumption quantitatively, finding a high degree of similarity across all social cognitive trait space matrices (face – familiar person trait space matrices, Spearman $\rho(103) = 0.841$, $\rho^2(103) = 0.707$, $p < 0.0001$; 95% CI = [0.774 , 0.889]; face – social group trait space matrices, Spearman $\rho(103) = 0.794$, $\rho^2(103) = 0.631$, $p < 0.0001$; 95% CI = [0.711 , 0.856]; familiar person – social group trait space matrices, Spearman $\rho(103) = 0.824$, $\rho^2(103) = 0.679$, $p < 0.0001$; 95% CI = [0.751 , 0.877]). These results confirm the assumption that there is indeed a common trait space in social perception[7]. To our knowledge, these results also provide a first quantitative assessment of the commonality between social perceptual trait spaces.

The relationship of conceptual and social perceptual trait space while controlling for valence.

Study 1. Given the apparent clustering by valence of traits in the similarity matrices (Fig. 2), we also conducted these analyses as a multiple linear regression controlling for the valence similarity matrix (based on the absolute difference of valence ratings of each trait term; $n = 69$; Supplementary Fig. 1). We found social perceptual trait space matrices were each predicted significantly by the conceptual trait space matrix over and above the valence matrix (conceptual matrix predicts: face trait space matrix, $t(102) = 7.049$, $p < .0001$, $r^2 = .328$, 95% CI = [0.185, 0.330]; familiar person trait space matrix, $t(102) = 6.553$, $p < .0001$, $r^2 = .296$, 95% CI = [0.179, 0.334]; social group trait space matrix, $t(102) = 6.910$, $p < .0001$, $r^2 = .320$, 95% CI = [0.184, 0.333]). As we would expect, we also found that valence similarity significantly predicted these social perceptual trait matrices as well (and in person knowledge and stereotypes, the effect sizes were smaller than predictions from conceptual trait space similarity; valence matrix predicts: face trait space matrix, $t(102) = 7.086$, $p < .0001$, $r^2 = .330$, 95% CI = [0.116, 0.206]; familiar person trait space matrix, $t(102) = 4.199$, $p < .0001$, $r^2 = .147$, 95% CI = [0.054, 0.151]; social group trait space matrix, $t(102) = 5.383$, $p < .0001$, $r^2 = .221$, 95% CI = [0.079, 0.172]). Indeed, valence has long been noted to be a major factor in the organization of social perceptions (e.g., even used as an alternative labeling to the 'trustworthiness' dimension in the two-factor model of face impressions[8]). Our theoretical account is agnostic to the valenced nature of particular traits. A similar or dissimilar valence among two traits surely would play a role in driving traits' conceptual similarity, but in our view it is only one contributor. By demonstrating strong effects of conceptual trait space after controlling for the contribution of valence (and equal if not stronger effects of conceptual trait space than valence space), the results cast doubt on the possibility that purely affective associations are driving the observed effects.

Study 2. In Study 2, effects of conceptual matrices on social perceptual matrices remained significant when controlling for the valence matrix, taken from Study 1, in multiple regression (conceptual matrix predicts: face trait space matrix, $t(102) = 3.196$, $p = .002$, $r^2 =$

