American Psychological Association

Journal of Experimental Psychology:
Human Perception and Performance

0096-1523/19/$12.00

2019, Vol. 45, No. 2, 224–236
http://dx.doi.org/10.1037/xhp0000601

Individual Differences in Perceptual Sensitivity and Representation of Facial Signals of Trustworthiness

Elisa Baccolo
Università degli Studi di Milano – Bicocca

Viola Macchi Cassia
Università degli Studi di Milano – Bicocca and
NeuroMI – Milan Center for Neuroscience, Milano, Italy

One of the most important sources of social information is the human face, on whose appearance we easily form social judgments: Adults tend to attribute a certain personality to a stranger based on minimal facial cues, and after a short exposure time. Previous studies shed light on the cognitive and neural mechanisms underlying the ability to discriminate facial properties conveying social signals, but the underlying processes supporting individual differences remain poorly understood. In the current study, we explored whether differences in sensitivity to facial cues to trustworthiness and in representing such cues in a multidimensional space are associated with individual variability in social attitude, as measured by the extraversion/introversion dimension. Participants performed a task where they assessed the similarity between faces that varied in the level of trustworthiness, and multidimensional scaling analyses were performed to describe perceptual similarity in a multidimensional representational space. Extraversion scores impacted RTs, but not accuracy or face representation, making less extraverted individuals slower in detecting similarity of faces based on physical cues to trustworthiness. These findings are discussed from an ontogenetic perspective, where reduced social motivation might constrain perceptual attunement to social cues from faces, without affecting the structuring of the face representational space.

Public Significance Statement

The ability to discriminate facial properties conveying social signals, such as trustworthiness, is a fundamental component of social interactions. Individual differences in this skill, however, remain almost unexplored, and are thus far from being understood. Our study revealed that individual variability in social attitudes is associated with differential sensitivity to perceptual differences among faces varying in the level of expressed trustworthiness. Results showed that more extraverted individuals are faster in assessing the similarity between faces based on trustworthiness cues. Nevertheless, differences on the extraversion dimension did not impact the representation of such cues in long-term memory. These results highlight the importance of individual differences related to one's social attitude in shaping social motivation and facial experience, which drive selective attention to faces and constrain perceptual attunement to social cues from faces across the life span.

Keywords: trustworthiness, perceptual sensitivity, face representational space, individual differences, extraversion

Faces are a fundamental vehicle of social information. As social animals, we are able to easily process and react to a wealth of social cues conveyed by a person's face, and use them to perform a number of evaluations about their mental states, emotions, intentions or even character. As a matter of fact, people are able to make personality judgments based on faces after a very short exposure time (e.g., Bar, Neta, & Linz, 2006; Todorov, Pakrashi, & Oosterhof, 2009; Willis & Todorov, 2006). Even relatively small differences between faces are easily perceived by the observer, leading to significantly different social judgments and evaluations (Jones, Kramer, & Ward, 2012). For example, facial cues with adaptive relevance, such as features of emotionally neutral faces that slightly resemble anger, happiness or other emotional states, are systematically used to infer behavioral intentions from other people (e.g., Fiske, Cuddy, & Glick, 2007; Todorov, Mende-Siedlecki, & Dotsch, 2013; Todorov, Said, Engell, & Oosterhof, 2008). In particular, one of the face-to-trait inferences that we typically perform when we first meet a stranger concerns their trustworthiness, that is the extent to which they could be

This article was published Online First December 27, 2018.

Elisa Baccolo, Department of Psychology, Università degli Studi di Milano-Bicocca; Viola Macchi Cassia, Department of Psychology, Università degli Studi di Milano-Bicocca, and NeuroMI – Milan Center for Neuroscience, Milano, Italy.

We thank Chiara Arossa for help in testing participants. This research was supported by a scholarship from the University of Milano-Bicocca to Elisa Baccolo and a grant from the same university to Viola Macchi Cassia.

Correspondence concerning this article should be addressed to Elisa Baccolo, Department of Psychology, Università degli Studi di Milano-Bicocca, piazza dell'Ateneo Nuovo 1, Milano 20154, Italy. E-mail: e.baccolo@campus.unimib.it

safely approached or better avoided (Adolphs, 2002; Willis & Todorov, 2006).

Although much research has been devoted to investigating the cognitive and neural mechanisms underlying the ability to make social trait inferences based on facial appearance in the average adult population (e.g., Ames, Fiske, & Todorov, 2011), still little is known about whether and how individuals differ in their sensitivity to the facial cues that drive those inferences and/or in their proneness to perform trait inferences from faces. There are many reasons to believe that individual differences related to personality and social behavior may be associated with differences in how social signals from faces are detected, represented and interpreted. Indeed, research suggests that neural responses to faces reflect social personality characteristics of the beholder, such as extraversion/introversion. Cheung and colleagues (Cheung, Rutherford, Mayes, & McPartland, 2010) reported a differential sensitivity to stimulus inversion at the level of the N170 event-related potential component for extraverts compared to introverts, which was marked by an enhanced N170 for inverted compared to upright faces in the former group, but not in the latter. Because the N170 inversion effect is a marker of perceptual expertise, its association with extraversion was interpreted by the Authors as suggesting that personality characteristics may affect social motivation, which would in turn affect the amount of perceptual experience that individuals acquire with faces through social interactions. Both these aspects would be diminished in typically developing individuals with high introversion scores, just like they are in individuals with clinical levels of social impairment (e.g., Dawson, Webb, & McPartland, 2005; Gepner, de Gelder, & Schonen, 1996; Grelotti, Gauthier, & Schultz, 2002; Klin et al., 1999). Likewise, Kirihara et al. (2012) reported an abnormal amplitude of the N170 for low-extravert schizophrenics in response to emotional faces.

Recent research suggests that individual differences in social personality characteristics are not only associated with the amount of neural effort involved in processing faces, but also to neural sensitivity to social cues conveyed by faces. Individuals with high anxiety traits retain a more fine-grained representation of untrustworthy faces in visual working memory compared to individuals with low anxiety traits (Meconi, Luria, & Sessa, 2014). These results were interpreted as oversusceptibility toward threats related to the high-anxious profile, which would lead to improved working memory processing of untrustworthy faces that are likely perceived as threatening. Accordingly, there is evidence that self-protection motives (i.e., mental states that tunes perceptual and cognitive processes to threatening information in the environment) positively impact performance accuracy in detecting facial cues associated with judgments of trustworthiness (Young, Slepian, & Sacco, 2015). Further support to the hypothesis of a link between an individual's social attitude and their disposition to derive social judgments from faces is provided by individuals with neurogenetic disorders. Individuals with autism spectrum disorders show atypical face-based judgments of trustworthiness in association with impaired social relation and communication abilities (Adolphs, Sears, & Piven, 2001; Forgeot d'Arc et al., 2016). Similarly, individuals with Williams Syndrome show an atypical positive bias in social judgments of unknown people in association with hypersocial behavior (Bellugi, Adolphs, Cassady, & Chiles, 1999).

If only a few studies investigated the association between individual variability in personality dimensions and sensitivity to physical cues of social traits, no studies at all have explored whether such variability is also associated with differences in the way social traits from faces are represented in long-term memory. According to the most influential model of how facial experience is stored in memory, our visual system extracts information from the social environment to build a representational model that maps the perceived properties of the faces we encounter into a multidimensional face–space that accounts for stimulus variation (Valentine, 1991). This mapping continues throughout the life span, and is constrained by the amount and quality of experience one person gains with faces (e.g., Gao, Maurer, & Nishimura, 2010; Humphreys & Johnson, 2007; Rhodes & Jeffery, 2006), so that the more experience we acquire with specific face types, the more fine-grained their representation will be in our perceptual space (see review by Valentine, Lewis, & Hills, 2016).

Although there is evidence that individual variability in face–space characteristics is related to individual differences in face recognition skills (e.g., Dennett, McKone, Edwards, & Susilo, 2012), to the best of our knowledge, no studies have explored whether individual differences in personality dimensions that are central to an individual's social drive and motivation are reflected in corresponding variations of the face–space organization. To fill this gap in the literature, in the current study, we explored whether differences among individuals in perceptual sensitivity and mental representation of facial features related to social judgments along the trustworthiness–untrustworthiness continuum are associated with individual differences in extraversion–introversion levels. We focused on the extraversion–introversion dimension, as it represents the personality attribute that best explains an individual's social motivation. Indeed, individuals characterized by an introvert profile are socially inhibited, avoid social situations and find time spent alone more rewarding than time spent with others. On the contrary, individuals who score high on extraversion actively seek social engagements and find time spent with others as more rewarding than time spent alone. Moreover, available evidence suggests that extraversion–introversion scores measured through self-report questionnaires modulate neural correlates of face processing (Cheung et al., 2010; Fink, 2005).

