Research Report

Psychological Science
2014, Vol. 25(5) 1132–1139

DOI: 10.1177/0956797614523297
pss.sagepub.com

Inferring Character From Faces: A Developmental Study

Emily J. Cogsdill[1], Alexander T. Todorov[2], Elizabeth S. Spelke[1], and Mahzarin R. Banaji[1]
[1]Harvard University and [2]Princeton University

Abstract

Human adults attribute character traits to faces readily and with high consensus. In two experiments investigating the development of face-to-trait inference, adults and children ages 3 through 10 attributed trustworthiness, dominance, and competence to pairs of faces. In Experiment 1, the attributions of 3- to 4-year-olds converged with those of adults, and 5- to 6-year-olds' attributions were at adult levels of consistency. Children ages 3 and above consistently attributed the basic mean/nice evaluation not only to faces varying in trustworthiness (Experiment 1) but also to faces varying in dominance and competence (Experiment 2). This research suggests that the predisposition to judge others using scant facial information appears in adultlike forms early in childhood and does not require prolonged social experience.

Keywords

cognitive development, social perception, social cognition, physical appearance, open data, open materials

Received 6/11/13; Revision accepted 1/15/14

Faces command attention and interest, and facial appearance has profound effects on social judgments (Todorov, Mende-Siedlecki, & Dotsch, 2013; Zebrowitz & Montepare, 2008). The speed and confidence with which observers dispatch character assessments such as "trustworthy" or "competent" in response to a face is impressive. Face-to-trait inference appears to be intuitive and automatic among human adults; its development in early childhood is the focus here.

Several aspects of face-to-trait inferences are noteworthy. First, prior research shows that they occur extremely rapidly, within 50 ms after exposure to a face (Todorov, Pakrashi, & Oosterhof, 2009). Second, these character attributions show broad and cross-cultural consensus (Rule et al., 2010). Third, they often result from overgeneralizing perceptions of facial configurations that signal ecologically valid information, such as emotional states (Said, Sebe, & Todorov, 2009) and fitness (Zebrowitz & Rhodes, 2004). Finally, face-to-trait inferences occur even in consequential settings, including criminal courtroom proceedings (Blair, Judd, & Chapleau, 2004), businesses (Rule & Ambady, 2008), and political elections (Todorov, Mandisodza, Goren, & Hall, 2005).

Here, we report an initial exploration testing whether young children infer character traits such as trustworthiness, competence, and dominance in response to two-dimensional static images of faces, and if so, how early in development they do so in an adultlike manner.[1] If adult-child agreement in face-to-trait inferences emerges gradually across development, one might infer that these inferences require prolonged social experience to reach an adultlike state. If instead young children's inferences are like those of adults, this would indicate that face-to-trait character inferences are a fundamental social cognitive capacity that emerges early in life. Thus, we simply investigated whether children and adults make similar trait inferences based on the same faces.

Research has shown that infants prefer to look at faces over nonfaces and form face preferences on the basis of attractiveness, gender, and race (Bar-Haim, Ziv, Lamy, & Hodes, 2006; Langlois et al., 2000; Quinn, Yahr, Kuhn,

Corresponding Author:
Mahzarin R. Banaji, 1520 William James Hall, 33 Kirkland St., Cambridge, MA 02138
E-mail: mahzarin_banaji@harvard.edu

Slater, & Pascalis, 2002; Ramsey, Langlois, Hoss, Rubenstein, & Griffin, 2004; Simion, Macchi Cassia, Turati, & Valenza, 2001). However, little is known about how older children use faces to make inferences about other people's character, and the existing research on this topic is mixed. Even though 3- to 4-year-olds predict behavior from information about mental states (Wellman, Cross, & Watson, 2001), children under age 7 usually fail simple behavior-to-behavior prediction tasks, which have been used to study the existence of early trait attributions (Rholes & Ruble, 1984), and are less likely than older children to use trait words to describe people (Barenboim, 1981).

Although some aspects of face-to-trait inference in children have been studied (see Antonakis & Dalgas, 2009; Clément, Bernard, Grandjean, & Sander, 2013; Keating & Bai, 1986; Montepare & Zebrowitz-McArthur, 1989), the present research explored the development of face-to-trait inferences within a wide age group and in a variety of domains. Our method enabled us to test and compare responses among not only adults but also 3- to 10-year-old children. In two experiments, we explored face-based attributions of basic evaluations (mean vs. nice), as well as of more specific traits (strength and intelligence).

