

# Robust Effects of Affective Person Learning on Evaluation of Faces

Sara C. Verosky  
Oberlin College

Jenny Porter  
Clayton Early Learning and University of Colorado Boulder

Joel E. Martinez and Alexander Todorov  
Princeton University

People form impressions of others from multiple sources of information. Facial appearance is one such source and judgments based on facial appearance are made after minimal exposure to faces. A more reliable source of information is affective person learning based on others' past actions. Here we investigated whether the effects of such appearance-independent learning on face evaluation emerge after rapid face exposure, a response deadline procedure, and a lack of explicit recognition of the faces. In three experiments, participants learned to associate novel faces with negative and positive behaviors, and then evaluated the faces presented on their own, without the behaviors. Even after extremely brief exposures (e.g., 35 ms), participants evaluated faces previously associated with negative behaviors more negatively than those associated with positive behaviors (Experiment 1). The learning effect persisted when participants were asked to evaluate briefly presented faces before a response deadline (Experiment 2), although the effect was diminished. Finally, although this learning effect increased as a function of face recognition (Experiment 3), it was present with only minimal recognition, suggesting that participants do not need to deliberately retrieve behavioral information for it to influence face evaluation. Together, the findings suggest that person learning unrelated to facial appearance is a powerful determinant of face evaluation.

*Keywords:* affect, face recognition, learning, minimal exposure, person knowledge

In *Baboon Metaphysics*, their book on the complexities of baboon life, primatologists Dorothy Cheney and Robert Seyfarth write, "Any way you look at it, most of the problems facing baboons can be expressed in two words: other baboons." This statement applies with even greater force to humans. People mostly care about other people and, accordingly, readily and effortlessly form impressions of others from minimal information. One such source of information with dubious validity is facial appearance (for a review, see Todorov, 2017; Todorov, Olivola, Dotsch, & Mende-Siedlecki, 2015). People form inferences about a person's character (e.g., trustworthiness, aggressiveness) after extremely brief exposure to the person's face. These inferences made after viewing faces for limited times correlate highly with inferences made after much longer times, suggesting that people are relying on the same information in both cases (Bar, Neta, & Linz, 2006; Rule, Ambady, & Adams, 2009; Todorov, Pakrashi, & Oosterhof, 2009; Willis & Todorov, 2006). Underscoring this

point, additional exposure time beyond 100 ms barely increases the magnitude of correlation with judgments made in the absence of time constraints, although it substantially increases confidence (Todorov et al., 2009; Willis & Todorov, 2006).

The effects of appearance can be explained by a correspondence between impressions and configurations of facial features. In fact, a number of studies have identified specific constellations of physical features that lead to the perception of different traits (Oosterhof & Todorov, 2008; Walker & Vetter, 2009, 2016; Zebrowitz, 2004; Zebrowitz, Fellous, Mignault, & Andreoletti, 2003; Zebrowitz, Kikuchi, & Fellous, 2010). For instance, the structural similarity of a neutral face to a face with an emotional expression has been found to influence a variety of trait judgments (Adams, Hess, & Kleck, 2015; Adams, Nelson, Soto, Hess, & Kleck, 2012; Montepare & Dobish, 2003; Oosterhof & Todorov, 2008, 2009; Said, Sebe, & Todorov, 2009; Zebrowitz, 2004; Zebrowitz et al., 2010). That is, the information that people use to make trait judgments is literally "coded" in the face, and people are able to extract this information after extremely brief exposure to faces.

People also form impressions from more reliable sources of information such as others' past actions. When faces are presented with descriptions of behaviors (e.g., John kicked the puppy), people readily learn to associate the traits implied by the behaviors (e.g., cruel) with the people pictured (Bliss-Moreau, Barrett, & Wright, 2008; Carlston & Skowronski, 1994; Todorov & Uleman, 2002, 2003, 2004). Predictably, such associations influence subsequent judgments of faces (Rudoy & Paller, 2009; Todorov & Olson, 2008; Verosky & Todorov, 2010, 2013). In a recent study, participants saw as many as 500 faces paired with negative, neu-

---

Sara C. Verosky, Department of Psychology, Oberlin College; Jenny Porter, Department of Research and Evaluation, Clayton Early Learning, Denver, CO, and Institute of Cognitive Science, University of Colorado Boulder; Joel E. Martinez and Alexander Todorov, Department of Psychology, Princeton University.

We thank Russell Dinkins, Virginia Falvello, Dongwon Oh, and Kimberly Solomon for their help with data collection.

Correspondence concerning this article should be addressed to Sara C. Verosky, Department of Psychology, Oberlin College, 120 West Lorain, Oberlin, OH 44074. E-mail: [sverosky@oberlin.edu](mailto:sverosky@oberlin.edu)

tral, or positive behaviors (Falvello, Vinson, Ferrari, & Todorov, 2015). Despite the large number of faces and behaviors, when subsequently asked to evaluate the faces presented alone, participants evaluated faces associated with negative behaviors more negatively than faces associated with positive behaviors.

In contrast to judgments based on physical appearance, affective associations derived from others' past actions are not "coded" in the face. At the moment of evaluation, these associations need to be triggered by the face. Rudoy and Paller (2009) describe judgments relying on these two types of information as perceptual and memory-based processes, respectively. In their work, perceptual processes refer to judgments of face trustworthiness, while memory-based processes refer to the influence of prior associations between the face and negative or positive adjectives. Despite the fact that these two types of information—appearance and learned associations with appearance—are integrated in behavioral judgments of faces, Rudoy and Paller (2009) found separate neural signatures for these processes using event-related potentials, with the signature for the perceptual process occurring earlier in time. Consistent with this finding, reducing the amount of time participants had to respond to faces increased the influence of appearance on face evaluation relative to that of learning (Rudoy & Paller, 2009).

Like adults, children use both perceptual and memory-based information when forming impressions of others. Studies where children fit faces to simple trait descriptions have found that children show an adult-like consensus in their judgments of facial appearance quite early in life (Cogsdill & Banaji, 2015; Cogsdill, Todorov, Spelke, & Banaji, 2014). Other studies where children choose which informant to trust have demonstrated that children are sensitive to information about both appearance and behavior (Bascandzic & Harris, 2014, 2016; Koenig, Clément, & Harris, 2004). In these studies, children were shown a novel object, along with a pair of informants who each offered a different name for the object. When the informants differed in their attractiveness, children paid attention to the more attractive informant (Bascandzic & Harris, 2014). When the informants differed in their history of providing accurate names for known objects, children paid attention to the more accurate of the two (Bascandzic & Harris, 2016; Koenig et al., 2004). When the two cues were pitted against each other, such that the more attractive informant was also less accurate, children paid attention to both cues (Bascandzic & Harris, 2016). Thus, although past history of accurate responses played a role in children's decisions about whom to trust, it was not enough to outweigh the effect of attractiveness.

While the studies examining the influence of behavior on face evaluation mentioned thus far differ in a variety of ways, they have one thing in common: The faces in these studies were presented for long amounts of time during the judgments, which, in some sense, renders the findings of the effect of behavior learning on face judgments unsurprising. People are extremely good at forming affective associations with faces (Bliss-Moreau et al., 2008; Falvello et al., 2015; Johnson, Kim, & Risse, 1985; Todorov & Olson, 2008; Todorov & Uleman, 2003; Tranel & Damasio, 1993) and when given sufficient time they can retrieve these associations, which in turn can influence face judgments.

In the current research, we investigated whether associative learning effects can be detected under conditions that should make the retrieval of the associations more difficult. These conditions included rapid presentation of faces and a response deadline procedure, which forces participants to respond within a limited time

window. We did not expect these conditions to interfere with judgments based on facial appearance (Ballew & Todorov, 2007; Rudoy & Paller, 2009). However, we were interested in how they would affect judgments based on past behavior. On the one hand, to the extent that using behavioral information in impression formation requires extensive cognitive processing, these conditions might prevent or decrease the influence of learning on face evaluation. On the other hand, to the extent that affective person knowledge becomes an inherent part of the face representation, its retrieval could instead be effortless and spontaneous.