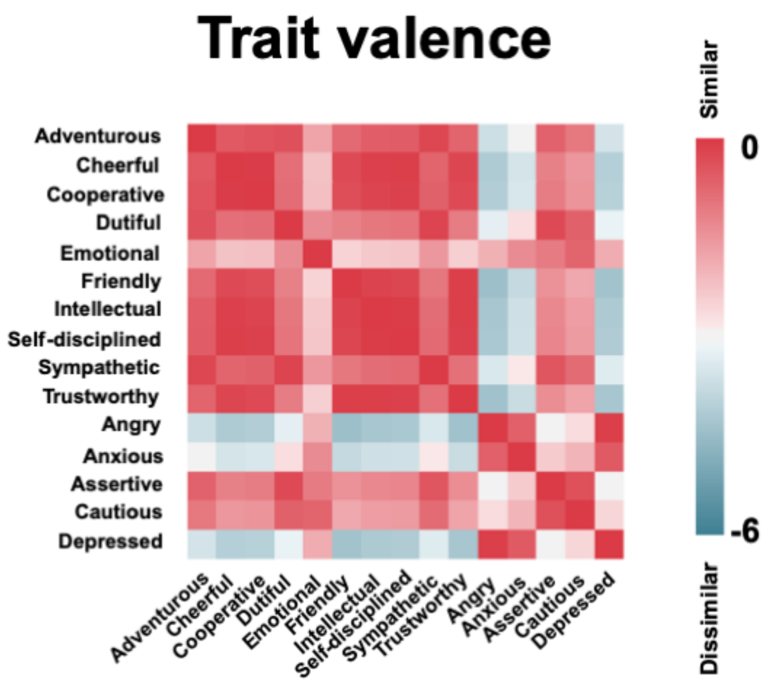
.091, 95% CI = [0.077, 0.328]; familiar person trait space matrix, $t(102) = 5.148$, $p < .0001$, $r^2 = .206$, 95% CI = [0.156, 0.352]; social group trait space matrix, $t(102) = 2.724$, $p = .008$., $r^2 = .068$, 95% CI = [0.036, 0.229]).

Study 3. In Study 3, we additionally performed the analysis controlling for the valence model collected in Study 1, finding subjective conceptual trait associations had a significant relation to face impressions over and above both group-average conceptual associations and their valenced structure ($b = 0.144$, $SE = .019$, $t(142.4) = 7.501$, $p < .0001$, 95% CI = [0.11, 0.18]).

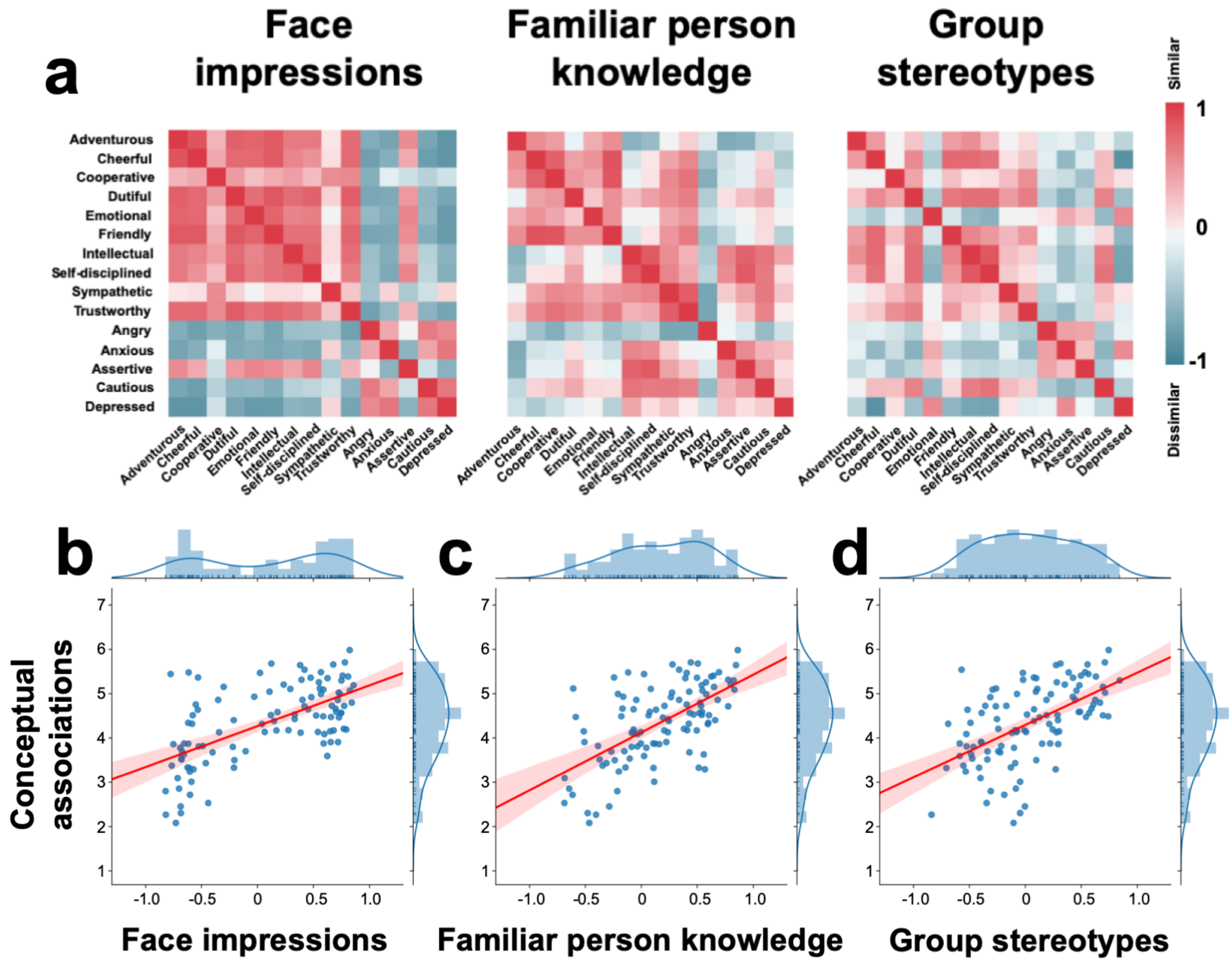
Social perceptions reflect actual personality structure. If conceptual trait space is applied to trait inferences (Studies 1-5), and perceivers' trait space mirrors actual personality structure as observed here, social perceptual trait spaces may also reflect the actual structure of personality. Indeed, we found this to be the case, as social perceptual trait space matrices from Study 1 were also strongly positively related to the NEOPI trait space matrix (NEOPI trait space matrix predicted the: face trait space matrix, Spearman $\rho(103) = 0.677$, $\rho^2(103) = 0.459$, $p < 0.0001$; 95% CI = [0.558 , 0.769]; familiar person trait space matrix, Spearman $\rho(103) = 0.644$, $\rho^2(103) = 0.415$, $p < 0.0001$; 95% CI = [0.516 , 0.744]; social group trait space matrix, Spearman $\rho(103) = 0.706$, $\rho^2(103) = 0.498$, $p < 0.0001$; 95% CI = [0.595 , 0.791]).