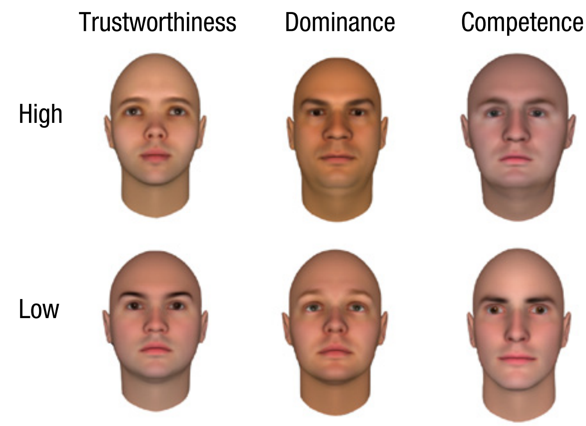

Fig. 1. Sample stimuli from the three stimulus sets used in Experiment 1. Each stimulus set included three faces that are perceived as high (+3 *SD* from the average face) and three faces that are perceived as low (−3 *SD* from the average face) on a given trait (trustworthiness, dominance, or competence).

General Method

Participants viewed computer-generated faces selected to be high or low on perceived trustworthiness, dominance, or competence. These extensively validated (Todorov, Dotsch, Porter, Oosterhof, & Falvello, 2013) faces were created in FaceGen Modeller 3.2 (Singular Inversions, www.facegen.com) and based on data-driven, computational models (derived from adults' judgments) of the respective traits (Oosterhof & Todorov, 2008; Todorov & Oosterhof, 2011). In our experiments, we used three sets of faces, each of which included six distinct face identities. Each set contained three faces that are perceived as high (3 *SD* above the average face) and three faces that are perceived as low (3 *SD* below the average face) on a given trait (trustworthiness, dominance, or competence; see Fig. 1).

In each trial, participants viewed two faces presented side by side; one of the faces was high and the other low in a given trait. In Experiment 1, face pairs appeared in three blocks (order counterbalanced across participants), with each block showing faces from a different set (i.e., from a different trait model). Each block contained nine trials in which all possible face pairs in the given set (i.e., each of the three high-trait faces paired with each of the three low-trait faces) appeared in a random order. Participants in Experiment 2 viewed only a single block of trials (i.e., all possible face pairs for a single set of faces).

On each trial, participants identified which of the two faces possessed a particular trait by answering a short question (e.g., "Which of these people is very nice"). Children were tested in person and answered the questions by pointing to one of the faces on the computer screen; adults participated online. The sides of the screen on which faces appeared and the order in which questions were asked were randomized with the constraint that the anticipated choices appeared on the left and right sides of the screen with approximately equal frequency.

Experiment 1

Method

One hundred forty-one children (mean age = 6 years 5 months, range = 3 years 1 month to 10 years 11 months; 68 females, 73 males) participated at local museums and in the laboratory; 99 adults (mean age = 30.23 years, range = 18–67; 54 females, 44 males, 1 unspecified) participated online through SocialSci (www.socialsci.com). Participants attributed trustworthiness ("mean"/"nice"), dominance ("strong"/"not strong"), or competence ("smart"/"not smart") to the pair of faces presented on each trial. Questions regarding a particular trait were asked only on trials in which faces manipulated on that trait dimension were presented. For analyses, children were divided into three age groups: 3- to 4-year-olds (n = 37), 5- to 6-year-olds (n = 50), and 7- to 10-year-olds (n = 54).

Results

Figures 2, 3, and 4 summarize results for all age groups and traits; higher percentages of expected responses (i.e., those predicted on the basis of prior data—e.g., that trustworthy faces would be identified as nice and untrustworthy faces as mean) indicate stronger consensus.