To investigate these possibilities, we conducted three experiments. In each of these experiments, we paired faces with negative and positive behaviors and then we asked participants to evaluate the rapidly presented faces on their own, without the behaviors. In the first experiment, we examined the influence of learning and behavior across different exposure times to faces. We were interested in whether the effect of learning would be detectable after the short exposure time and in whether additional visual information would be useful for making judgments based on appearance and learning. In the second experiment, in addition to manipulating exposure time, we introduced the response deadline procedure. This allowed us to test whether the influence of learning would persist when participants were given only a short time to process the available visual information. Finally, in the third experiment, we explored the relationship between prior affective learning and explicit face recognition. Participants evaluated a series of briefly presented faces. Immediately after evaluating each face, participants indicated whether they recognized it from the first part of the experiment. The brief presentation times ensured a range of recognition responses, which allowed us to examine whether behavior would continue to influence face evaluation even when participants reported difficulty recognizing the faces.

## Experiment 1

In the first experiment, we investigated the influence of both behavior learning and face appearance on judgments made under limited exposure. We selected trustworthy- and untrustworthy-looking groups of faces based on pilot ratings and paired them with negative, positive, or no behavioral information. To test how the contributions of learning and the physical properties of the face change over time, we manipulated exposure time between participants. First, as shown in multiple studies (Bar et al., 2006; Todorov, Loehr, & Oosterhof, 2010; Todorov et al., 2009; Willis & Todorov, 2006), we expected that the effects of appearance would be detectable after brief exposures. Second and more importantly, we tested whether the effects of learning would be also detectable. For both appearance and learning, we expected that their effects would increase with the increase in exposure to faces. These increases are informative about the underlying evaluation processes: if additional visual information is useful for making these judgments, then we would expect stronger effects with longer exposure times.

## Method

**Participants.** Forty undergraduates and members of the Princeton University community participated for partial course credit or payment (21 female, 19 male;  $M_{\text{age}} = 19.18$ ,  $SD_{\text{age}} = 1.22$ ). This sample size was based on the power analysis described below. Data from one participant were excluded because that participant gave the same response for 35 of the 36 faces.

**Statistical power.** We determined an appropriate sample size based on the results of a pilot study where participants ( $N = 13$ ) viewed nine learned faces paired with negative, neutral, and positive behaviors following the procedure described in Verosky and Todorov (2010) and then evaluated those faces after presentation times of 40 ms. We did not systematically manipulate facial appearance (e.g., face trustworthiness) or presentation time (as all faces were presented for this same brief amount of time) in this pilot study. Participants evaluated faces that were previously paired with negative behaviors ( $M = 4.36$ ,  $SD = 1.82$ ) significantly more negatively than those paired with positive behaviors ( $M = 5.97$ ,  $SD = 1.46$ ;  $t(12) = 2.61$ ,  $p = .02$ ,  $d_{rm} = .97$ ) and we used this mean difference in evaluation ( $M = 1.62$ ,  $SD = 2.24$ ) to compute sample size. We found that to have 80% power to detect an effect of behavior the size of that observed in this pilot study, we would need a sample size of 18. Although the learning procedure used in the current experiment was less stringent than that used in the pilot study, we used a greater number of learned faces to account for this difference. Because exposure time in the current experiment was manipulated between participants, this meant we needed a total sample size of at least 36.

**Stimuli.** A separate group of participants ( $N = 18$ ) evaluated the trustworthiness of 45 photographs of men with neutral facial expressions taken from a set of digital black-and-white photographs (Goldstone & Steyvers, 2001; originally from Kayser, 1997). We selected the 18 faces with the highest and 18 with the lowest mean ratings to use as the trustworthy- and untrustworthy-looking faces.

The behaviors were selected from a large database of behaviors based on goodness ratings (Fuhrman, Bodenhausen, & Lichtenstein, 1989). The negative behaviors ( $M_{words} = 8.89$ ,  $SD_{words} = 2.19$ ) did not differ from the positive behaviors in their length ( $M_{words} = 9.39$ ,  $SD_{words} = 1.93$ ;  $t(70) = 1.03$ ,  $p = .31$ ,  $d = .25$ ). In addition, the goodness ratings for the negative behaviors ( $M = 1.66$ ,  $SD = .40$ ) did not differ from the goodness ratings for the positive behaviors ( $M = 8.25$ ,  $SD = .36$ ) in their distance from the midpoint of the scale (midpoint = 5; independent samples  $t$  test on distance:  $t(70) = 1.02$ ,  $p = .31$ ,  $d = .24$ ). However, even though the negative and positive behaviors were equally extreme in terms of valence, the negative behaviors ( $M = 3.00$ ,  $SD = 1.07$ ) were seen as less typical than the positive behaviors ( $M = 6.11$ ,  $SD = 1.37$ ;  $t(70) = 10.73$ ,  $p < .001$ ,  $d = 2.57$ ). This discrepancy in the perceived typicality of the negative versus positive behaviors is consistent with prior work demonstrating that immoral behaviors are perceived as less frequent (Mende-Siedlecki, Baron, & Todorov, 2013) and treated as more diagnostic of traits than moral behaviors (Skowronski & Carlston, 1987, 1989).

The trustworthy- and untrustworthy-looking sets of faces were each divided into three groups of six faces and the groups were paired with negative behaviors, positive behaviors, or they were shown without behaviors. The face and behavior groups were counterbalanced across participants. Stimuli were presented on 17 in. CRT monitors with a refresh rate of 85 Hz.

#### Procedures.

**Learning.** During the first part of the experiment, participants saw faces paired with negative and positive behaviors or faces presented alone. The behaviors were short sentences such as “He

hit a car and left the scene of the accident” (negative) or “He volunteered to stay late to help a coworker” (positive). Each face was shown for 5 s. For the faces presented with behaviors, participants were instructed to form impressions of the people pictured by imagining them performing the behaviors. Participants saw each face three times. The faces paired with behaviors were presented with three behaviors of the same valence. The faces were blocked such that each face was shown once before any of the faces were shown a second time. Within each block, the faces appeared in a different random order.

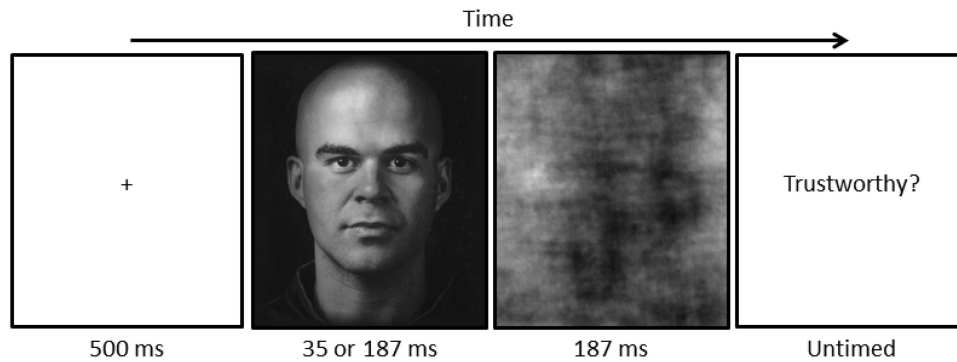
**Face evaluation.** Immediately after seeing the faces paired with behaviors, participants were asked to evaluate the trustworthiness of the faces presented on their own, without the accompanying behaviors. Trustworthiness judgments were used because they have been shown to be a very good proxy for valence evaluation (Oosterhof & Todorov, 2008; Sutherland et al., 2013). Participants were instructed that each face would remain on the screen for a short amount of time. They were told that the task was designed to be difficult, and they were asked to rely on their gut feeling when making their response. Each trial began with 1,500 ms of a blank screen. The blank screen was followed by a 500-ms fixation cross, which indicated that a face was about to appear (see Figure 1). Presentation time was manipulated between participants such that participants viewed all of the faces for 35 ms (3 refresh cycles) or 187 ms. After each face, a phase-scrambled version of the face was shown for 187 ms to prevent further visual processing. After the mask, the word “Trustworthy?” appeared along with a response scale. The scale remained on the screen until participants responded using the number keys 1 (*extremely untrustworthy*) to 9 (*extremely trustworthy*). The order of the faces was randomized for each participant.

After evaluating all of the faces under limited presentation time, participants evaluated the faces under unlimited presentation time. These ratings were collected to ensure that learning had occurred. Each trial began with a 500-ms fixation cross. Then, a face was presented and remained on the screen until participants responded using the same scale as above.