Additional analyses. In Studies 1, 2, and 7 we performed analyses testing the overlap between aggregate similarity matrices per each social cognitive model: conceptual, face, familiar person, and group matrices. There are two limitations of this statistical approach. In one, there is inherent dependency between elements of the matrix, as the same items and sometimes data are sometimes used to compute different elements (e.g., 'trustworthy' as input to both the 'trustworthy-angry' and 'trustworthy-anxious' cells). In another, random effects inference of any kind are not possible given matrices are aggregated across all subjects. Given entire conceptual matrices were collected per subjects in the conceptual task of Study 1, we performed additional analyses to address these limitations. These analyses were of course only possible where conceptual matrices are included in analysis. First, we conducted first-level analyses as RSA per conceptual task subject, predicting their unique conceptual matrices from the aggregate face, familiar person, and group matrices (as these were only able to be computed through aggregation across subjects in Studies 1 and 2). Second, we performed a group level analysis to test whether the similarity coefficients from first-level analyses were significant. Specifically, for each RSA reported, Spearman correlation coefficients were computed per subject, Fisher's z transformed, then submitted to a one-sample t-test against 0. Statistical test results are reported as performed on Fisher's z transformed Spearman correlation coefficients. For interpretation, descriptive statistics and confidence intervals of the original Spearman correlation coefficients are reported. All analyses reported were significant when tested with this method in Study 1 (conceptual – face RSA, mean $\rho = .525$, ρ SD = .202, $t(115) = 24.342$, $p < .0001$, $r^2 = .837$, mean ρ 95% CI = [0.488, 0.562]; conceptual – familiar person RSA, mean $\rho = .494$, ρ SD = .184, $t(115) = 25.760$, $p < .0001$, $r^2 = .852$, mean ρ 95% CI = [0.460, 0.528]; conceptual – group RSA, mean $\rho = .520$, ρ SD = .195, $t(115) = 25.180$, $p < .0001$, $r^2 = .846$, mean ρ 95% CI = [0.484, 0.556]), Study 2 (conceptual – face RSA, mean $\rho = .386$, ρ SD = .170, $t(115) = 22.661$, $p < .0001$, $r^2 = .817$, mean ρ 95% CI = [0.355, 0.417]; conceptual – familiar person RSA, mean $\rho = .398$, ρ SD = .143, $t(115) = 28.004$, $p < .0001$, $r^2 = .872$, mean ρ 95% CI = [0.372, 0.424]; conceptual – group RSA, mean $\rho = .378$, ρ SD = .162, $t(115) = 22.948$, $p < .0001$, $r^2 = .821$, mean ρ 95% CI = [0.348, 0.408]), and Study 7 (conceptual – NEOPI RSA, mean $\rho = .517$, ρ SD = .192, $t(115) = 26.203$, $p < .0001$, $r^2 = .857$, mean ρ 95% CI = [0.482, 0.552]).