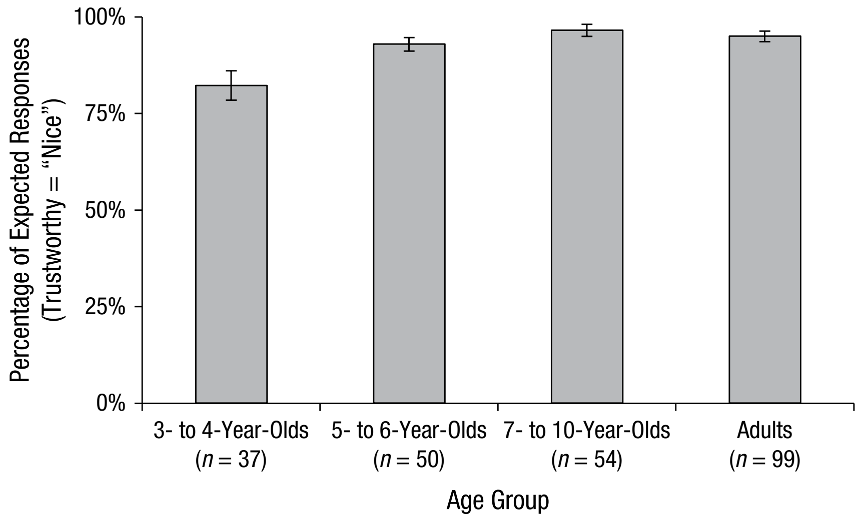

Fig. 2. Results from Experiment 1: average rate at which trustworthy faces were identified as nice and untrustworthy faces were identified as mean in each of the four age groups. Error bars represent ±1 *SEM*.

Combined, all four age groups showed significant consensus compared with chance (50%) when identifying faces as mean or nice (93%; Fig. 2), strong or not strong (85%; Fig. 3), and smart or not smart (76%; Fig. 4).

Critically, all age groups attributed all three traits with significant consensus, $ps < .001$, $ds > 1.08$. However, an analysis of variance (ANOVA) revealed a significant main effect of age group, $F(3, 236) = 17.91$, $p < .001$. Although 3- to 4-year-olds responded with robust and adultlike consensus (72% across all traits), they were less consistent than 5- to 6-year-olds (81%), 7- to 10-year-olds (88%), and adults (89%). One-way ANOVAs followed by post hoc tests with Sidak corrections for multiple comparisons were used to analyze age differences for each trait. These analyses revealed that when attributing both trustworthiness and dominance, 3- to 4-year-olds were less consistent than all other age groups (all $ps < .01$, $ds > 0.59$), which exhibited equivalent consistency (all $ps > .23$, $ds < 0.40$).

Attributions of competence showed a different developmental pattern whereby consensus increased primarily between the ages of 5 to 6 and 7 to 10. Consensus of 3- to 4-year-olds (68%) was identical to that of 5- to 6-year-olds (66%; $p = 1.00$, $d = 0.07$) but lower than that of 7- to

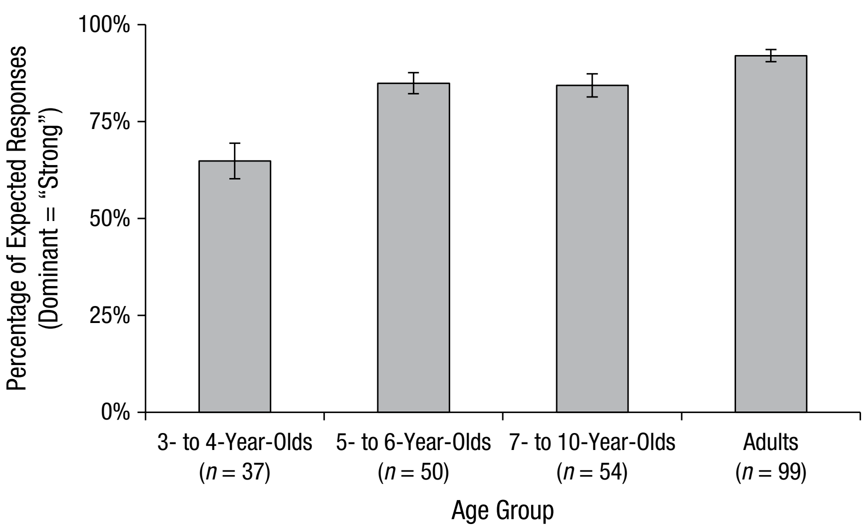

Fig. 3. Results from Experiment 1: average rate at which dominant faces were identified as strong and submissive faces were identified as not strong in each of the four age groups. Error bars represent ±1 *SEM*.