The overall design of the timed presentation portion of the experiment was a 2 (appearance: trustworthy or untrustworthy)  $\times$  3 (behavior: none, negative, or positive)  $\times$  2 (time: 35 or 187 ms) mixed analysis of variance (ANOVA), with the first two factors within-subject and the last between subjects. The data used in these analyses and in the analyses for subsequent experiments are available on the Open Science Framework website (<https://osf.io/9jwv2/>).

## Results

**Manipulation check.** To ensure that learning had occurred, we examined face evaluation after unlimited presentation times. As expected, we found an effect of behavior on face evaluation ( $F(2, 74) = 59.71$ ,  $p < .001$ ,  $\eta_p^2 = .62$ ). Participants evaluated faces that were presented with negative behaviors during the first part of the experiment ( $M = 3.90$ ,  $SD = .97$ ) more negatively than faces that were presented without behaviors ( $M = 4.79$ ,  $SD = .55$ ) or with positive behaviors ( $M = 5.83$ ,  $SD = .89$ ;  $t(38) > 5.79$ ,  $p < .001$ ,  $d_{rm} > 1.09$  for all pairwise comparisons). There was also a significant effect of appearance ( $F(1, 37) = 41.53$ ,  $p < .001$ ,  $\eta_p^2 = .53$ ), such that participants evaluated untrustworthy-looking faces ( $M = 4.44$ ,  $SD = .73$ ) more negatively than trustworthy-looking



*Figure 1.* Face evaluation task for Experiment 1. Each trial began with 1,500 ms of a blank screen, which was followed by the events shown above. When the “trustworthy?” probe appeared, participants evaluated the face according to how trustworthy they thought it looked. Photo © Alex Kayser, from the book *Heads* by Alex Kayser, Abbeville Press, 1985.

faces ( $M = 5.24$ ,  $SD = .58$ ). None of the other effects reached significance (Appearance  $\times$  Behavior:  $F(2, 74) = 2.03$ ,  $p = .14$ ,  $\eta_p^2 = .05$ ; Appearance  $\times$  Time Group for previous part of the experiment:  $F(1, 37) = 1.67$ ,  $p = .20$ ,  $\eta_p^2 = .04$ ;  $F < 1$  in both of the other cases). The average correlation between ratings made after limited presentation times and unlimited presentation times was  $.43$  ( $SD = .28$ ).

**Findings.** Collapsing across the limited presentation time conditions, appearance influenced face evaluation ( $F(1, 37) = 13.29$ ,  $p = .001$ ,  $\eta_p^2 = .26$ ), such that participants evaluated the untrustworthy-looking faces ( $M = 4.71$ ,  $SD = .70$ ) more negatively than the trustworthy-looking faces ( $M = 5.14$ ,  $SD = .64$ ). More importantly, behavior also influenced face evaluation ( $F(2, 74) = 38.51$ ,  $p < .001$ ,  $\eta_p^2 = .51$ ). Participants evaluated faces that were previously presented with negative behaviors ( $M = 4.21$ ,  $SD = .80$ ) more negatively than faces presented without behaviors ( $M = 5.04$ ,  $SD = .67$ ) or with positive behaviors ( $M = 5.53$ ,  $SD = .90$ ;  $t(38) > 3.53$ ,  $p < .002$ ,  $d_{rm} > .61$  for all pairwise comparisons). There was also a significant interaction between appearance and behavior ( $F(2, 74) = 4.37$ ,  $p = .02$ ,  $\eta_p^2 = .11$ ). This interaction reflected a stronger effect of appearance for faces presented without behaviors (trustworthy—untrustworthy difference:  $M = .70$ ,  $SD = .92$ ) than for faces presented with positive behaviors ( $M = .22$ ,  $SD = .93$ ;  $t(38) = 2.76$ ,  $p = .009$ ,  $d_{rm} = .52$ ) and a marginally stronger effect of appearance for faces presented without behaviors than for faces presented with negative behaviors ( $M = .37$ ,  $SD = 1.05$ ;  $t(38) = 1.79$ ,  $p = .08$ ,  $d_{rm} = .33$ ). The effect of appearance did not differ for faces presented with negative versus positive behaviors,  $t(38) = 1.13$ ,  $p = .27$ ,  $d_{rm} = .15$ .

Next, we examined how judgments based on appearance and learning change over time. We were interested in whether we would be able to detect these effects in the short presentation time condition and if additional exposure time would be useful for making these judgments. With increasing presentation times, we found that the influence of both appearance and behavior on face evaluation increased (see Figure 2), as reflected by a marginally significant interaction between presentation time and appearance ( $F(1, 37) = 3.23$ ,  $p = .08$ ,  $\eta_p^2 = .08$ ) and a significant interaction between presentation time and behavior ( $F(2, 74) = 3.81$ ,  $p = .03$ ,  $\eta_p^2 = .09$ ). After presentation times of 35 ms, participants evalu-

ated the untrustworthy-looking faces ( $M = 4.75$ ,  $SD = .75$ ) more negatively than the trustworthy-looking faces ( $M = 4.97$ ,  $SD = .72$ ), but this difference did not reach significance ( $F(1, 19) = 1.73$ ,  $p = .20$ ,  $\eta_p^2 = .08$ ). After 187 ms, this difference became larger (untrustworthy faces:  $M = 4.67$ ,  $SD = .65$ ; trustworthy faces:  $M = 5.32$ ,  $SD = .50$ ) and significant ( $F(1, 18) = 14.66$ ,  $p = .001$ ,  $\eta_p^2 = .45$ ).

After 35 ms, there was already a significant effect of behavior on evaluation ( $F(2, 38) = 9.14$ ,  $p = .001$ ,  $\eta_p^2 = .32$ ). Participants evaluated faces presented with negative behaviors ( $M = 4.38$ ,  $SD = .89$ ) significantly more negatively than faces presented without behaviors ( $M = 4.90$ ,  $SD = .82$ ;  $t(19) = 2.72$ ,  $p = .01$ ,  $d_{rm} = .61$ ) or with positive behaviors ( $M = 5.31$ ,  $SD = .84$ ;  $t(19) = 3.78$ ,  $p = .001$ ,  $d_{rm} = 1.07$ ); participants also evaluated faces presented without behaviors more negatively than those presented with positive behaviors, although this difference was only marginally significant,  $t(19) = 1.92$ ,  $p = .07$ ,  $d_{rm} = .49$ . After 187 ms, these effects were stronger ( $F(2, 36) = 33.49$ ,  $p < .001$ ,  $\eta_p^2 = .65$ ). Participants again evaluated faces presented with negative behaviors ( $M = 4.03$ ,  $SD = .68$ ) more negatively than faces presented without behaviors ( $M = 5.18$ ,  $SD = .43$ ) or with positive behaviors ( $M = 5.77$ ,  $SD = .92$ ); all pairwise comparisons were significant ( $t(18) > 3.17$ ,  $p < .006$ ,  $d_{rm} > .73$ ). The three-way interaction between appearance, behavior, and presentation time did not reach significance ( $F < 1$ ).

## Discussion

Even minimal exposure to a person’s face is enough for people to make judgments about their character (Bar et al., 2006; Todorov et al., 2009; Willis & Todorov, 2006). In the first experiment, we found that appearance is not the whole story: like judgments made after unlimited time, judgments made after only 35 ms exposure to the faces reflect knowledge about a person’s past actions. With longer exposure times, both the influence of appearance and the influence of behavior learning increased, indicating that the additional visual information was useful for making both types of judgment.