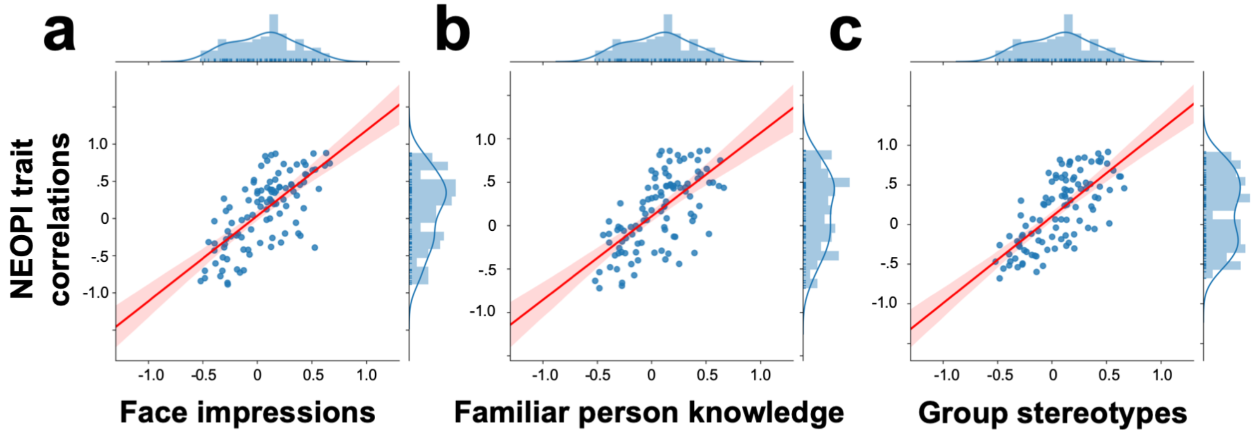
Supplementary Figures



Supplementary Figure 1. **Trait valence matrix**. In Study 1, we collected a trait space matrix of the absolute difference in valence ratings (dissimilar/blue to similar/red) of each trait adjective stimulus (*n* = 69). This was used as a control in Studies 1, 2, and 3 analyses to measure the redundancy of conceptual trait space and the valence similarity of trait terms, given the large contribution of valence to trait inferences and conceptual knowledge[8]. Control allowed analyses to measure the independent contribution of non-valence related conceptual similarities in trait-pairs to trait inferences.