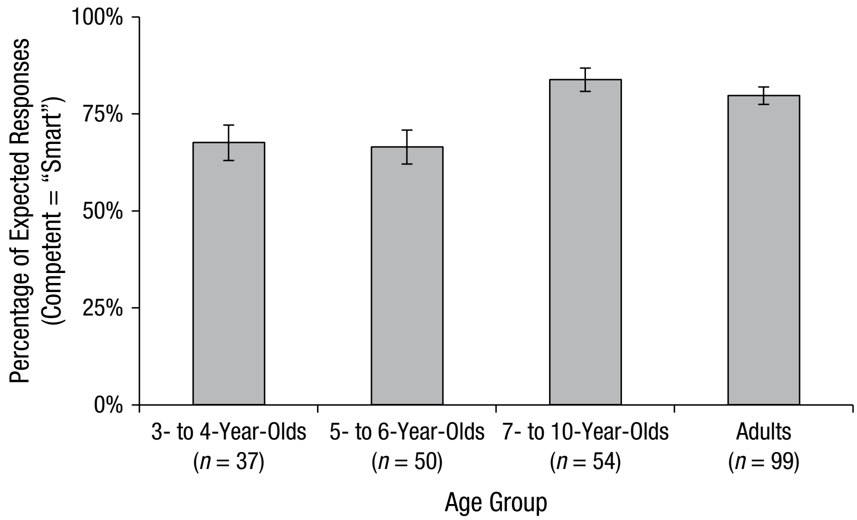

Fig. 4. Results from Experiment 1: average rate at which competent faces were identified as smart and incompetent faces were identified as not smart in each of the four age groups. Error bars represent ±1 *SEM*.

10-year-olds (84%; $p < .05$, $d = 0.64$) and marginally lower than that of adults (80%; $p = .08$, $d = 0.47$).[2] In similar fashion, 5- to 6-year-olds were less consistent than both 7- to 10-year-olds ($p < .01$, $d = 0.67$) and adults ($p < .05$, $d = 0.51$). Seven- to 10-year-olds and adults attributed competence with similar consensus ($p = .91$, $d = 0.18$).

Data were further analyzed using a 3 (trait: trustworthiness vs. dominance vs. competence) × 4 (age group: 3- to 4-year-olds vs. 5- to 6-year-olds vs. 7- to 10-year-olds vs. adults) mixed-model ANOVA with repeated measures on the first factor. This analysis revealed main effects of both trait, $F(2.028,$[3] $472) = 42.66$, and age group, $F(3, 236) = 18.09$, $ps < .001$. These main effects were qualified by a Trait × Age Group interaction, $F(6, 472) = 4.031$, $p < .01$. Within-subjects contrasts revealed that consensus was highest for judgments of trustworthiness (91.6%). Consistency for judgments of trustworthiness was significantly higher than that for judgments of dominance (81.5%), $F(1, 236) = 54.24$, $p < .001$, which was higher than that for judgments of competence (74.4%), $F(1, 236) = 10.10$, $p < .01$.

Overall, the data suggest that children's face-to-trait inferences reach adultlike consensus at an impressively early age. For all three traits tested, children in the youngest age group responded with striking consistency that greatly exceeded chance responding, although they were typically less consistent than older participants.

Although consensus was consistently high across all age groups and traits, the consensus that emerged for ratings of trustworthiness was significantly greater than that obtained for ratings of dominance and competence. This suggests that "mean" and "nice" judgments might emerge earlier than other judgments. If so, such judgments might be fundamental to face-to-trait inference and therefore broadly applied to faces varying in trait dimensions other than trustworthiness. We explored this possibility in Experiment 2.

Experiment 2

In Experiment 2, we tested whether "mean"/"nice" judgments emerge when observers view faces that vary in dominance and competence instead of trustworthiness. Given the primacy of valence evaluations in social judgments, children might robustly apply basic "mean"/"nice" judgments to faces varying in traits other than trustworthiness. If such evaluations rely on specific features that uniquely vary among faces perceived as differing in trustworthiness, however, consensus should not emerge when this global evaluation is applied to faces that vary in traits other than trustworthiness.