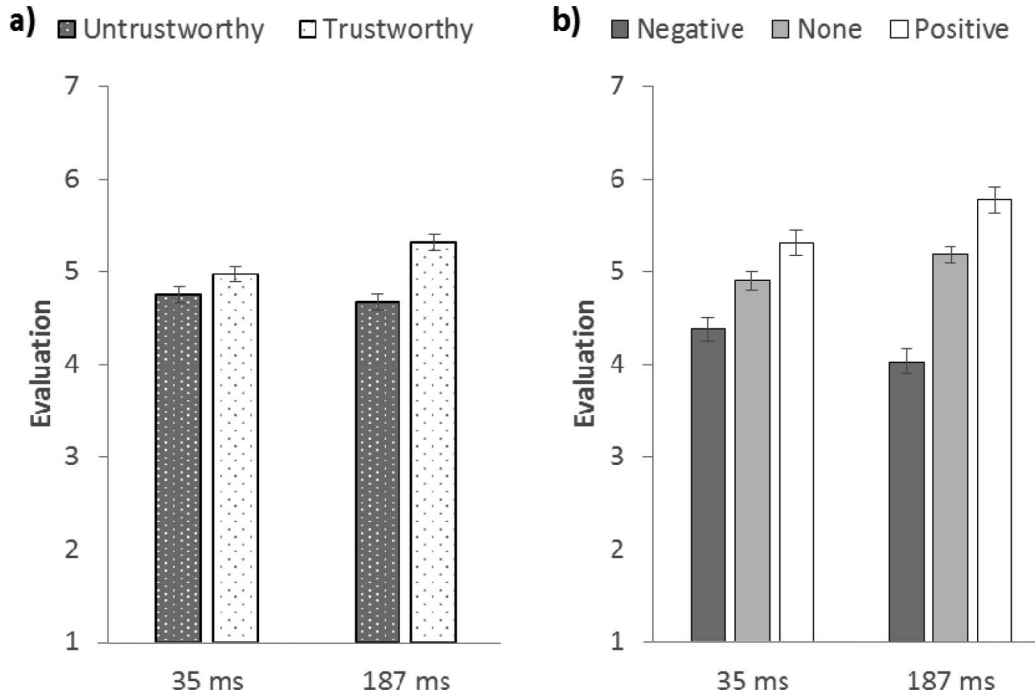


Figure 2. Influence of appearance and behavior on face evaluation after brief exposures ( $N = 39$ ). (a) Mean trustworthiness ratings of untrustworthy- (gray with white dots) and trustworthy-looking faces (white with gray dots). (b) Mean trustworthiness ratings of faces previously presented with negative behaviors (dark gray), no behaviors (light gray), or positive behaviors (white). Error bars represent SEM.

## Experiment 2

In the first experiment, we manipulated the exposure time to faces by displaying the faces for brief periods of time and then replacing them with visual masks. While these visual masks prevented further extraction of visual information from the faces, they did not block processing of the information that had already been extracted. Therefore, it is possible that participants are only able to make judgments of briefly presented faces after extensive processing of the available visual information. To explore how readily participants are able to judge faces presented for brief times, we decided to simultaneously limit both the exposure time to faces and the response time during face evaluation. We did this by modifying the previous experimental design to include a response deadline, which we adapted from the work of Rudoy and Paller (2009).

In Rudoy and Paller's (2009) experiment, participants saw faces paired with negative or positive adjectives, and then they were asked to evaluate faces within 1,500 ms of when they first appeared on the screen or after they had been on the screen for 3,000 ms. Unlike in our study, faces in their study remained on the screen for the entire period of face evaluation. The authors found that the influence of face trustworthiness on evaluation remained constant across response times, but that there was a stronger effect of learning after the longer response time. This led them to conclude that perceptual information is available earlier than memory-based information.

In the current experiment, we were interested in whether the influence of learning observed in the previous experiment would

persist even when participants were asked to evaluate faces within a response deadline. As in the previous experiment, we manipulated exposure time between participants. In addition, we manipulated response deadline within participants, such that participants were asked to evaluate faces either within 1,500 ms or after 3,000 ms of when they first appeared.

## Method

**Participants.** Sixty undergraduates and members of the Princeton University community participated for partial course credit or payment (44 female, 14 male, 2 other;  $M_{\text{age}}^1 = 22.5$ ,  $SD_{\text{age}} = 4.63$ ; primary racial/ethnic background: 13% Black/African American, 25% East Asian/East Asian American, 10% Latino/Hispanic American, 2% Middle Eastern/Arab American, 7% South Asian/South Asian American, 40% White/European American, 3% other; 12% of participants reported that they were multiracial). This sample size was based on the power analysis described below.

**Statistical power.** The sample size for Experiment 1 was based on the expected effect size for behavior from a pilot study. For Experiment 2, we were able to use data from the previous experiment to update this estimate so that it was also based on the effect size for appearance. We found that to have 80% power to detect an effect of the size of the main effect of appearance in

<sup>1</sup> Participants reported their age in an online survey by selecting a value between 18 and 35 or an "over 35" option. Two participants in Experiment 2 reported that they were over 35 and their age was entered as 35 when calculating descriptive statistics for the sample.

Experiment 1, we would need a sample size of 29. Because we manipulated exposure time between participants in Experiment 2 and we wanted to be able to detect an effect of appearance in each timing condition, this meant that we needed a total sample size of 58.

**Stimuli.** This experiment used the same stimuli as Experiment 1.

**Procedures.**

**Learning.** Participants followed the learning procedure from Experiment 1, where they saw faces paired with negative or positive behaviors or faces presented alone.

**Face evaluation.** The face evaluation procedure was very similar to that used in Experiment 1, with the only exception being the introduction of a response deadline, adapted from Rudoy and Paller (2009). Immediately after the learning part of the experiment, participants were asked to evaluate the trustworthiness of the faces presented on their own, without the accompanying behaviors. As in the previous experiment, presentation time was manipulated between participants, such that participants viewed all of the faces for 35 ms or for 187 ms. The response deadline was manipulated within participants: participants were required to either evaluate the faces within 1,500 ms of their presentation or to wait 3,000 ms before evaluating them. On the long deadline trials, a red X remained on the screen until it was time for the participants to respond.

Participants evaluated each of the 36 faces two times in the same response deadline condition. The faces were divided into eight blocks of nine faces each, and the faces assigned to the blocks and the order of the blocks were randomized for each participant. At the beginning of each block, participants were told whether that block would be a short deadline or a long deadline block. Participants evaluated each of the 36 faces one time before rating any of them a second time. The average correlation between first and second ratings of the faces was .41 ( $SD = .24$ ). During data analysis, the two ratings for each face were averaged together. The faces assigned to the short versus long deadline condition were counterbalanced across participants. Because Experiment 1 successfully demonstrated that learning influences face evaluation after limited presentation times, participants were not asked to evaluate the faces under unlimited presentation time.

The overall design of the experiment was a 2 (appearance: trustworthy or untrustworthy)  $\times$  3 (behavior: none, negative, or positive)  $\times$  2 (response deadline: respond within 1,500 ms or after 3,000 ms)  $\times$  2 (time: 35 or 187 ms) mixed ANOVA, with the first three factors within-subject and the last between subjects.

## Results

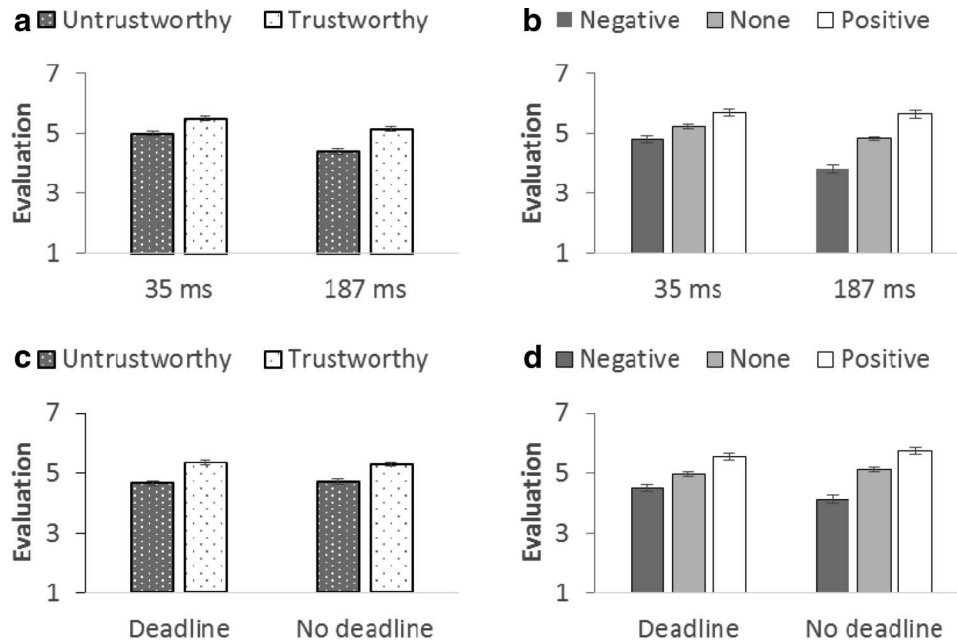
As in the previous experiment, both appearance ( $F(1, 58) = 27.47, p < .001, \eta_p^2 = .32$ ) and behavior ( $F(2, 116) = 49.42, p < .001, \eta_p^2 = .46$ ) influenced face evaluation. Once again, untrustworthy-looking faces ( $M = 4.68, SD = .90$ ) were rated more negatively than trustworthy-looking faces ( $M = 5.31, SD = .90$ ). Faces that were presented with negative behaviors during the first part of the experiment ( $M = 4.31, SD = 1.12$ ) were also evaluated more negatively than faces presented without behaviors ( $M = 5.03, SD = .81$ ) or faces presented with positive behaviors ( $M = 5.65, SD = 1.03; t(59) > 5.43, p < .001, d_{rm} > .65$  for all pairwise comparisons).