Supplementary Figure 2. **Study 2 results**. First depicted are all social perceptual trait space similarity matrices from Study 2 (panel a), each made of the pairwise similarity values between each trait-pair. Each matrix is sorted by the k-means cluster solution of the conceptual trait space matrix, as to most intuitively depict their similar structure. Importantly, Study 2 used different descriptors for traits in each domain, for instance, while 'friendly' was used in the conceptual task, 'likely to agree with others' was used in face impressions. Study 2 used the same conceptual association data as Study 1 (see Figure 2, panel a). Second, we see evidence that conceptual trait space (n = 116; y-axis) substantially overlaps with social perceptual trait space across domains (x-axes; face impressions, panel b, n = 496, Spearman $\rho(103) = 0.575$, $\rho^2(103) = 0.331$, $p < 0.0001$; 95% CI = [0.431 , 0.691]; person knowledge, panel c, n = 478, Spearman $\rho(103) = 0.576$, $\rho^2(103) = 0.332$, $p < 0.0001$; 95% CI = [0.432 , 0.691]; and group stereotypes, n = 489, Spearman $\rho(103) = 0.574$, $\rho^2(103) = 0.329$, $p < 0.0001$; 95% CI = [0.430 , 0.690]). Error ribbons reflect standard error of effect estimates. While Pearson correlations are plotted for ease of interpretation, statistical analyses were of rank ordered data points. In each plot, trait space matrices (panel a) are flattened into their unique pair-wise similarity values and plotted against one another (conceptual on the y-axis, social perceptual matrices along the x-axes). Each data point is a trait-pair (e.g., 'friendly'-'self-disciplined'; 105 trait-pairs make up data points per panel). In each comparison, as two traits become more associated in conceptual knowledge, they become more correlated in trait inferences across domains.



Supplementary Figure 3. NEOPI trait space predicts social perceptual trait spaces. In Study 7, we find perceivers' conceptual trait associations (*y*-axes) are strikingly reflective of the actual correlation structure of personality traits (n = 307,313; *x*-axes; in face impressions, panel a, n = 484, Spearman $\rho(103) = 0.677$, $\rho^2(103) = 0.459$, $p < 0.0001$; 95% CI = [0.558 , 0.769]; person knowledge, panel b, n = 503, Spearman $\rho(103) = 0.644$, $\rho^2(103) = 0.415$, $p < 0.0001$; 95% CI = [0.516 , 0.744]; and group stereotypes, panel c, n = 488, Spearman $\rho(103) = 0.706$, $\rho^2(103) = 0.498$, $p < 0.0001$; 95% CI = [0.595 , 0.791]). Error ribbons display standard error around effect estimates, and there are 105 trait-pairs as data points per panel. While Pearson correlations are plotted for ease of interpretation, statistical analyses were of rank ordered data points. These results suggest a possibility that actual trait correlations are learned conceptually, and thereafter influence social perception. This does not necessarily entail accuracy in social perception ipso facto. This point is addressed in detail in the discussion.

Supplementary References

1 Kriegeskorte, N., Mur, M. & Bandettini, P. Representational similarity analysis - connecting the branches of systems neuroscience. *Frontiers in systems neuroscience* **2**, 4, doi:10.3389/neuro.06.004.2008 (2008).

2 Stolier, R. M., Hehman, E. & Freeman, J. B. A Dynamic Structure of Social Trait Space. *Trends in cognitive sciences* **22**, 197-200, doi:https://doi.org/10.1016/j.tics.2017.12.003 (2018).

3 Tskhay, K. O. & Rule, N. O. Accuracy in categorizing perceptually ambiguous groups: A review and meta-analysis. *Personality and Social Psychology Review* **17**, 72-86 (2013).

4 Johnson, J. A. Measuring thirty facets of the Five Factor Model with a 120-item public domain inventory: Development of the IPIP-NEO-120. *Journal of Research in Personality* **51**, 78-89 (2014).

5 Goldberg, L. R. A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. *Personality psychology in Europe* **7**, 7-28 (1999).

6 Schneider, D. J. Implicit personality theory: A review. *Psychological bulletin* **79**, 294 (1973).

7 Fiske, S. T., Cuddy, A. J. & Glick, P. Universal dimensions of social cognition: Warmth and competence. *Trends in cognitive sciences* **11**, 77-83 (2007).

8 Oosterhof, N. N. & Todorov, A. The functional basis of face evaluation. *Proceedings of the National Academy of Sciences* **105**, 11087-11092 (2008).