Method

A total of 213 children (mean age = 5 years 11 months, range = 3 years 1 month to 10 years 8 months; 110 females, 101 males, 2 unspecified) participated at museums and in the laboratory; 301 adults (mean age = 28.9 years, range = 18–72; 142 females, 153 males, 6 unspecified) participated online through SocialSci and Qualtrics (www.qualtrics.com).

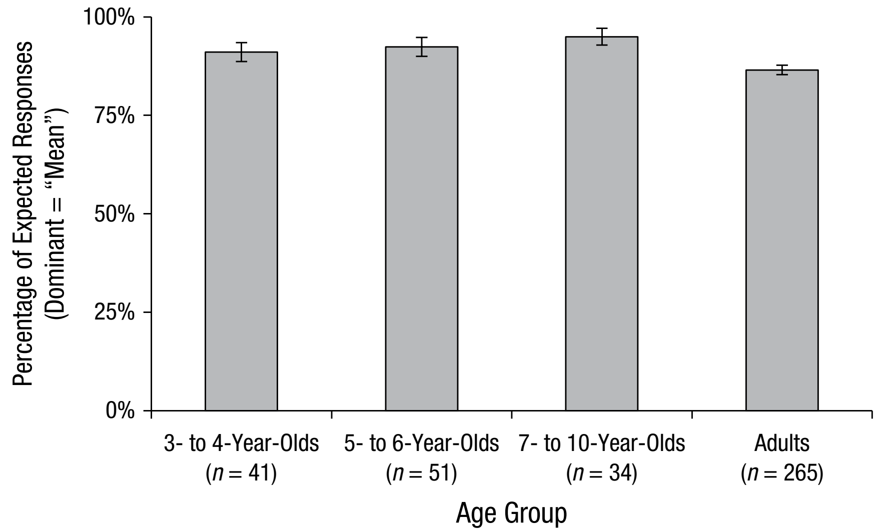

Fig. 5. Results from Experiment 2: average rate at which dominant faces were identified as mean and submissive faces were identified as nice in each of the four age groups. Error bars represent ±1 *SEM*.

Participants viewed two of the face sets used in Experiment 1: the faces varying in perceived dominance and those varying in perceived competence. Verbal prompts elicited solely judgments of which face in each pair was "mean" or "nice" (i.e., "Which of these people is very mean/nice?"). Unlike in Experiment 1, face trait was a between-subjects variable. Sample sizes for all age groups for each face trait are displayed in Figures 5 and 6.

Results

As in Experiment 1, consensus of judgment was strikingly high, vastly exceeding chance responding (50%) for all age groups and both traits, *p*s < .001, *d*s > 2.15 (see Figs. 5 and 6). Consensus of judgments based on facial dominance (i.e., dominant = mean) showed developmental invariance, ranging from 87% to 95%, with no significant pairwise differences between any age groups,

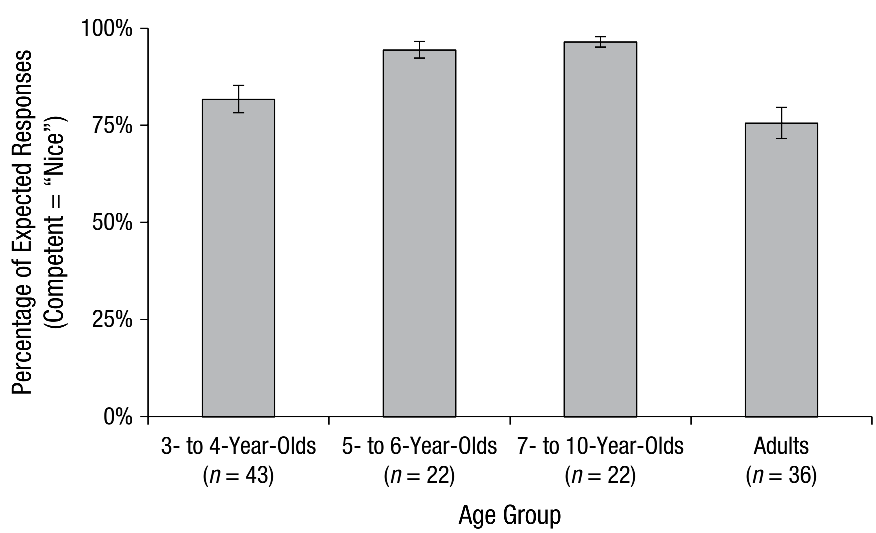

Fig. 6. Results from Experiment 2: average rate at which competent faces were identified as nice and incompetent faces were identified as mean in each of the four age groups. Error bars represent ±1 *SEM*.