For the presentation time manipulation, there was a main effect of presentation time ( $F(1, 58) = 6.38, p = .01, \eta_p^2 = .10$ ), such that participants evaluated faces more positively after presentation times of 35 ms ( $M = 5.23, SD = .79$ ) versus 176 ms ( $M = 4.75, SD = .68$ ). Once again, we were interested in whether the effects of appearance and learning would be present under the short presentation time condition and in whether additional visual information would increase the strength of these effects. With increasing presentation times, the influence of appearance on face evaluation became stronger (Figure 3a), although the interaction between presentation time and appearance did not reach significance ( $F(1, 58) = 1.52, p = .22, \eta_p^2 = .03$ ). After presentation times of 35 ms, participants evaluated the untrustworthy-looking faces ( $M = 4.99, SD = .87$ ) significantly more negatively than the trustworthy-looking faces ( $M = 5.48, SD = .93; F(1, 29) = 8.97, p = .006, \eta_p^2 = .24$ ). After presentation times of 187 ms, the difference in evaluation of untrustworthy-looking ( $M = 4.36, SD = .84$ ) versus trustworthy-looking faces ( $M = 5.15, SD = .85$ ) was more pronounced ( $F(1, 29) = 18.97, p < .001, \eta_p^2 = .40$ ).

As expected, with increasing presentation times, the influence of behavior on face evaluation also increased (Figure 3b), as reflected by a significant interaction between presentation time and behavior ( $F(2, 116) = 6.55, p = .002, \eta_p^2 = .10$ ). After presentation times of only 35 ms, there was an effect of behavior on face evaluation ( $F(2, 58) = 10.58, p < .001, \eta_p^2 = .27$ ). Faces presented with negative behaviors ( $M = 4.81, SD = 1.18$ ) were evaluated significantly more negatively than faces presented without behaviors ( $M = 5.22, SD = .84$ ) or faces presented with positive behaviors ( $M = 5.67, SD = .89; t(29) > 2.57, p < .02, d_{rm} > .37$  for all pairwise comparisons). After presentation times of 187 ms, the influence of behavior on evaluation was stronger ( $F(2, 58) = 43.90, p < .001, \eta_p^2 = .60$ ). Faces presented with negative behaviors ( $M = 3.81, SD = .80$ ) were again evaluated significantly more negatively than faces presented without behaviors ( $M = 4.83, SD = .74$ ) or faces presented with positive behaviors ( $M = 5.63, SD = 1.16; t(29) > 4.99, p < .001, d_{rm} > .75$  for all pairwise comparisons).

For the response deadline manipulation, we were interested in whether the effect of appearance, and especially the effect of learning, would be present under the deadline condition. The response deadline manipulation did not have a strong influence on the effect of appearance (Figure 3c;  $F < 1$  for the interaction between response deadline and appearance). After a short response deadline, untrustworthy-looking faces ( $M = 4.65, SD = 1.01$ ) were rated more negatively than trustworthy faces ( $M = 5.34, SD = .98; F(1, 59) = 22.14, p < .001, \eta_p^2 = .27$ ). After waiting to respond for three seconds, untrustworthy-looking faces ( $M = 4.70, SD = 1.01$ ) were again rated more negatively than trustworthy faces ( $M = 5.28, SD = .99$ ), and this effect was not appreciably different in size ( $F(1, 59) = 19.92, p < .001, \eta_p^2 = .25$ ).

In contrast, increasing the amount of time participants had to respond to faces increased the influence of behavior on face evaluation (Figure 3d; response deadline by behavior:  $F(2, 116) = 6.49, p = .002, \eta_p^2 = .10$ ). After a short response deadline, the influence of behavior on face evaluation was already significant ( $F(2, 118) = 21.99, p < .001, \eta_p^2 = .27$ ). Faces presented with negative behaviors ( $M = 4.49, SD = 1.13$ ) were evaluated significantly more negatively than faces without behaviors ( $M = 4.96, SD = .91$ ) or faces presented with positive behaviors ( $M = 5.54,$



**Figure 3.** Influence of appearance and behavior on face evaluation after brief presentation times and a response deadline ( $N = 60$ ). Mean trustworthiness ratings of untrustworthy- (gray with white dots) and trustworthy-looking faces (white with gray dots), and of faces previously presented with negative behaviors (dark gray), no behaviors (light gray), or positive behaviors (white). Panels (a) and (b) show mean ratings by exposure time. Panels (c) and (d) show mean ratings by whether participants responded within 1,500 ms (deadline) or after 3,000 ms (no deadline) of when a face appeared. Error bars represent *SEM*.

$SD = 1.20$ ;  $t(59) > 3.29$ ,  $p < .002$ ,  $d_{mm} > .44$  for all pairwise comparisons). However, after waiting to respond for at least 3 s, the influence of behavior on evaluation was stronger ( $F(2, 118) = 46.96$ ,  $p < .001$ ,  $\eta_p^2 = .44$ ). Participants evaluated the faces presented with negative behaviors ( $M = 4.13$ ,  $SD = 1.33$ ) significantly more negatively than those presented without behaviors ( $M = 5.09$ ,  $SD = .92$ ) or with positive behaviors ( $M = 5.75$ ,  $SD = 1.15$ ;  $t(59) > 4.68$ ,  $p < .001$ ,  $d_{mm} > .62$  for all pairwise comparisons).

None of the remaining effects reached significance (Appearance  $\times$  Behavior:  $F(1, 116) = 2.55$ ,  $p = .08$ ,  $\eta_p^2 = .04$ ; Deadline  $\times$  Time:  $F(1, 58) = 2.69$ ,  $p = .11$ ,  $\eta_p^2 = .04$ ; Appearance  $\times$  Behavior  $\times$  Deadline:  $F(2, 116) = 1.28$ ,  $p = .28$ ,  $\eta_p^2 = .02$ ; Appearance  $\times$  Behavior  $\times$  Time:  $F(2, 116) = 2.05$ ,  $p = .13$ ,  $\eta_p^2 = .03$ ; Appearance  $\times$  Deadline  $\times$  Time:  $F(1, 58) = 1.04$ ,  $p = .31$ ,  $\eta_p^2 = .02$ ;  $F < 1$  in all other cases).

## Discussion

In the second experiment, we found an effect of behavior on face judgments made within 1,500 ms of a brief exposure to a face, indicating that behavioral associations are readily available after exposure to minimal visual information. Thus, judgments made soon after a brief glance at a person's face already reflect knowledge about that person's past behavior. With an increase in response time from 1,500 to 3,000 ms, we saw an increase in the influence of behavior learning on evaluation, suggesting that while the additional processing time is not necessary for the retrieval of behavioral associations, it does facilitate this retrieval. In contrast,

increasing response time did not increase the influence of appearance on face evaluation. These results are consistent with those of Rudoy and Paller (2009), although they observed a stronger difference between learning and appearance across the response deadline manipulation.

## Experiment 3

In the third experiment, we investigated whether the influence of behavior on face evaluation depends on conscious recognition of the learned faces. Previous work indicates that evaluative responding can occur under a range of conditions indicating automaticity, including the absence of recognition (for discussions on the independence of affect and cognition see Zajonc, 1980, 2000, for a review on evaluative responding see De Houwer, 2009). For example, participants show greater liking for stimuli that they have been previously exposed to even when recognition for those stimuli is at chance (for a review of the role of stimulus recognition in the mere exposure effect see Bornstein, 1989).