To explore individual differences in perceptual sensitivity to facial cues to trustworthiness, we used as stimulus material the trustworthiness continuum selected from the inventory developed by Oosterhof and Todorov (2008), composed of computer-generated faces slightly varying on seven levels of perceived trustworthiness. To obtain measures of perceptual sensitivity and perceived dissimilarity, we tested a group of typically developing adult participants in a perceptual similarity task in which they were required to provide similarity judgments by selecting among two probe faces the one they perceived as more similar to a simultaneously presented target face. Participants' accuracy and RTs in selecting the correct probe face, which is nearer to the position of the target face along the trustworthiness continuum, were measured as indexes of perceptual sensitivity to physical cues to trustworthiness from faces. Pairwise dissimilarity scores, obtained for each participant from their specific pattern of probe choices, were used to perform a multidimensional scaling (MDS) analysis, which provided a measure of how physical cues to trustworthiness are represented in memory. MDS allows to visualize how, on average, faces are represented and clustered by visualizing their data-driven similarity as points in a two-dimensional space (Robert, 2007; Shepard, 1980). This method has been used in prior studies investigating the organization of face representation in adults and children (e.g., Nishimura, Maurer, & Gao, 2009) to provide a measure of how participants represent similarities between perceived faces (Edelman, 1998).

We hypothesized that, if individual differences in social attitude and motivation are able to shape our sensitivity to social signals from faces, more extraverted people would perform better and/or faster at matching faces based on physical cues to trustworthiness. Moreover, if individual differences along the introversion/extraversion dimension impact the representation of trustworthy and untrustworthy faces, we would expect to observe a more fine-grained representation in more extraverted people compared to introverts.

To test for the stability and consistency of the similarity judgments obtained from the perceptual similarity task, participants also completed a multiarrangement dissimilarity task (Kriegeskorte & Mur, 2012), which provided us with a second, additional measure of the participants' representation of trustworthiness cues in long-term memory. In this task, pairwise dissimilarities scores were inferred from participants' active grouping, on multiple subsequent trials, of two-dimensional arrangements of the whole set and multiple subsets of faces from the trustworthiness continuum. The item set context in which dissimilarities are judged thus varies across trials, yielding a deeper reflection of the participants' mental representation. The finding of similarities between the MDS solutions of participants' dissimilarity scores obtained through the perceptual similarity task and the multiarrangement dissimilarity task would provide evidence for the robustness and cross-task stability of participants' judgments about the similarity of faces varying along the trustworthiness continuum.

Finally, explicit subjective judgments of facial trustworthiness (trustworthy vs. untrustworthy) and emotional expression (happy vs. angry) were obtained from each subject to explore the presence of an association with extraversion levels. As faces that vary on trustworthiness are typically perceived as emotionally expressive (Caulfield, Ewing, Burton, Avard, & Rhodes, 2014; Todorov et al., 2008), we expected to observe similar results for the two explicit evaluations.

Materials and Method

Participants

The sample size was based on previous studies showing how individual differences in social personality traits modulate face processing skills (Cheung et al., 2010; Meconi et al., 2014). Additionally, a power analysis for a multiple regression model with two predictors (target trustworthiness and extraversion score) revealed that about 34 participants would be required to have an 80% chance to observe a significant effect with an alpha level of .05 and a medium effect size. Our final sample included 43 young adult participants (33 females; M_{age} = 24.93 years; range = 22–32). All subjects had normal or corrected-to-normal visual ability and were without history of any psychiatric or neurological disorders. An additional 3 participants were tested but excluded from the final sample as they were identified as RT outliers by using both the interquartile method (upper quartile + 3 × interquartile range) and the standard deviation (±2 *SD* from the mean). Informed written consent was obtained from all participants before testing. The protocol was carried out in accordance with the ethical standards of the Declaration of Helsinki and approved by the Ethics Committee of Università degli Studi di Milano-Bicocca.

Stimuli

Prior to the experiment, an independent sample of 25 young adults (19 females; M_{age} = 24.96 years; range = 19–35) provided nine-step ratings of facial trustworthiness (trustworthy vs. untrustworthy) for 11 computer-generated emotionally neutral male identities taken from the Todorov database (Oosterhof & Todorov, 2008). These extensively validated (Todorov, Dotsch, Porter, Oosterhof, & Falvello, 2013) face stimuli were created using FaceGen Modeler 3.2 (Singular Inversions, Toronto, Ontario, Canada, www.facegen.com) based on data-driven, computational models of trustworthiness judgments. For each face identity, seven variations were generated along the trustworthiness dimension to reflect a continuum ranging from −3 *SD* (lowest level of trustworthiness) to +3 *SD* (highest level of trustworthiness), with the neutral version of each identity located at 0 *SD*. By filling in a digital questionnaire via SurveyMonkey (SurveyMonkey Inc., San Mateo, California, USA, www.surveymonkey.com), participants rated the −3 *SD* and the +3 *SD* versions of the 11 selected identities on a scale ranging from 1 (*I wouldn't trust this person at all*) to 9 (*I would definitely trust this person*). This procedure allowed us to select the identity that yielded the highest (*M* = 6.41, *SD* = 1.41) and lowest (*M* = 2.91, *SD* = 1.42) trustworthy ratings (i.e., Identity fi_002 of the Todorov database; Figure 1).

Apparatus and Procedure

Participants were tested individually in a quiet, dedicated room. They all completed the perceptual similarity task followed by the multiarrangement dissimilarity task while seated 60 cm from a 17.3-in. touchscreen monitor with a resolution of 1080p, onto which the stimuli were presented in color. Stimulus presentation and response collection were controlled by a MATLAB script interfaced with Mousetracker (Freeman & Ambady, 2010). At the beginning of the testing session, participants completed a Qualtrics-delivered (Qualtrics, Provo, Utah, USA, https://www.qualtrics.com) Italian version of the Big Five Questionnaire (BFQ; Caprara, Barbaranelli, Borgogni, & Perugini, 1993); at the end of the testing session, they filled in two questionnaires aimed to obtain explicit subjective judgments of facial trustworthiness and emotional expres-

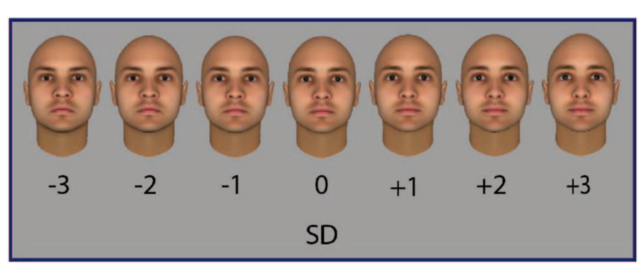

Figure 1. The seven variations of the computer-generated face identity (i.e., fi_002; Oosterhof & Todorov, 2008) used in the study. See the online article for the color version of this figure.

sion, which were administered in a counterbalanced order. The datasets generated during the current study are available in the Open Science Framework (OSF) repository, https://osf.io/645ky/?view_only=3f90ec0e66ee4837a784e259fd2ed63b. (Baccolo & Macchi Cassia, 2018).

BFQ

The Italian version of the BFQ (Caprara et al., 1993) is a 44-item self-report questionnaire designed to measure the Big Five dimensions of personality: extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience. The BFQ has high internal consistency, temporal stability, convergent and discriminant validity. Participants responded using a 5-point Likert-type scale. Only the 24 items contributing to the Extraversion scale were scored, and internal consistency of this scale was high (α = .916). The obtained extraversion scores were standardized by converting them into z scores.

Trustworthiness and Happiness Ratings

Participants provided nine-step ratings of facial trustworthiness (trustworthy vs. untrustworthy) and emotional expression (happy vs. angry) for each of the seven faces by filling in two separate Qualtrics-delivered questionnaires, presented in a counterbalanced order. Participants' judgments could range from 1 (*I wouldn't trust this person at all*/*This person does not look happy at all*) to 9 (*I would definitely trust this person*/*This person looks very happy*).

Perceptual Similarity Task

In the perceptual similarity task participants were told that one target face would appear on the screen and that they would be asked to recognize which face from two probes appearing after the initial presentation was more similar to the target face. Participants controlled the start of each trial by pressing a START button that appeared centrally at the bottom of the screen, and was replaced upon participants' pressure by a target face. After 1,000 ms, two probes appeared on the right and left side of the upper portion of the screen. Participants were instructed to keep the cursor on the START position until the probes appeared and to decide which of the two probes looked more like the target by moving the cursor toward the chosen probe. Participants were required to respond as fast as possible. The probes remained on the screen until a response was made, and an intertrial interval of 500 ms elapsed before the START button reappeared on the screen. If a response was not made by 3,000 ms after the probes' appearance, a pop-up message appeared on the screen that warned the participant to respond faster in subsequent trials. The target and the two probes were always different variations on the trustworthiness continuum (see Figure 2). Each of the seven variations was presented as target 15 times, resulting in a total of 105 trials. The position (left/right) at which each probe appeared on the screen was randomized across participants. In nine trials (i.e., catch trials) the two probes were equally distant from the target along the trustworthiness continuum; on all the remaining, experimental, trials one probe was closer to target than the other. Catch trials were introduced to obtain a full-factorial design were each variation was compared the same number of times with all the others; they were included in the analyses on the dissimilarity scores but excluded from the analyses on response accuracy and RTs. Responses on experimental trials were coded as correct when participants selected the probe that was closer to the target along the trustworthiness continuum. Participants performed a few trials with an example face identity to assure they understood the task fully.