all *p*s > .06, *d*s < 0.50. Children of all ages also showed robust consensus in judgments based on facial competence (82%–96%); adults, however, showed markedly lower consensus (76%) than did 5- to 6-year-olds (94%) and 7- to 10-year-olds (96%), *p*s < .01, *d*s > 1.03. Consensus in judgments based on facial competence increased with age among children, with 7- to 10-year-olds responding significantly more consistently than 3- to 4-year-olds (82%), $p < .05$, $d = 0.83$.

Data were further analyzed using a 2 (face trait: dominance vs. competence) × 4 (age group: 3- to 4-year-olds vs. 5- to 6-year-olds vs. 7- to 10-year-olds vs. adults) between-subjects ANOVA. Main effects emerged for both age group, $F(3, 506) = 10.804$, $p < .001$, and face trait, $F(1, 506) = 3.721$, $p = .054$, and there was a significant Age Group × Face Trait interaction, $F(3, 506) = 2.674$, $p < .05$. When we collapsed the data across face trait, we found an age-related increase in consensus among children, with 7- to 10-year-olds showing significantly greater consensus than 3- to 4-year-olds (Sidak post hoc $p < .05$). Adults showed the same consensus as 3- to 4-year-olds ($p > .99$) but were less consistent than 5- to 6-year-olds and 7- to 10-year-olds (*p*s < .01).

One-way ANOVAs on combined data from the two experiments explored the possibility that participants were more consistent in their "mean" and "nice" judgments than in their more specific attributions of strength and intelligence. These analyses compared consensus for "mean"/"nice" evaluations (Experiment 1) with consensus for specific trait judgments (Experiment 2) for faces varying in dominance and competence. Overall, "mean"/"nice" judgments were significantly more consistent than the more specific judgments for faces varying in dominance, $F(1, 629) = 5.332$, $p < .05$, and competence, $F(1, 361) = 10.709$, $p < .01$. All groups of children were significantly more consistent in evaluating the faces as mean or nice than in attributing strength and intelligence to the faces, *p*s < .05. However, adults' "mean"/"nice" judgments of faces varying in dominance were less consistent than their attributions of strength to those faces, $F(1, 362) = 6.441$, $p < .05$, and adults were equally consistent whether they attributed strength to faces varying in competence or evaluated those faces as mean or nice, $F(1, 133) = 0.816$, n.s.

General Discussion

Children in both experiments—even at the earliest ages tested—made reliable inferences about character that approached adult levels of consensus, and by age 7, children showed a level of consensus that matched adults'. In particular, participants of all ages robustly generated basic "mean"/"nice" judgments in response to a variety of facial characteristics.

In both experiments, judgments based on facial competence showed different developmental patterns than those based on facial trustworthiness and dominance. Experiment 2 also produced the seemingly anomalous result that adults were *less* consistent than 5- to 6-year-olds and 7- to 10-year-olds when evaluating faces varying in competence as mean or nice. It is possible that the face-to-trait judgments of adults are more differentiated than those of children, who rely more on global valence. Increasing sensitivity to features other than those affecting global valence might also account for developmental increases in reliability in attributions of strength and intelligence to faces (Experiment 1). In addition, the model for facial competence that we used may be less effective than the models for facial trustworthiness and dominance.

The striking consensus in "mean"/"nice" judgments that we observed for faces varying on all three trait dimensions suggests that such evaluations might underlie the consensus in face-to-trait inferences observed in Experiment 1. In fact, principal component analyses of trait judgments of faces show that trustworthiness judgments are strongly correlated with the first principal component ($r > .90$), interpreted as valence (Oosterhof & Todorov, 2008), and the computational model of face trustworthiness closely resembles a valence model based on multiple social judgments (Said, Dotsch, & Todorov, 2010). Further research probing the relationship between "mean"/"nice" judgments and specific trait inferences will be necessary to evaluate this possibility.