In the current study, we were interested in whether learning would continue to influence face evaluation even when participants reported difficulty recognizing the faces. As in the previous experiments, participants viewed faces paired with negative or positive behaviors. Then, participants evaluated the trustworthiness of a series of faces, which included the learned faces intermixed with novel faces. Immediately after evaluating each face, participants indicated whether they recognized the face from the first part of the experiment using a continuous scale. To ensure that each participant would report a range of

recognition responses, we manipulated the exposure time of the faces within participants. If participants need to deliberately retrieve behavioral information for it to influence face evaluation, then behavior should only influence evaluation when participants recognize the learned faces. If, however, behavioral information is retrieved relatively automatically, then we should see an influence of learning even in the absence of recognition.

## Method

**Participants.** Sixty-five undergraduates at Princeton University participated in this experiment for partial course credit (27 female, 38 male;  $M_{\text{age}} = 19.37$ ,  $SD_{\text{age}} = 1.10$ ; primary racial/ethnic background: 6% Black/African American, 22% East Asian/East Asian American, 3% Latino/Hispanic American, 2% Middle Eastern/Arab American, 2% Native/American Indian, 6% South Asian/South Asian American, 54% White/European American, 6% other; 14% of participants reported that they were multiracial). While all of the manipulations in this experiment occurred within participants, we chose to use a sample size comparable to the previous experiment because we planned to run analyses on subsets of the data where participants reported low levels of face recognition.

**Stimuli.** Stimuli consisted of 72 black-and-white digital photographs of bald males. Fifty-four of the photographs were from the face set used in the previous experiments (Goldstone & Steyvers, 2001) and the remaining 18 photographs were taken from Kayser (1997), the source of the original set of faces. Thirty-six of the faces were selected to be used as the learned faces. Because of our specific theoretical focus and also for reasons of statistical power, we chose not to include face appearance (trustworthy vs. untrustworthy) as a factor in this experiment. Therefore, using the pilot ratings of face trustworthiness from Experiment 1, we selected the 36 faces out of the set of 45 rated faces with the least extreme mean ratings of face trustworthiness.

Behaviors were selected from the database of behaviors used in the previous experiments (Fuhrman et al., 1989). As in the previous experiments, the negative and positive behaviors did not differ in their length (negative:  $M_{\text{words}} = 9.28$ ,  $SD_{\text{words}} = 2.00$ ; positive:  $M_{\text{words}} = 9.39$ ,  $SD_{\text{words}} = 2.09$ ;  $t < 1$ ) or in their distance from the midpoint of the goodness rating scale (negative behaviors:  $M_{\text{goodness}} = 1.95$ ,  $SD_{\text{goodness}} = .69$ ; positive behaviors:  $M_{\text{goodness}} = 8.09$ ,  $SD_{\text{goodness}} = .50$ ; independent samples  $t$  test on distance from the midpoint:  $t < 1$ ). Once again, the negative behaviors ( $M = 3.30$ ,  $SD = 1.13$ ) were seen as less typical than the positive behaviors ( $M = 6.25$ ,  $SD = 1.42$ ;  $t(106) = 11.90$ ,  $p < .001$ ,  $d = 2.32$ ).

The 36 faces were divided into two groups equated on mean face trustworthiness, and participants learned to associate each group with negative or positive behaviors. The face and behavior groups were counterbalanced across participants. Stimuli were presented on 17 in. CRT monitors with a refresh rate of 75 Hz.

### Procedures.

**Learning.** The only difference in the learning procedure between this experiment and the previous experiments is that in the current experiment all faces were presented with behaviors.

**Face evaluation and recognition.** Immediately after seeing the faces paired with behaviors, participants were asked to (a) evaluate the trustworthiness of a series of faces presented without behaviors and (b) indicate whether they recognized each face from the first part of the experiment. The instructions for the task were kept as similar as possible to those in the previous experiments, and the blank screen and fixation cross at the start of each trial were shown for the same amount of time as before. The presentation time of the faces was manipulated within participants, such that each participant viewed 24 faces for 27 ms (2 refresh cycles), 24 faces for 40 ms (3 refresh cycles), and 24 faces for 147 ms (11 refresh cycles). For each presentation time, 6 of the faces were previously presented with negative behaviors, 6 were presented with positive behaviors, and 12 were novel faces. The faces shown at each presentation time were counterbalanced across participants. After each face, a phase-scrambled mask was shown for 187 ms. After the mask, participants evaluated each face on trustworthiness as in the previous experiments. Immediately after evaluating each face, participants were asked to indicate whether they recognized it from the first part of the experiment. The word “Recognize?” appeared along with a response scale. The scale remained on the screen until participants responded using the number keys 1 (*not at all*) to 9 (*definitely*). Participants rated each of the 72 faces two times. The faces were blocked such that participants saw each face once before seeing any face a second time. The average correlation between the first and second ratings of the faces was .41 ( $SD = .20$ ). The order of the faces was randomized for each participant.

The experiment had a 2(behavior: negative or positive)  $\times$  3 (time: 27, 40, or 147 ms)  $\times$  2 (judgment: trustworthiness, recognition) repeated measures design. We used a repeated measures ANOVA to examine recognition ratings for learned versus novel faces. Next, we used a linear mixed-effects model to address our main question of interest: whether the influence of behavior on face evaluation depends on conscious recognition of a face. Because most participants did not use all of the values on the recognition scale, there were a large number of missing values and this approach allowed us to cope with the missing data. We performed the linear mixed-effects analysis predicting face evaluation from behavior, presentation time, recognition, and the interactions between these factors using the lme4 package Version 1.1–12 (Bates, Mächler, Bolker, & Walker, 2015) for R (R Core Team, 2015). Behavior and presentation time were treated as categorical predictors and recognition was treated as continuous. We ran the maximal model with participants as random effects with varying intercepts and the factor interactions as varying slopes (Barr, Levy, Scheepers, & Tily, 2013). We obtained  $p$  values using Satterthwaite approximations to degrees of freedom from the package lmerTest Version 2.0–33 (Kuznetsova, Brockhoff, & Christensen, 2016). Planned comparisons and confidence intervals (CIs) were computed with the lsmeans package Version 2.26–3 (Lenth, 2016). The script for this analysis is available on the Open Science Framework (<https://osf.io/9jwv2/>). Finally, we used a paired-samples  $t$  test to examine the influence of behavior on face evaluation for those participants who reported recognition values of one for at least one positive and one negative learned face.



## Results

**Face recognition.** Participants rated the learned faces ( $M = 5.40$ ,  $SD = 1.22$ ) as more familiar than the novel faces ( $M = 3.90$ ,  $SD = 1.19$ ;  $F(1, 64) = 78.22$ ,  $p < .001$ ,  $\eta_p^2 = .55$ ; Figure 4). Participants also rated faces presented for longer times as more familiar ( $F(2, 128) = 85.03$ ,  $p < .001$ ,  $\eta_p^2 = .57$ ). This effect of presentation time was limited to the learned faces (learned:  $F(2, 128) = 49.58$ ,  $p < .001$ ,  $\eta_p^2 = .44$ ; novel:  $F < 1$ ), as reflected by a significant interaction between presentation time and face type ( $F(2, 128) = 126.14$ ,  $p < .001$ ,  $\eta_p^2 = .66$ ). Participants rated the learned faces presented for 27 ms as less familiar than those presented for 40 or 147 ms ( $t(64) > 5.40$ ,  $p < .001$ ,  $d_{rm} > .50$  for all pairwise comparisons). However, participants did not rate the novel faces differently depending on presentation time ( $t(64) < 1.04$ ,  $p > .31$ , and  $d_{rm} < .13$  for all comparisons). A second analysis revealed that recognition did not differ for learned faces associated with negative versus positive behaviors (behavior:  $F(1, 64) = 2.19$ ,  $p = .14$ ,  $\eta_p^2 = .03$ ; Behavior  $\times$  Presentation Time:  $F < 1$ ).

### Face evaluation by face recognition.

**Linear mixed-effects analysis.** To investigate whether the influence of behavior on face evaluation depends on conscious recognition of the learned faces we constructed a linear mixed-effects regression model. In the model, face evaluation was the dependent variable and behavior, presentation time, recognition, and the interactions between these factors were predictors (see Figure 5). The effect of behavior,  $F(1, 49.67) = 16.77$ ,  $p < .001$ , the effect of recognition,  $F(1, 70.81) = 5.16$ ,  $p = .03$ , the two-way interaction between behavior and recognition,  $F(1, 61.37) = 113.29$ ,  $p < .001$ , and the three-way interaction between behavior, recognition, and presentation time,  $F(2, 75.73) = 6.66$ ,  $p = .002$  were all significant. The remaining effects did not reach significance (Behavior  $\times$  Time:  $F(2, 96.13) = 1.07$ ,  $p = .35$ ;  $F < 1$  for both other cases).