Multiarrangement Dissimilarity Task

In the multiarrangement dissimilarity task (adapted from Kriegeskorte & Mur, 2012), participants were instructed to

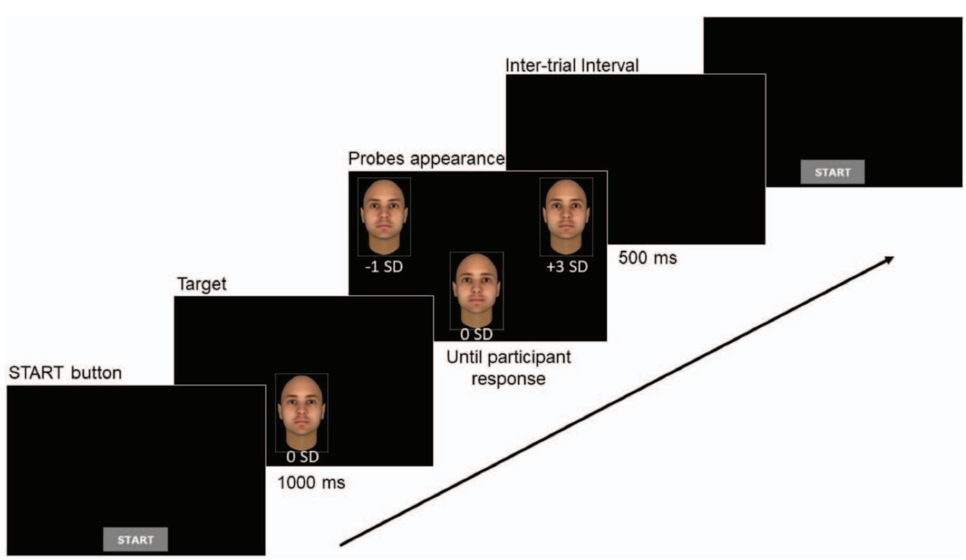

Figure 2. Example of an experimental trial from the perceptual similarity task, in which the probe presented on the left side was closer to the target face than the one presented on the right. Standard deviations values for each face denoting the position of the face along the trustworthiness continuum are superimposed for clarity. See the online article for the color version of this figure.

arrange the whole set and multiple subsets of the face stimuli presented as icons on a computer screen according to their similarity, by means of mouse drag-and-drop operations. Indeed, the multiarrangement method works via an optimization process (i.e., lift-the-weakest algorithm for adaptive design of item subsets), whose aim is to obtain the same amount of information on pairwise dissimilarity for each pair of items in a set by keeping record of what has already been specified by the subject during the arrangement procedure, and by building subsequent item subsets to validate those dissimilarities for which the weakest evidence has been obtained so far. On each trial, the faces were initially presented in random order at regular angular intervals around a circular arena. Participants were instructed to arrange each face within the arena by dragging and dropping it with the mouse, using the entire arena to express the dissimilarity between the faces; they were told that the position of each face in the arena indicated its similarity relationship with every other face. At the beginning of the first trial, the icons of all the seven faces appeared at the border of the arena; during the following trials, different subsets of the seven faces, selected by the lift-the-weakest algorithm for adaptive design of item subsets based on weakest evidence of dissimilarity, appeared at the borders of the arena: the participant's task remained that of arranging the faces within the arena to reflect their perceptual similarity (see Kriegeskorte & Mur, 2012, for a detailed description of the algorithm; Figure 3). Acquisition was terminated after 5 min from the start have elapsed, and not before the completion of at least the first trial. Participants completed an average of 4.5 trials (range = 1–11).

Results

Perceptual Similarity Task

Response accuracy. To examine whether participants' accuracy in selecting which probe was more similar to the target face was modulated by the position of the target on the trustworthiness continuum and the participant's extraversion score, we conducted a repeated-measures analysis of covariance (ANCOVA) with target trustworthiness (−3 *SD*, −2 *SD*, −1 *SD*, 0 *SD*, +1 *SD*, +2 *SD*, +3 *SD*) as the within-subjects factor and extraversion score entered as covariate. The ANCOVA on mean accuracy showed a main effect of target, $F(6, 246) = 5.345$, $p < .001$, $\eta^2 = 0.115$. Post hoc analysis (Bonferroni corrected) revealed that participants performed more accurately on −3 *SD* target trials than on 0 *SD*, $p = .002$, and +2 *SD* target trials, $p = .020$, on −2 *SD* target trials than on 0 *SD* trials, $p = .005$, and on +3 *SD* target trials compared to 0 *SD* trials, $p = .019$ (see Figure 4). A test of within-subjects contrasts revealed a significant quadratic trend, $F(1, 41) = 17.370$, $p < .001$, $\eta^2 = 0.982$.

Response times. A repeated-measures ANCOVA with target trustworthiness (−3 *SD*, −2 *SD*, −1 *SD*, 0 *SD*, +1 *SD*, +2 *SD*, +3

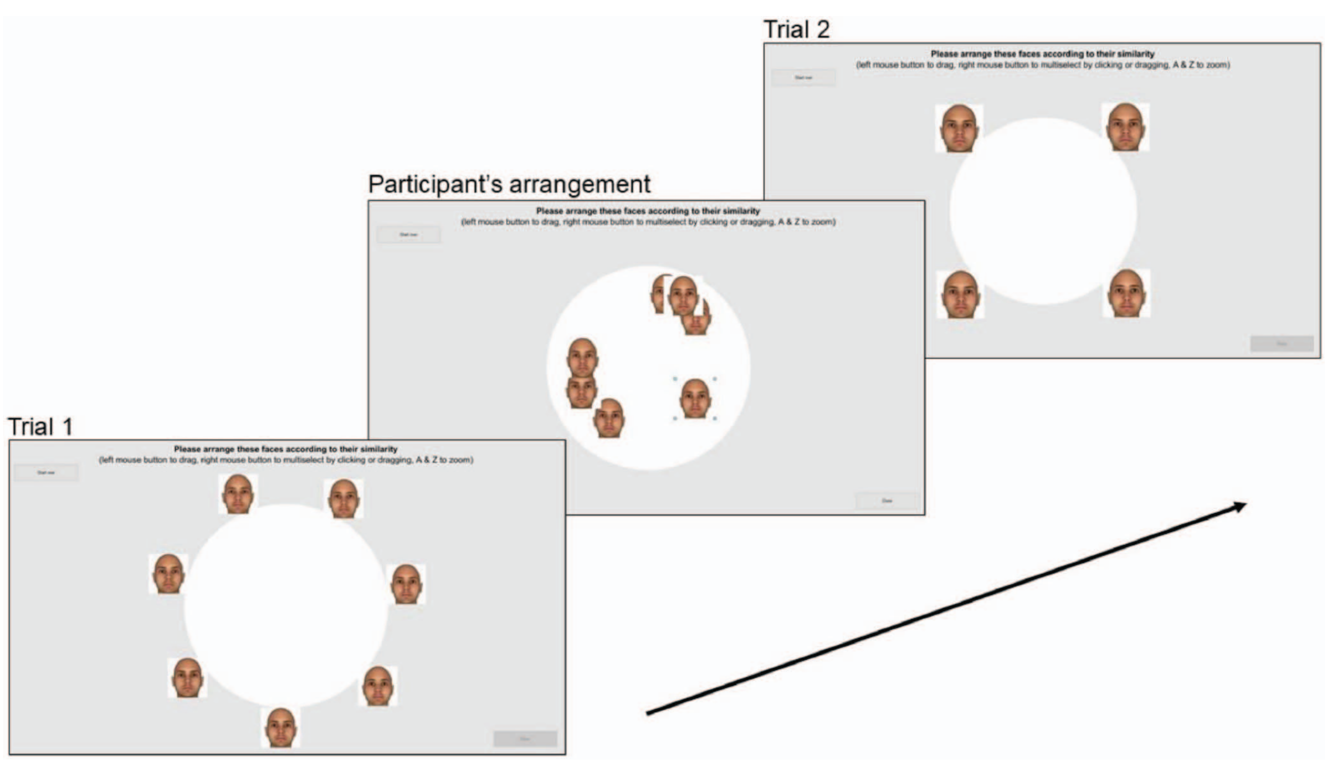

Figure 3. An example of three subsequent two-dimensional arrangements of stimulus sets from the multiarrangement dissimilarity task, corresponding to (a) the starting point for Trial 1, (b) the stimulus arrangement generated by the participant, and (c) the starting point for Trial 2. Subsets of faces presented in trials following the first one were selected by the lift-the-weakest algorithm for adaptive design of item subsets (Kriegeskorte & Mur, 2012) based on weakest evidence of dissimilarity. See the online article for the color version of this figure.