These two experiments provide a clear demonstration that children as young as 3 to 4 years of age show an adultlike tendency both to attribute specific traits to faces and to evaluate faces as mean or nice on the basis of the faces' appearance. It is possible that attractiveness underlies character inferences, particularly for faces varying in trustworthiness and competence. However, recent work has shown that the facial features manipulated in the computational models elicit divergent trait judgments irrespective of attractiveness (Todorov, Dotsch, et al., 2013).

These data leave open the question of when face-to-trait inference first emerges during development. Animation-based stimuli may enable researchers to study even younger populations. If such inferences take root early in development, as the data suggest, even infants might associate faces with trait-consistent behaviors, such as those conveying prosociality (Hamlin, Wynn, & Bloom, 2007) or dominance (Mascaro & Csibra, 2012).

The predisposition to make rapid and unreflective judgments based on scant facial information is a pervasive form of social judgment. Prior work suggests that such inferences have important real-world consequences. We have demonstrated that face-to-trait judgments are robust by age 3, and that certain judgments reach fully adultlike levels when children are 5 to 6 years of age. By revealing the young age at which children make face-to-trait inferences, these data challenge accounts positing slow social-learning mechanisms that develop through the gradual detection and internalization of environmental regularities (e.g., Fazio, Sanbonmatsu, Powell, & Kardes, 1986; Smith & DeCoster, 2000).

Author Contributions

E. J. Cogsdill and M. R. Banaji developed the study concept and design. Data collection was performed by E. J. Cogsdill. Data analysis and interpretation were performed by E. J. Cogsdill under advisement from the other three authors. E. J. Cogsdill drafted the manuscript, and all authors provided critical revisions and approved the final manuscript for submission.

Acknowledgments

We thank Larisa Heiphetz and Steven Lehr for comments on an earlier draft and members of the Implicit Social Cognition Laboratory and the Laboratory for Developmental Studies at Harvard University for their support.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

This work was supported by grants and fellowships from the E. J. Safra Center for the Study of Ethics and the Santa Fe Institute to M. R. Banaji and by a grant from the National Institutes of Health to E. S. Spelke (HD 23103).

Open Practices

All data, face stimuli, and prompts used have been made publicly available via Open Science Framework and can be accessed at https://osf.io/c5kme/. The complete Open Practices Disclosure for this article can be found at http://pss.sagepub.com/content/by/supplemental-data. This article has received badges for Open Data and Open Materials. More information about the badges can be found at https://osf.io/tvyxz/wiki/view/ and http://pss.sagepub.com/content/25/1/3.full/.

Notes

1. In this article, we do not address the veridicality of face-to-trait inference, as other researchers have (e.g., Carré, McCormick, & Mondloch, 2009). Although that is an important topic, we focus on the development of such inferences from the earliest ages at which children can be tested.

2. This result was statistically significant before correction for multiple comparisons, $t(134) = 2.607$, $p = .01$.

3. Degrees of freedom were adjusted using a Greenhouse-Geisser correction after the trait variable failed a test of sphericity, Mauchly's $W = .726$, $p < .01$.

References

Antonakis, J., & Dalgas, O. (2009). Predicting elections: Child's play! *Science*, *323*, 1183.

Barenboim, C. (1981). The development of person perception in childhood and adolescence: From behavioral comparisons to psychological constructs to psychological comparisons. *Child Development*, *52*, 129–144. doi:10.2307/1129222

Bar-Haim, Y., Ziv, T., Lamy, D., & Hodes, R. M. (2006). Nature and nurture in own-race face processing. *Psychological Science*, *17*, 159–163. doi:10.1111/j.1467-9280.2006.01679.x

Blair, I. V., Judd, C. M., & Chapleau, K. M. (2004). The influence of Afrocentric facial features in criminal sentencing. *Psychological Science*, *15*, 674–679. doi:10.1111/j.0956-7976.2004.00739.x

Carré, J. M., McCormick, C. M., & Mondloch, C. J. (2009). Facial structure is a reliable cue of aggressive behavior. *Psychological Science*, *20*, 1194–1198. doi:10.1111/j.1467-9280.2009.02423.x

Clément, F., Bernard, S., Grandjean, D., & Sander, D. (2013). Emotional expression and vocabulary learning in adults and children. *Cognition & Emotion*, *27*, 539–548.