As shown in Figure 5, the interaction between behavior and recognition indicated that the differences between faces previously associated with negative and faces previously associated with positive behaviors increased as recognition increased. This was the

case for the 27 ms (positive  $B = .19$ ,  $SE = .04$ ; negative  $B = -.08$ ,  $SE = .04$ ; difference  $B = -.28$ , 95% CI  $[-.38, -.17]$ ,  $t(71.22) = 5.31$ ,  $p < .001$ ), 40 ms (positive  $B = .19$ ,  $SE = .04$ ; negative  $B = -.08$ ,  $SE = .04$ ; difference  $B = -.37$ , 95% CI  $[-.47, -.27]$ ,  $t(75.50) = 7.07$ ,  $p < .001$ ), and 147 ms time conditions (positive  $B = .19$ ,  $SE = .04$ ; negative  $B = -.08$ ,  $SE = .04$ ; difference  $B = -.52$ , 95% CI  $[-.63, -.41]$ ,  $t(53.30) = 9.23$ ,  $p < .001$ ). The three-way interaction indicated that the two-way interaction was augmented by the increase in presentation time. The difference in the valence slopes was larger for the 147 ms condition relative to the 40 ms (difference  $B = .15$ ,  $SE = .07$ , 95% CI  $[.02, .28]$ ,  $t(74.42) = 2.22$ ,  $p = .03$ ) and 27 ms conditions (difference  $B = .25$ ,  $SE = .07$ , 95% CI  $[.11, .38]$ ,  $t(73.48) = 3.62$ ,  $p < .001$ ). However, the difference between 40 and 27 ms presentations was not significant (difference  $B = .10$ ,  $SE = .07$ , 95% CI  $[-.04, .38]$ ,  $t(76.56) = 1.39$ ,  $p = .17$ ), indicating that valence showed the largest divergence across recognition when faces were presented for 147 ms.

To investigate whether conscious recognition of a learned face is necessary for behavior to influence evaluation, we examined the predicted effect of behavior on evaluation at each value of recognition on the nine-point scale. For presentation times of 27 ms, we found that there was a significant effect of behavior starting at a recognition value of only three, such that faces previously associated with negative behaviors were predicted to be evaluated significantly more negatively ( $M = 4.67$ ,  $SE = .12$ ) than those previously presented with positive behaviors ( $M = 5.00$ ,  $SE = .12$ ; difference =  $-.33$ , 95% CI  $[-.56, -.10]$ ,  $t(547.86) = 2.86$ ,  $p = .004$ ). This was also true for presentation times of 40 ms (negative:  $M = 4.69$ ,  $SE = .12$ ; positive:  $M = 5.10$ ,  $SE = .14$ ; difference =  $-.41$ , 95% CI  $[-.68, -.13]$ ,  $t(121.46) = 2.93$ ,  $p = .004$ ). For presentation times of 147 ms, the predicted values were in the expected direction, but the difference between them was only marginally significant (negative:  $M = 4.72$ ,  $SE = .19$ ; positive:  $M = 5.20$ ,  $SE = .17$ ; difference =  $-.47$ , 95% CI  $[-.95, .00]$ ,  $t(49.08) = 2.01$ ,  $p = .05$ ). For a recognition value of four, the predicted effect of behavior was significant across all presentation times ( $t > 5.02$  and  $p < .001$  in all cases).

Interestingly, the model predicted a reversal of the effect of behavior at the lowest level of recognition, such that faces previously associated with negative behaviors were expected to be evaluated more positively than those associated with positive behaviors, although these differences were no more than marginally significant ( $t < 1.74$  and  $p > .08$  in all cases). As previously mentioned, many participants did not use all of the recognition values on the scale and this effect may be due in part to interpolating over those missing values. To explore the effect of behavior on evaluation among those participants who did use a recognition value of one, we preformed the analysis described below.

**Analysis by recognition response.** We compared evaluation of the faces paired with negative versus positive behaviors for only those trials with a recognition response of one. Because splitting the data by presentation time resulted in a large number of missing values, we collapsed across this factor. When we collapsed across time, 42 out of 65 participants (65%) had complete data, meaning that they used a recognition response of one for both negative and positive faces at least one time. Although these participants reported not recognizing the faces at all, they still evaluated faces previously paired with negative

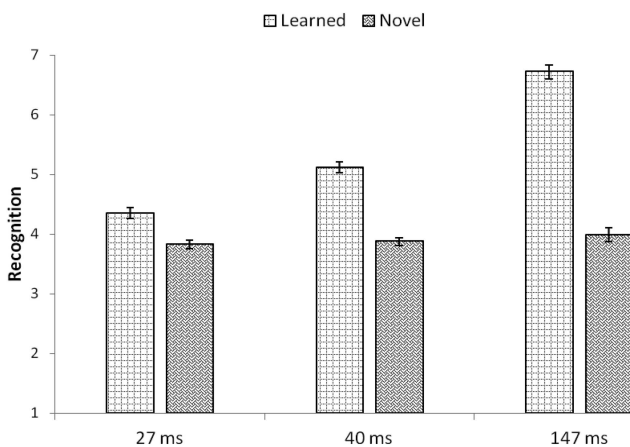


Figure 4. Recognition by exposure time ( $N = 65$ ). Mean recognition ratings of learned (light gray crosshatching) and novel faces (dark gray crosshatching). Error bars represent SEM.

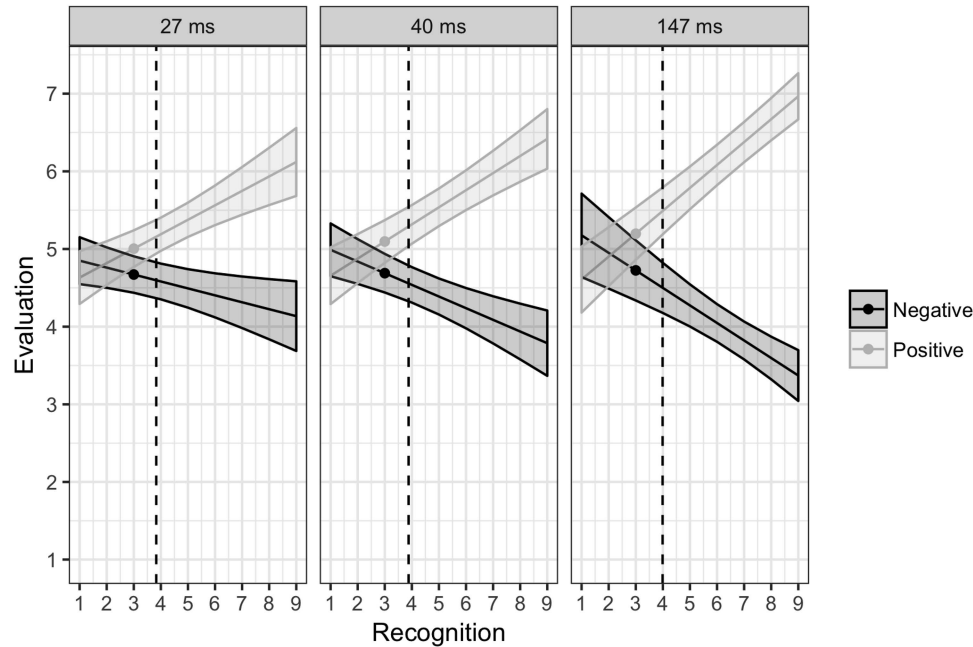


Figure 5. Influence of behavior on face evaluation by recognition and presentation time ( $N = 65$ ). Predicted trustworthiness ratings of faces previously presented with negative behaviors (dark gray) or positive behaviors (light gray). Error bars represent 95% confidence intervals of the regression slope. The dashed vertical lines represent the average recognition ratings for the novel faces and the dots indicate the recognition values where the effect of behavior became significant.

behaviors ( $M = 4.10$ ,  $SD = 1.23$ ) more negatively than those paired with positive behaviors ( $M = 5.07$ ,  $SD = 1.52$ ;  $t(41) = 4.21$ ,  $p < .001$ ,  $d_{rm} = .70$ ).