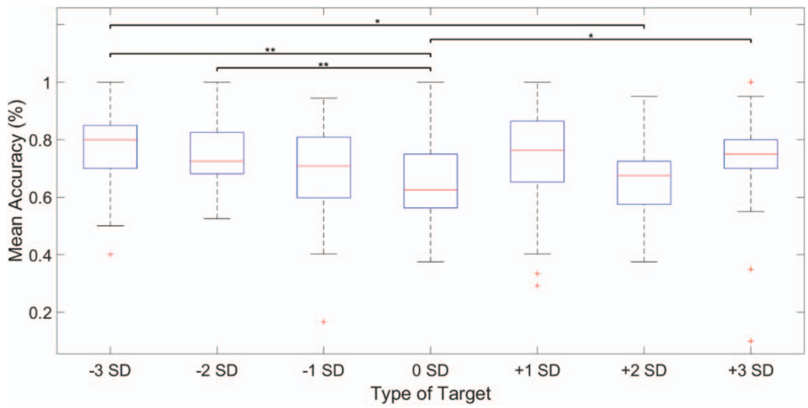

Figure 4. Boxplot of mean accuracy rates plotted as a function of the position of the target face on the trustworthiness continuum. Red crosses represent outliers. * $p < .05$. ** $p < .01$. See the online article for the color version of this figure.

SD) as the within-subjects factor and extraversion score as covariate, performed on mean correct response times (RTs), revealed a main effect of extraversion score, $F(1, 41) = 5.039$, $p = .030$, $\eta^2 = 0.109$, which we followed up by means of correlational analysis. Such analysis revealed a negative association between participants' RTs and their extraversion score, $r = -0.33$, $p = .031$, 95% confidence interval (CI) [−0.574, −0.034]: The higher their level of extraversion, the faster they were in selecting the probe that was more similar to the target face, irrespective of the intensity of the target's expressed trustworthiness (see Figure 5).

Dissimilarity scores. Pairwise dissimilarity scores were derived for each participant from correct and incorrect responses on all trials, by attributing a distance score of 0 (minimum dissimilarity) to the face pair composed of the target face and the selected probe, and a distance score of 1 (maximum dissimilarity) to the face pair composed of the target face and the nonselected probe. The sum of the scores obtained for each face pair was scaled to 0–1 by dividing it by the number of pairwise judgments provided ($N = 10$). Dissimilarity scores for each subject were used to derive a 7 × 7 representational dissimilarity matrix (RDM) illustrating the level of perceived dissimilarity between face pairs. Each column and row represents the dissimilarity judgment for one trustworthiness intensity with respect to every other trustworthiness intensities, with the diagonal representing the extent of dissimilarity within the same trustworthiness intensity, which will therefore be a diagonal of zeros, and the upper part of the matrix being specular to the lower part (Kriegeskorte, Mur, & Bandettini, 2008). The obtained RDMs were averaged across subjects, resulting in a single RDM representing the average perceived dissimilarity between face pairs differing in trustworthiness intensity (see Figure 6).

Using the pairwise dissimilarity scores averaged across participants, a MDS analysis was performed using a MATLAB script adapted from Kriegeskorte and Mur (2012) to represent the per-

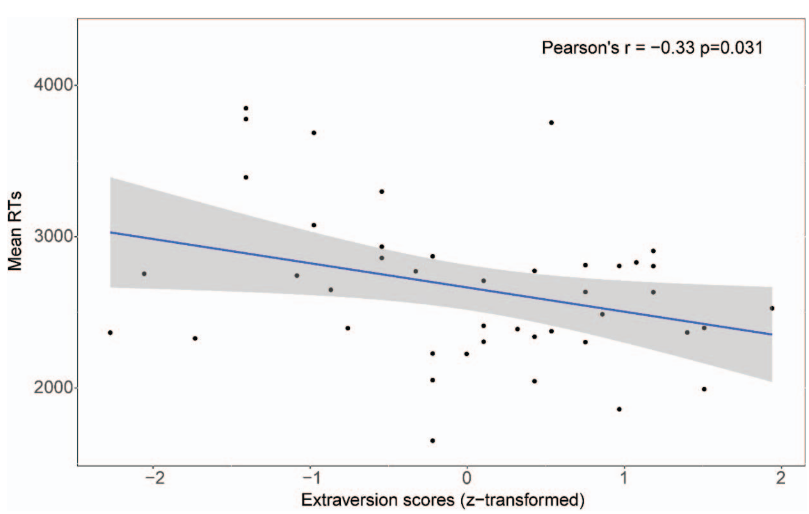

Figure 5. The significant negative relation between participants' mean response times (RTs) for correct responses in the perceptual similarity task and their level of extraversion, $r = -0.33$, $p = .031$. See the online article for the color version of this figure.

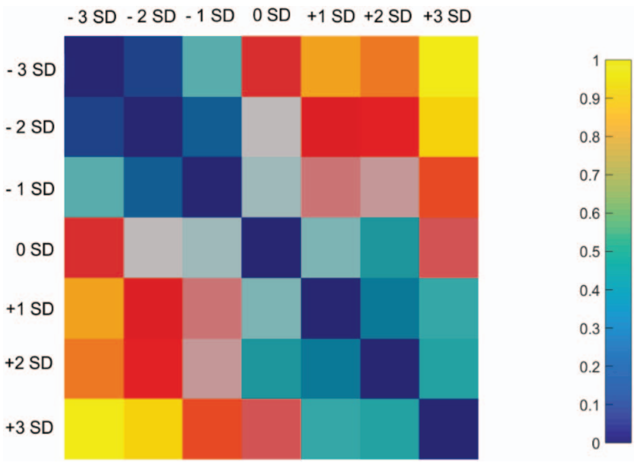

Figure 6. Average representational dissimilarity matrix derived from pairwise dissimilarity scores obtained from participants' responses in the perceptual similarity task (yellow [light gray] = maximum dissimilarity, blue [dark gray] = minimum dissimilarity). See the online article for the color version of this figure.

ceived similarity of the seven faces in a multidimensional space. The goodness-of-fit value for the MDS solution, measured by Kruskal stress formula (Kruskal & Wish, 1978), was 0.087. The scatterplot shown in Figure 7 represents the distribution of the dissimilarity scores against the fitted distances of the MDS solution. The dissimilarity scores and the fitted distances of MDS were highly correlated, Spearman's $r = .98$, $p < .001$, 95% CI [0.979, 0.993], which provides a further qualitative diagnostic for the MDS. To determine how the faces clustered together in the MDS solution, we performed a cluster analysis (Sireci & Geisinger, 1992) using the dendrogram MATLAB function (see Figure 8).

To test whether participants consistently agreed in attributing dissimilarity judgments for specific items, we computed a repeated-measure ANOVA on cosine distances between participants' pairwise dissimilarity scores for each trustworthiness intensity. Cosine distance is defined as one minus the cosine of the angle between two vectors of an inner product space: Two vectors that have the same orientation have a cosine distance of 0, while two orthogonal vectors have a cosine distance of 1, thus cosine distance ranges between 0 and 1. In a RDM, each row represents the dissimilarity scores obtained for one trustworthiness intensity with respect to every other trustworthiness intensities: We computed the cosine distance between each row of the RDM of each single participant and the corresponding row of the RDM of every other participant. A repeated-measures ANOVA on cosine distances with trustworthiness intensity as the within-subjects factor proved significant, $F(6, 5412) = 125.735$, $p < .001$, $\eta^2 = 0.122$, showing that dissimilarity scores were more consistent across participants for extremely (±3 SD) and moderately (±2 SD) trustworthy or untrustworthy faces than for all the others, $ps < .05$ (see Figure 9).

To determine whether participants' representation of physical cues to trustworthiness varied as a function of their level of extraversion, we computed the cosine distance between the whole RDMs (i.e., dissimilarity judgments between all face pairs) for all participant pairs and the Euclidean distance between the extraversion scores for all participant pairs. We hypothesized that, if the way participants represent faces along the trustworthiness dimension varies as a function of their extraversion scores, a greater difference in extraversion scores should be associated with a greater difference in the RDMs. On the contrary, a correlational analysis revealed a significant, though very weak, negative correlation (performed on a total of 903 pairs, which represent all possible pairwise combinations of the 43 participants), $r = -0.08$, $p = .009$, 95% CI [−0.151, −0.021], between the Euclidean distances of participants' extraversion scores and the cosine distances of the same participants' RDMs. As a matter of fact, as apparent from Figure 10, participants showed high agreement in attributing dissimilarity judgments independently of their extraversion score, as 95% of observations remained below a cosine distance of 0.177, with mean and standard deviation equal, respectively, to 0.067 and 0.045.