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*, 229–238. doi:10.1037/0022-3514.50.2.229

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature*, *450*, 557–559. doi:10.1038/nature06288

Keating, C. F., & Bai, D. L. (1986). Children's attributions of social dominance from facial cues. *Child Development*, *57*, 1269–1276. doi:10.2307/1130449

Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, *126*, 390–423. doi: 10.1037//0033-2909.126.3.390

Mascaro, O., & Csibra, G. (2012). Representation of stable social dominance relations by human infants. *Proceedings of the National Academy of Sciences, USA*, *109*, 6862–6867. doi:10.1073/pnas.1113194109

Montepare, J. M., & Zebrowitz-McArthur, L. (1989). Children's perceptions of babyfaced adults. *Perceptual & Motor Skills*, *69*, 467–472.

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

Quinn, P., Yahr, J., Kuhn, A., Slater, A., & Pascalis, O. (2002). Representation of the gender of human faces by infants: A preference for female. *Perception*, *31*, 1109–1121. doi:10.1068/p3331

Ramsey, J. L., Langlois, J. H., Hoss, R. A., Rubenstein, A. J., & Griffin, A. (2004). Origins of a stereotype: Categorization of facial attractiveness by 6-month-old infants. *Developmental Science*, *7*, 201–211. doi:10.1111/j.1467-7687.2004.00339.x

Rholes, W. S., & Ruble, D. N. (1984). Children's understanding of dispositional characteristics of others. *Child Development*, *55*, 550–560. doi:10.2307/1129966

Rule, N. O., & Ambady, N. (2008). The face of success: Inferences from chief executive officers' appearance predict company profits. *Psychological Science*, *19*, 109–111. doi:10.1111/j.1467-9280.2008.02054.x

Rule, N. O., Ambady, N., Adams, R. B., Jr., Ozono, H., Nakashima, S., Yoshikawa, S., & Watabe, M. (2010). Polling the face: Prediction and consensus across cultures. *Journal of Personality and Social Psychology*, *98*, 1–15. doi:10.1037/a0017673

Said, C. P., Dotsch, R., & Todorov, A. (2010). The amygdala and FFA track both social and non-social face dimensions. *Neuropsychologia*, *48*, 3596–3605. doi: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. doi:10.1037/a0014681

Simion, F., Macchi Cassia, V., Turati, C., & Valenza, E. (2001). The origins of face perception: Specific versus non-specific mechanisms. *Infant and Child Development*, *10*, 59–65. doi:10.1002/icd.247

Smith, E. R., & DeCoster, J. (2000). Dual-process models in social and cognitive psychology: Conceptual integration and links to underlying memory systems. *Personality and Social Psychology Review*, *4*, 108–131. doi:10.1207/S15327957PSPR0402_01

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. doi:10.1037/a0032335

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science*, *308*, 1623–1626. doi:10.1126/science.1110589

Todorov, A., Mende-Siedlecki, P. M., & Dotsch, R. (2013). Social judgments from faces. *Current Opinion in Neurobiology*. Advance online publication. doi:10.1016/j.conb.2012.12.010

Todorov, A., & Oosterhof, N. N. (2011). Modeling social perception of faces. *IEEE Signal Processing Magazine*, *28*, 117–122. doi:10.1109/MSP.2010.940006

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

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, *72*, 655–684. doi:10.1111/1467-8624.00304

Zebrowitz, L. A., & Montepare, J. M. (2008). Social psychological face perception: Why appearance matters. *Social & Personality Psychology Compass*, *2*, 1497–1517. doi:10.1111/j.1751-9004.2008.00109.x.Social

Zebrowitz, L. A., & Rhodes, G. (2004). Sensitivity to "bad genes" and the anomalous face overgeneralization effect: Cue validity, cue utilization, and accuracy in judging intelligence and health. *Journal of Nonverbal Behavior*, *28*, 167–185.