Multiarrangement Dissimilarity Task

Pairwise dissimilarity scores estimated from participants' grouping of the seven faces and the multiple face subsets selected by the lift-the-weakest algorithm for adaptive design of item subsets (Kriegeskorte & Mur, 2012) in the multiarrangement task were used to generate an average RDM similar to that obtained from the dissimilarity scores derived from the perceptual similarity

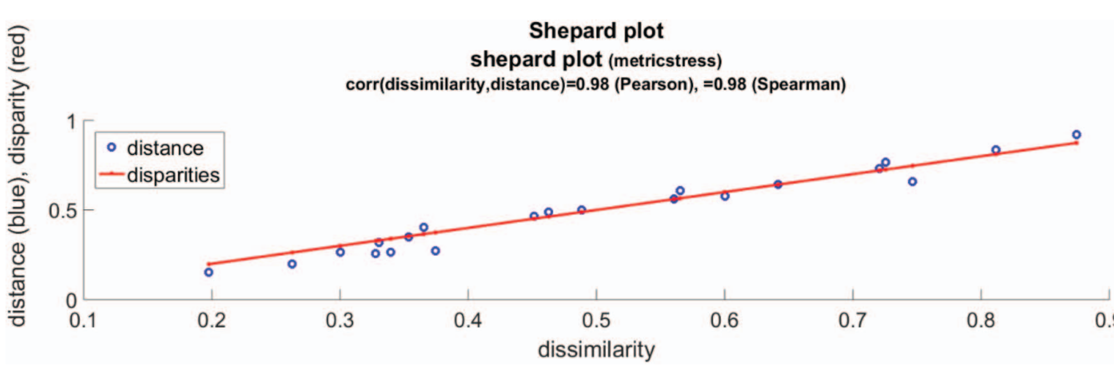

Figure 7. Scatterplot of the average dissimilarities scores obtained from participants' responses in the perceptual similarity task against the fitted distances of the multidimensional scaling solution. See the online article for the color version of this figure.

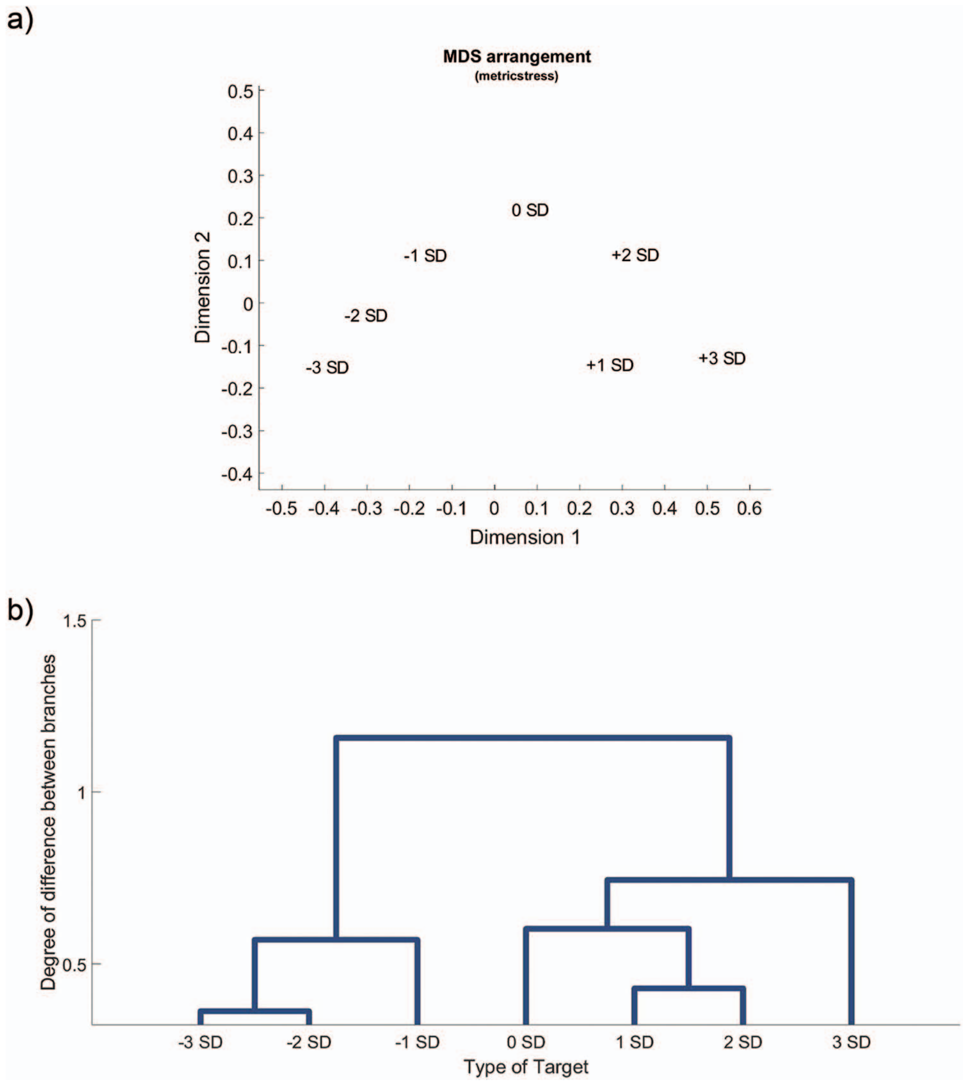

Figure 8. Visualization of (a) the two-dimensional arrangement of the faces in the multidimensional scaling (MDS) solution of the average dissimilarity scores obtained from participants' responses in the perceptual similarity task and (b) the corresponding hierarchical plot describing the results of the cluster analysis. See the online article for the color version of this figure.

task. Correlational analysis revealed a significant positive association between the two matrices, $r = .83$, $p < .001$, 95% CI [0.628, 0.930].

A MDS analysis was performed, and the goodness-of-fit value for the MDS solution, measured by Kruskal stress formula, was very similar (stress = 0.07) to that of the MDS solution from the perceptual similarity task (stress = 0.087), indicating that the similarity judgments collected through the two tasks are highly comparable, and that participants attributed similarity judgments to faces expressing different degrees of trustworthiness with stability and consistency, irrespective of the task context.

Trustworthiness and Happiness Ratings

The level of interrater agreement for the trustworthiness questionnaire, computed using Krippendorff's alpha (Krippendorff, 2004), was $\alpha = .472$. To determine how participants' explicit judgments of perceived trustworthiness varied as a function of the faces position along the trustworthiness continuum, we performed a repeated-measures ANCOVA with trustworthiness intensity (−3 *SD*, −2 *SD*, −1 *SD*, 0 *SD*, +1 *SD*, +2 *SD*, +3 *SD*) as the within-subjects factor and extraversion score entered as covariate. The analysis revealed a main effect of trustworthiness intensity, $F(6, 264) = 71.157$, $p < .001$, $\eta^2 = 0.634$. A test of within-subjects contrasts revealed a significant linear trend, $F(1, 41) = 166.056$, $p < .001$, $\eta^2 = 0.802$, showing that participants explicitly judged trustworthiness intensity from each of the seven face stimuli they were presented with as a function of the stimulus position along the trustworthiness continuum.

The level of interrater agreement for the happiness questionnaire was $\alpha = .692$. The repeated-measures ANCOVA with trustworthiness intensity as the within-subjects factor and extraversion

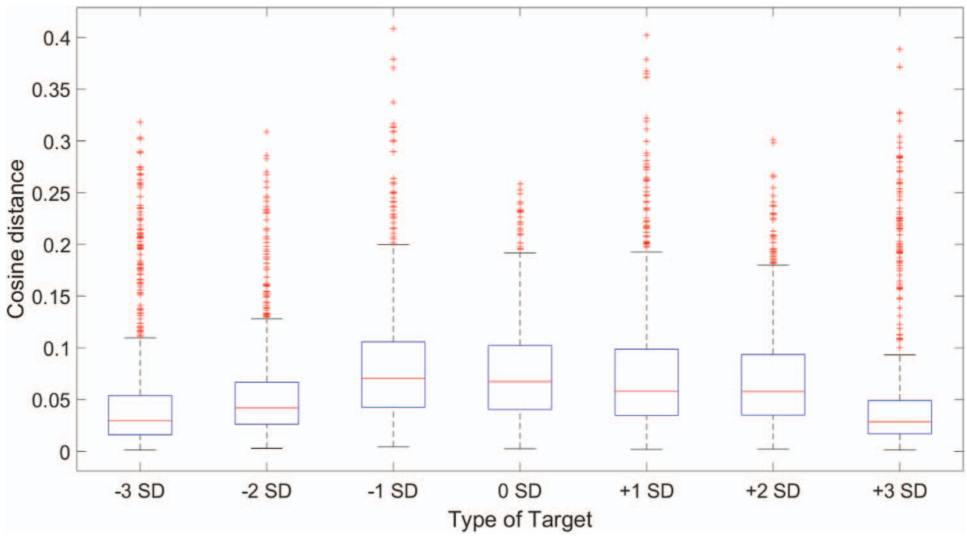

Figure 9. Boxplot showing the cosine distances between participants' pairwise dissimilarity scores for each trustworthiness intensity. Crosses represent outliers. See the online article for the color version of this figure.

score entered as covariate showed a main effect of trustworthiness intensity, $F(6, 264) = 199.571$, $p < .001$, $\eta^2 = 0.819$, with all post hoc comparisons (Bonferroni corrected) reaching statistical significance, *p*s $< .05$. A test of within-subjects contrasts revealed a significant linear trend, $F(1, 41) = 522.098$, $p < .001$, $\eta^2 = 0.922$.

To explore the relationship between explicit judgments of perceived trustworthiness and explicit judgments of perceived happiness of the seven faces, we ran a correlational analysis between the two judgments scores, which proved significant, $r = .99$, $p < .001$, 95% CI [0.918, 0.998]. Perceived trustworthiness was positively related to perceived happiness: participants' judgments of happiness increased by 1.17683 points for each increase in judgments of trustworthiness. Moreover, a paired-sample *t* test on participants' judgments of perceived trustworthiness and happiness revealed that happiness in the faces was judged as overall less intense than trustworthiness, $t(6) = 3.963$, $p = .007$.

Discussion

The present study aimed to explore whether individual differences in social attitudes are associated with differences in the way social cues from faces are discriminated and represented. Analyses of participants' RTs in the perceptual similarity task revealed that individuals who scored higher on extraversion on a self-report questionnaire were faster at successfully discriminating narrow differences in the level of trustworthiness expressed by the faces. This finding extends previous evidence that social attitudes, as measured through introversion-extraversion scores, are reflected in

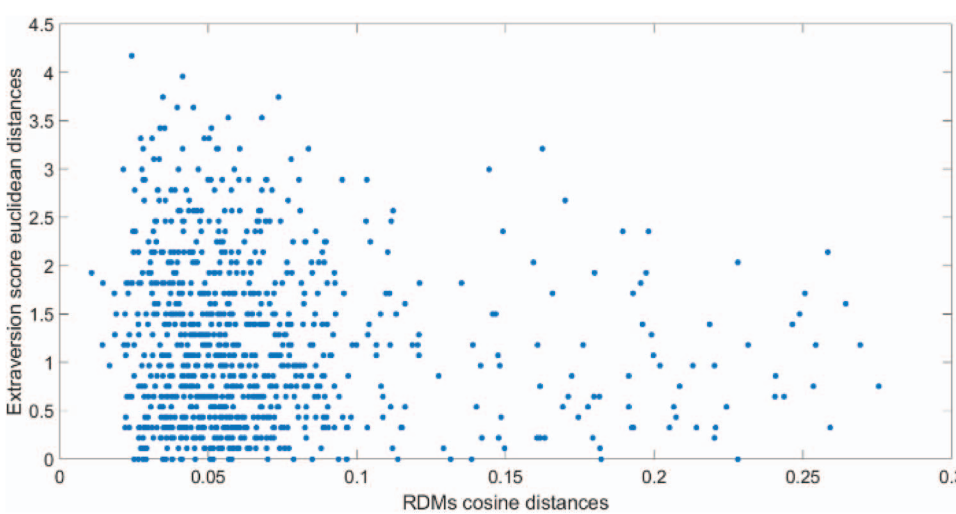

Figure 10. Scatterplot of the cosine distances between the representational dissimilarity matrices (RDMs) of participant pairs plotted as a function of Euclidean distances between participants' extraversion scores. See the online article for the color version of this figure.

electroencephalographic (EEG) neural activation during emotional face processing (Fink, 2005) or in the amplitude of EEG markers of perceptual expertise (Cheung et al., 2010). Indeed, in the current study introversion-extraversion scores modulated a specific aspect of face processing, that is the ability to promptly detect the social signals that faces convey, and to discriminate subtle differences in the intensity of such signals.

The association between social personality attributes and neural face processing abilities has been interpreted as a secondary effect of primary variations in social motivation, which would constrain the development of face processing expertise across the life span (Cheung et al., 2010). A similar interpretation may be extended to the current findings: If normal, subclinical, variations in social motivation associated with personality attributes constrain the amount of social contact and consequent facial experience accumulated by an individual, our findings of a negative relation between extraversion scores and RTs to detect subtle differences in the intensity of facial cues to trustworthiness may reflect greater perceptual expertise in extraverts than in introverts.

Indeed, recent models have emphasized the role of an individual's social (and cultural) experience in shaping their spontaneous tendency to infer other people's traits from facial attributes (Over & Cook, 2018). Developmental research suggests that early in development infants are extremely sensitive to social cues hidden in other people's appearance or actions, and modulate their own behavior based on such cues (Mascaro & Csibra, 2014; Van de Vondervoort & Hamlin, 2018). Although research on face-traits inferences has focused mainly on adults, there is some evidence that even preverbal infants can detect facial signs of trustworthiness (Jessen & Grossmann, 2016, 2017), and by the age of 3, children can use such information to make explicit judgments about how "nice" or "mean" a face appears to be (Cogsdill, Todorov, Spelke, & Banaji, 2014). This suggests that, although mature personality trait concepts may emerge later in childhood, sensitivity to those facial cues on which specific trait inferences are mapped appears early in development, and is continuously refined throughout the life span as a result of each individual's own experience within their social environment. Indeed, infant research has shown that two fundamental dimensions of infant temperament that are strongly linked to the introversion–extraversion personality continuum, namely approach to novelty and fearfulness, modulate infants' neural processing of emotional faces, as a result of their impact on the way infants interact with their social environment (e.g., de Haan, Belsky, Reid, Volein, & Johnson, 2004; Rajhans, Missana, Krol, & Grossmann, 2015; Taylor-Colls & Pasco Fearon, 2015). Following from this evidence, our finding of faster responding to facial cues to trustworthiness in extraverts relative to introverted adults may represent a developmental outcome of differential ontogenetic experience with the social environment.

It is worth noting that, unlike RTs, response accuracy in the perceptual similarity task was not modulated by participants' personality characteristics. Rather, participants' accuracy varied as a function of the trustworthiness intensity of the target face following a quadratic trend: Participants exhibited higher accuracy when the target face belonged to one of the continuum extremes (-3 *SD*, $+3$ *SD*), independently of the valence, compared to the neutral (0 *SD*) face. This shows that the intensity of the facial cues that are relevant to trustworthiness judgments affects our perceptual discrimination abilities, and suggests that faces including more intense social cues enjoy a processing advantage over those including less intense cues. This finding is congruent with the widely reported attentional and processing advantage of angry (e.g., LoBue, 2009) or fearful/threatening (e.g., Holmes, Green, & Vuilleumier, 2005) faces over neutral ones. Unlike these earlier findings, though, the processing advantage in our data was not restricted to faces with negative valence, as it extended to extremely trustworthy, as well as untrustworthy, faces. Indeed, similar evidence of valence-independent sensitivity to trustworthiness cues has been reported in infants, who showed neural discrimination between neutral faces and both very trustworthy ($+3$ *SD*) and very untrustworthy (-3 *SD*) faces, but not between trustworthy and untrustworthy faces (Jessen & Grossmann, 2016). In adults, the amygdala shows stronger activation in response to both positively valenced and negatively valenced faces at the extremes of the trustworthiness continuum than to faces near the center of the continuum, following a quadratic response pattern similar to that observed for accuracy rates in the current study (Said, Baron, & Todorov, 2009; Said, Dotsch, & Todorov, 2010). Although we did not collect typicality judgments in the current study, previous studies using the same artificial faces that we used showed that neutral faces are perceived as more prototypical than faces at both extremes of the trustworthiness continuum, and that deviations in typicality explain the amygdala response better than valence (Said et al., 2010).

In addition to individual differences in perceptual sensitivity to facial signs of trustworthiness, the current study also aimed to examine, for the first time, how we represent perceived similarity of faces that vary in expressed trustworthiness in a multidimensional space, and whether social attitude has an impact on how this representation is organized. We obtained no evidence that participants' extraversion scores modulated how trustworthiness cues from faces are represented in long-term memory. Indeed, both the MDS and the cluster analysis revealed that our subjects, as a group, represented faces as a function of the intensity of the social signal they express, forming a cluster composed of untrustworthy faces (-3, -2, -1 *SD*) and another cluster composed of trustworthy faces ($+1$, $+2$, $+3$ *SD*), to which the neutral face (0 *SD*) is associated. Indeed, neutral faces were perceived as being more similar to the trustworthy branch of the continuum. This finding resonates well with those showing that neutral faces are often associated with emotional states (e.g., Adams, Nelson, Soto, Hess, & Kleck, 2012; Lee, Kang, Park, Kim, & An, 2008), and suggests that, when it comes to faces, there is not such a thing as absolute neutrality.

As a further indication that neutral faces engender ambiguous social cues, it is worth nothing that the 0 *SD* face, together with the -1 *SD* and $+1$ *SD*, is the one for which participants showed the lowest agreement in attributing dissimilarity judgments. As a matter of fact, variance followed a parabolic trend, reaching its maximum around the central hub of the trustworthiness continuum, that is the same part of the continuum for which participants' accuracy steadily decreased. On the contrary, participants mostly agreed on attributing dissimilarity judgments for faces at the extreme ends ($+3$ and -3 *SD*), further proving that social facial cues of greater intensity are more easily detected and processed than less intense cues, independently of their valence (i.e., whether the face is very trustworthy or very untrustworthy).

An important aspect of our results is that the same personality characteristics that modulated sensitivity to physical cues to trustworthiness from faces did not affect explicit judgments of perceived trustworthiness from the same faces. Participants judgments of perceived trustworthiness steadily increased from the negative extreme end (−3 *SD*) of the trustworthiness dimension to the positive extreme end (+3 *SD*), in accord to the continuum conceived by Oosterhof and Todorov (Oosterhof & Todorov, 2008). However, it is worth noting that trustworthiness judgments did not vary among each single step of the continuum, as participants did not attribute significantly different judgments for −3 *SD* and −2 *SD*, −2 *SD* and −1 *SD*, +1 *SD* and +2 *SD*, and +2 *SD* and +3 *SD*. This suggests that even the slightest physical cues from faces are sufficient to generate a robust evaluation of a person's trustworthiness, which is thus independent from the intensity of the facial cue and the social attitude of the beholder.

Explicit judgments of perceived happiness followed a pattern very similar to that observed for trustworthiness judgments. Indeed, judgments on the two dimensions were strongly correlated, confirming earlier demonstrations of a robust association between perceived trustworthiness and the attribution of emotional states (Adams, Ambady, Neil Macrae, & Kleck, 2006; Oosterhof & Todorov, 2008; Todorov et al., 2008), and providing further support to the emotion overgeneralization hypothesis (Said, Sebe, & Todorov, 2009; Zebrowitz, Fellous, Mignault, & Andreoletti, 2003), according to which spontaneous trait inferences are overgeneralized responses to facial configurations resembling emotional expressions. However, although strongly correlated, happiness judgments were significantly lower than trustworthiness judgments, suggesting that, albeit being an important cue to trustworthiness, emotional valence does not fully explain the variance of trustworthiness evaluation. Indeed, other elements, such as face typicality and/or attractiveness (Said et al., 2010), might contribute as well to forming our impression of others (e.g., Jones et al., 2012; Sofer, Dotsch, Wigboldus, & Todorov, 2015).

The current study has few important limitations. First of all, because a large number of trials was required to obtain dissimilarity judgments from the perceptual similarity task, we only used one face identity, thus limiting the generalizability of the obtained results. Second, we used computer-generated faces, which we obviously do not encounter in our everyday social interactions. The use of more natural and ecological stimuli might elicit different patterns of performance. Third, the absence of individual differences in response accuracy for the perceptual similarity task could be due to the task not being challenging enough. Indeed, including a response deadline, which was absent in the current task procedure, would have increased task difficulty by pushing participants to provide faster responses. Nevertheless, because the goal of the task was to obtain dissimilarity judgments from all possible pairwise comparisons between trustworthiness variations, we needed to acquire a response on each trial, and avoid null responses at any trial. Still, participants were encouraged to respond as fast as possible through task instructions and warning messages that prompted them to respond faster if their RT on any single trial slowed down to 3,000 ms or more.

To sum up, this study provides novel evidence for the existence of a significant relationship between individual differences in social motivation and perceptual sensitivity to facial properties conveying social signals. More extraverted individuals proved to be faster in successfully assessing similarity between faces that slightly varied in the level of expressed trustworthiness. Nevertheless, not performance accuracy nor the long-term representation of trustworthiness facial cues were impacted by participants' extraversion scores. Overall, these findings add to those reported by the only other existing study on individual differences in perceptual sensitivity to facial cues associated with trustworthiness (Young et al., 2015), in indicating that motivational states deriving from one's own personality characteristics (current study) or fundamental motives (i.e., self-protective motives; Young et al., 2015) are associated with sensitivity to such cues, ultimately affecting social perception. In line with social–cognitive theories of face recognition biases (see review by Hugenberg, Young, Bernstein, & Sacco, 2010), we hypothesize that perceivers' social motivation and perceptual experience in discriminating among individual faces may work together to drive selective attention during face encoding, thereby affecting discrimination of social cues from faces.

References

Adams, R. B., Ambady, N., Neil Macrae, C., & Kleck, R. (2006). Emotional expressions forecast approach-avoidance behavior. *Motivation and Emotion, 30,* 177–186. http://dx.doi.org/10.1007/s11031-006-9020-2

Adams, R. B., Jr., Nelson, A. J., Soto, J. A., Hess, U., & Kleck, R. E. (2012). Emotion in the neutral face: A mechanism for impression formation? *Cognition and Emotion, 26,* 431–441. http://dx.doi.org/10.1080/02699931.2012.666502

Adolphs, R. (2002). Recognizing emotion from facial expressions: Psychological and neurological mechanisms. *Behavioral and Cognitive Neuroscience Reviews, 1,* 21–62. http://dx.doi.org/10.1177/1534582302001001003

Adolphs, R., Sears, L., & Piven, J. (2001). Abnormal processing of social information from faces in autism. *Journal of Cognitive Neuroscience, 13,* 232–240. http://dx.doi.org/10.1162/089892901564289

Ames, D. L., Fiske, S. T., & Todorov, A. T. (2011). Impression formation: A focus on others' intents. In J. Decety & J. Cacioppo (Eds.), *The Oxford handbook of social neuroscience* (pp. 419–433). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/oxfordhb/9780195342161.013.0028

Baccolo, E., & Macchi Cassia, V. (2018). *Individual differences in perceptual sensitivity and representation of facial signals of trustworthiness*. Retrieved from https://osf.io/645ky/?view_only=3f90ec0e66ee4837a784e259fd2ed63b

Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion, 6,* 269–278. http://dx.doi.org/10.1037/1528-3542.6.2.269

Bellugi, U., Adolphs, R., Cassady, C., & Chiles, M. (1999). Towards the neural basis for hypersociability in a genetic syndrome. *Neuroreport, 10,* 1653–1657. http://dx.doi.org/10.1097/00001756-199906030-00006

Caprara, G. V., Barbaranelli, C., Borgogni, L., & Perugini, M. (1993). The "big five questionnaire": A new questionnaire to assess the five factor model. *Personality and Individual Differences, 15,* 281–288. http://dx.doi.org/10.1016/0191-8869(93)90218-R

Caulfield, F., Ewing, L., Burton, N., Avard, E., & Rhodes, G. (2014). Facial trustworthiness judgments in children with ASD are modulated by happy and angry emotional cues. *PLoS ONE, 9,* e97644. http://dx.doi.org/10.1371/journal.pone.0097644

Cheung, C. H. M., Rutherford, H. J. V., Mayes, L. C., & McPartland, J. C. (2010). Neural responses to faces reflect social personality traits. *Social Neuroscience, 5,* 351–359. http://dx.doi.org/10.1080/17470911003597377

Cogsdill, E. J., Todorov, A. T., Spelke, E. S., & Banaji, M. R. (2014). Inferring character from faces: A developmental study. *Psychological Science, 25,* 1132–1139. http://dx.doi.org/10.1177/0956797614523297

Dawson, G., Webb, S. J., & McPartland, J. (2005). Understanding the nature of face processing impairment in autism: Insights from behavioral and electrophysiological studies. *Developmental Neuropsychology, 27,* 403–424. http://dx.doi.org/10.1207/s15326942dn2703_6

de Haan, M., Belsky, J., Reid, V., Volein, A., & Johnson, M. H. (2004). Maternal personality and infants' neural and visual responsivity to facial expressions of emotion. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 45,* 1209–1218. http://dx.doi.org/10.1111/j.1469-7610.2004.00320.x

Dennett, H. W., McKone, E., Edwards, M., & Susilo, T. (2012). Face aftereffects predict individual differences in face recognition ability. *Psychological Science, 23,* 1279–1287. http://dx.doi.org/10.1177/0956797612446350

Edelman, S. (1998). Representation is representation of similarities. *Behavioral and Brain Sciences, 21,* 449–467. http://dx.doi.org/10.1017/S0140525X98001253

Fink, A. (2005). Event-related desynchronization in the EEG during emotional and cognitive information processing: Differential effects of extraversion. *Biological Psychology, 70,* 152–160. http://dx.doi.org/10.1016/j.biopsycho.2005.01.013

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

Forgeot d'Arc, B., Ramus, F., Lefebvre, A., Brottier, D., Zalla, T., Moukawane, S., . . . Delorme, R. (2016). Atypical social judgment and sensitivity to perceptual cues in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 46,* 1574–1581. http://dx.doi.org/10.1007/s10803-014-2208-5

Freeman, J. B., & Ambady, N. (2010). MouseTracker: Software for studying real-time mental processing using a computer mouse-tracking method. *Behavior Research Methods, 42,* 226–241. http://dx.doi.org/10.3758/BRM.42.1.226

Gao, X., Maurer, D., & Nishimura, M. (2010). Similarities and differences in the perceptual structure of facial expressions of children and adults. *Journal of Experimental Child Psychology, 105,* 98–115. http://dx.doi.org/10.1016/j.jecp.2009.09.001

Gepner, B., de Gelder, B., & Schonen, S. (1996). Face processing in autistics: Evidence for a generalised deficit? *Child Neuropsychology, 2,* 123–139. http://dx.doi.org/10.1080/09297049608401357

Grelotti, D. J., Gauthier, I., & Schultz, R. T. (2002). Social interest and the development of cortical face specialization: What autism teaches us about face processing. *Developmental Psychobiology, 40,* 213–225. http://dx.doi.org/10.1002/dev.10028

Holmes, A., Green, S., & Vuilleumier, P. (2005). The involvement of distinct visual channels in rapid attention towards fearful facial expressions. *Cognition and Emotion, 19,* 899–922. http://dx.doi.org/10.1080/02699930441000454

Hugenberg, K., Young, S. G., Bernstein, M. J., & Sacco, D. F. (2010). The categorization-individuation model: An integrative account of the other-race recognition deficit. *Psychological Review, 117,* 1168–1187. http://dx.doi.org/10.1037/a0020463

Humphreys, K., & Johnson, M. H. (2007). The development of "face-space" in infancy. *Visual Cognition, 15,* 578–598. http://dx.doi.org/10.1080/13506280600943518

Jessen, S., & Grossmann, T. (2016). Neural and behavioral evidence for infants' sensitivity to the trustworthiness of faces. *Journal of Cognitive Neuroscience, 28,* 1728–1736. http://dx.doi.org/10.1162/jocn_a_00999

Jessen, S., & Grossmann, T. (2017). Neural evidence for the subliminal processing of facial trustworthiness in infancy. *Neuropsychologia*. Advance online publication. http://dx.doi.org/10.1016/j.neuropsychologia.2017.04.025

Jones, A. L., Kramer, R. S. S., & Ward, R. (2012). Signals of personality and health: The contributions of facial shape, skin texture, and viewing angle. *Journal of Experimental Psychology: Human Perception and Performance, 38,* 1353–1361. http://dx.doi.org/10.1037/a0027078

Kirihara, K., Kasai, K., Tada, M., Nagai, T., Kawakubo, Y., Yamasaki, S., . . . Araki, T. (2012). Neurophysiological impairment in emotional face processing is associated with low extraversion in schizophrenia. *Progress in Neuro-Psychopharmacology & Biological Psychiatry, 37,* 270–275. http://dx.doi.org/10.1016/j.pnpbp.2012.02.012

Klin, A., Sparrow, S. S., de Bildt, A., Cicchetti, D. V., Cohen, D. J., & Volkmar, F. R. (1999). A normed study of face recognition in autism and related disorders. *Journal of Autism and Developmental Disorders, 29,* 499–508. http://dx.doi.org/10.1023/A:1022299920240

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

Kriegeskorte, N., Mur, M., & Bandettini, P. (2008). Representational similarity analysis: Connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience, 2,* 4.

Krippendorff, K. (2004). Reliability in content analysis some common misconceptions and recommendations. *Human Communication Research, 30,* 411. http://dx.doi.org/10.1111/j.1468-2958.2004.tb00738.x

Kruskal, J. B., & Wish, M. (1978). *Multidimensional scaling*. Beverly Hills, CA: SAGE. http://dx.doi.org/10.4135/9781412985130

Lee, E., Kang, J. I., Park, I. H., Kim, J.-J., & An, S. K. (2008). Is a neutral face really evaluated as being emotionally neutral? *Psychiatry Research, 157,* 77–85. http://dx.doi.org/10.1016/j.psychres.2007.02.005

LoBue, V. (2009). More than just another face in the crowd: Superior detection of threatening facial expressions in children and adults. *Developmental Science, 12,* 305–313. http://dx.doi.org/10.1111/j.1467-7687.2008.00767.x

Mascaro, O., & Csibra, G. (2014). Human infants' learning of social structures: The case of dominance hierarchy. *Psychological Science, 25,* 250–255. http://dx.doi.org/10.1177/0956797613500509

Meconi, F., Luria, R., & Sessa, P. (2014). Individual differences in anxiety predict neural measures of visual working memory for untrustworthy faces. *Social Cognitive and Affective Neuroscience, 9,* 1872–1879. http://dx.doi.org/10.1093/scan/nst189

Nishimura, M., Maurer, D., & Gao, X. (2009). Exploring children's face-space: A multidimensional scaling analysis of the mental representation of facial identity. *Journal of Experimental Child Psychology, 103,* 355–375. http://dx.doi.org/10.1016/j.jecp.2009.02.005

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

Over, H., & Cook, R. (2018). Where do spontaneous first impressions of faces come from? *Cognition, 170,* 190–200. http://dx.doi.org/10.1016/j.cognition.2017.10.002

Rajhans, P., Missana, M., Krol, K. M., & Grossmann, T. (2015). The association of temperament and maternal empathy with individual differences in infants' neural responses to emotional body expressions. *Development and Psychopathology, 27,* 1205–1216. http://dx.doi.org/10.1017/S0954579415000772

Rhodes, G., & Jeffery, L. (2006). Adaptive norm-based coding of facial identity. *Vision Research, 46,* 2977–2987. http://dx.doi.org/10.1016/j.visres.2006.03.002

Robert, H. (2007). Warren S. Torgerson, Theory and methods of scaling. *Behavioral Science, 4,* 245–247.

Said, C. P., Baron, S. G., & Todorov, A. (2009). Nonlinear amygdala response to face trustworthiness: Contributions of high and low spatial frequency information. *Journal of Cognitive Neuroscience, 21,* 519–528. http://dx.doi.org/10.1162/jocn.2009.21041

Said, C. P., Dotsch, R., & Todorov, A. (2010). The amygdala and FFA track both social and non-social face dimensions. *Neuropsychologia, 48,* 3596–3605. http://dx.doi.org/10.1016/j.neuropsychologia.2010.08.009

Said, C. P., Sebe, N., & Todorov, A. (2009). Structural resemblance to emotional expressions predicts evaluation of emotionally neutral faces. *Emotion, 9,* 260–264. http://dx.doi.org/10.1037/a0014681

Shepard, R. N. (1980). Multidimensional scaling, tree-fitting, and clustering. *Science, 210,* 390–398. http://dx.doi.org/10.1126/science.210.4468.390

Sireci, S., & Geisinger, K. (1992). Analyzing test content using cluster analysis and multidimensional scaling. *Applied Psychological Measurement, 16,* 17–31. http://dx.doi.org/10.1177/014662169201600102

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

Taylor-Colls, S., & Pasco Fearon, R. M. (2015). The effects of parental behavior on infants' neural processing of emotion expressions. *Child Development, 86,* 877–888. http://dx.doi.org/10.1111/cdev.12348

Todorov, A., Dotsch, R., Porter, J. M., Oosterhof, N. N., & Falvello, V. B. (2013). Validation of data-driven computational models of social perception of faces. *Emotion, 13,* 724–738. http://dx.doi.org/10.1037/a0032335

Todorov, A., Mende-Siedlecki, P., & Dotsch, R. (2013). Social judgments from faces. *Current Opinion in Neurobiology, 23,* 373–380. http://dx.doi.org/10.1016/j.conb.2012.12.010

Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating faces on trustworthiness after minimal time exposure. *Social Cognition, 27,* 813–833. http://dx.doi.org/10.1521/soco.2009.27.6.813

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. http://dx.doi.org/10.1016/j.tics.2008.10.001

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 43,* 161–204. http://dx.doi.org/10.1080/14640749108400966

Valentine, T., Lewis, M. B., & Hills, P. J. (2016). Face–space: A unifying concept in face recognition research. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology, 69,* 1996–2019. http://dx.doi.org/10.1080/17470218.2014.990392

Van de Vondervoort, J. W., & Hamlin, J. K. (2018). The early emergence of sociomoral evaluation: Infants prefer prosocial others. *Current Opinion in Psychology, 20,* 77–81. http://dx.doi.org/10.1016/j.copsyc.2017.08.014

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

Young, S. G., Slepian, M. L., & Sacco, D. F. (2015). Sensitivity to perceived facial trustworthiness is increased by activating self-protection motives. *Social Psychological & Personality Science, 6,* 607–613. http://dx.doi.org/10.1177/1948550615573329

Zebrowitz, L. A., Fellous, J.-M., Mignault, A., & Andreoletti, C. (2003). Trait impressions as overgeneralized responses to adaptively significant facial qualities: Evidence from connectionist modeling. *Personality and Social Psychology Review, 7,* 194–215. http://dx.doi.org/10.1207/S15327957PSPR0703_01

Received May 7, 2018
Revision received August 3, 2018
Accepted September 18, 2018 ■