## Discussion

In the third experiment, we found that the difference in the evaluation of faces previously associated with negative versus positive behaviors increased with increasing recognition. This was true for presentation times of 27, 40, and 147 ms, with the difference in the evaluation of faces as a function of recognition becoming more pronounced for longer presentation times. Not surprisingly, these results indicate that recognizing a face aids participants in retrieving and using relevant behavioral information. However, at the same time, we also found that the influence of learning was already significant at a recognition value of three on a 9-point scale, which is below the average value of recognition for novel faces (see Figure 5). Moreover, across those participants who reported the lowest value of recognition, we still found an effect of behavior on face evaluation. Thus, although recognition magnifies the effect of behavior on face evaluation, only minimal recognition is necessary to retrieve affective associations, suggesting it is possible for such retrieval to occur relatively automatically.

## General Discussion

When forming impressions of other people, perceivers rely on both facial appearance (Todorov et al., 2015) and past behavior (Falvello et al., 2015; Rudoy & Paller, 2009; Todorov & Olson,

2008; Verosky & Todorov, 2010, 2013). Although perceivers integrate these two sources of information in their judgments of other people, they differ in a fundamental way. On the one hand, the information used to make trait judgments from appearance is readily accessible when a perceiver views a face. In contrast, affective associations based on past behavior must be retrieved upon seeing the face in order for them to influence face evaluation. Given this distinction, we were interested in the extent to which use of learned information in face evaluation would require extensive cognitive processing. To investigate this, we examined whether learning would continue to influence face evaluation even under conditions that made the retrieval of associations with the face more difficult.

While previous research has shown that people make trait judgments after minimal exposure to faces, it has focused on how the physical appearance of faces leads to these judgments (Bar et al., 2006; Todorov et al., 2009; Todorov et al., 2010; Willis & Todorov, 2006). In three experiments, we find evidence that a person's past behavior is important for these judgments as well. After presentation times of as little as 27 ms, we found that faces associated with negative behaviors were judged as looking less trustworthy than faces associated with positive behaviors. After the longest presentation time of 187 ms, the influence of learning on judgments increased, but the same was also true for the influence of face trustworthiness. Moreover, when we limited both the exposure to faces and the response time during face evaluation, learning continued to influence evaluation of faces. Finally, even when participants reported difficulty recognizing the learned faces, learning still influenced evaluation of the faces. Thus, although

one possibility was that learning would only exert its influence on judgments of appearance after long presentation and deliberation times, we did not find this to be the case. Instead, learned information seemed to be readily accessible.

Our first experiment suggested that similar amounts of visual information are required for retrieving behavioral associations and making judgments based solely on facial appearance. However, the second experiment suggested that additional time for deliberation after exposure to a face may benefit the retrieval of behavioral associations slightly more than it benefits judgments based on appearance. Our results are in the same direction as those of [Rudoy and Paller \(2009\)](#), although they found a more pronounced difference between memory- and appearance-based effects. While a number of factors could account for the discrepancy between the experiments, it is worth noting that the faces in the current study were presented for a fraction of the time of the faces in [Rudoy and Paller \(2009\)](#). The limited exposure times could have increased the difficulty of making judgments based on facial appearance, thereby minimizing the difference between memory- and appearance-based effects. Another interesting possibility is that our learning procedure, which involved viewing faces paired with behavior descriptions rather than adjectives, facilitated the learning and retrieval of associations.

In the two experiments that compared the influence of learning and appearance, learning had a stronger effect on face evaluation than appearance. Because a person's past behavior is more diagnostic of their underlying character than their appearance, it would be logical to weight information about behavior more heavily in face judgments than information about appearance. Such weighting could explain why the effect of learning was stronger than that of appearance. Alternatively, the stronger learning effect could simply be because of the relative extremity of the behaviors versus the appearance of the faces. In line with this possibility, the negative behaviors were rated below the midpoint of the scale on normality and the positive behaviors were rated only slightly above the midpoint, meaning that both sets of behaviors were likely seen as quite diagnostic of character. Future work could use an experimental design similar to the one used here to test whether more typical behaviors would yield similar results.

Previous work suggests that retrieval of learned information is not always intentional. For instance, reading behavioral descriptions about a person has been found to influence the speed with which participants learn subsequent person information that is congruent or incongruent with what is already known ([Carlston & Skowronski, 1994](#)), to lead to false recognition of traits implied by the behavior when faces are paired with adjectives during a recognition test ([Todorov & Uleman, 2002, 2003, 2004](#)), and to influence judgments of faces that resemble the original faces ([Verosky & Todorov, 2010, 2013](#)). In these studies, retrieval of learned information can be seen as interfering with participants' goals: it slowed learning, led to false recognition of items, and affected judgments of similar looking faces, despite instructions to disregard physical similarity. Because learning can be seen as impairing performance, it seems unlikely that participants intentionally retrieved the learned information. Meanwhile, other work suggests that people do not need to be able to consciously retrieve learned information for its effects to be apparent. Specifically, amnesia patients have been found to acquire affective reactions to faces even in the absence of memory for the biographies previ-

ously associated with them, indicating that conscious awareness is not necessary for learning to influence face evaluation ([Croft et al., 2010](#); [Johnson et al., 1985](#); [Todorov & Olson, 2008](#); [Tranel & Damasio, 1993](#)).

The current work raises the question of how it could be possible to be influenced by behaviors previously associated with a face without recognizing the face. In theory, this could happen either as the result of parallel paths for affect and for face recognition or simply as a result of different thresholds for decisions regarding recognition and affect, such that more evidence is required for recognition. Work with prosopagnosics, who show skin conductance responses to familiar faces in the absence of explicit recognition, and with Capgras syndrome patients, who recognize faces but fail to show such affective responding, provides evidence that affect and face recognition can be separated ([Ellis & Lewis, 2001](#)). However, as [Calder and Young \(2005\)](#) argue with regard to the processing of facial identity and facial expression, some degree of dissociation between two processes does not necessarily mean those processes occur via independent pathways. Many apparent dissociations can be explained by different decision thresholds so that some judgments (e.g., affect-based person judgments) may require less evidence than other judgments (e.g., person recognition).

Regardless of the extent to which affect and recognition are processed separately, learning could exert its influence on face evaluation by behaviors becoming associated with the morphological characteristics of the face. One piece of evidence that learning is tightly tied to appearance is that behaviors are generalized from learned faces to similar-looking faces ([Verosky & Todorov, 2010, 2013](#)). Another piece of evidence is that while people tend to show a high level of agreement in their appearance-based judgments of faces, they also show stable idiosyncratic preferences ([Engell, Haxby, & Todorov, 2007](#); [Hönekopp, 2006](#)). A twin study examining judgments of facial attractiveness suggests that such individual face preferences are not heritable, but rather come from previous experience ([Germine et al., 2015](#)).

A related question is whether the observed learning effects would hold for other types of person learning. The behaviors in the current experiments were selected based on their goodness ratings, and as such they mapped onto the valence dimension of person perception. Because affective responding is thought to be more fundamental than other types of responding ([Zajonc, 1980, 2000](#)), one possibility is that the learning effects we observed would only be present for other behaviors to the extent that those behaviors relate to general negativity/positivity. However, going against this possibility, people have been found to be as adept at forming associations between large numbers of face-behavior pairs for competence-related as for morality-related behaviors ([Falvello et al., 2015](#)). Moreover, the perceived everyday frequency of behaviors has been found to predict behavioral and neural responding during updating of person impressions, regardless of the domain of the behaviors ([Mende-Siedlecki et al., 2013](#)), so that infrequent behaviors (e.g., a great achievement in the ability domain or a dishonest behavior in the morality domain) are weighed more heavily in person evaluation than frequent behaviors.

In conclusion, the current experiments contribute to a view of learning as a powerful influence on face processing. Together, our data suggest that people are able to retrieve the knowledge associated with faces even under circumstances that prevent careful



reflection. We found that learning does not take time to exert its influence on face evaluation, but rather that its effects are apparent after as little as 27 ms face exposure. In addition, the influence of learning on face evaluation was apparent when participants were only provided 1,500 ms to evaluate the faces. Finally, we found an effect of behavior with only minimal face recognition, suggesting that the influence of learning on evaluation does not require deliberate retrieval of the learned information. Thus, our data show that what you know about someone affects the way you feel about their face almost immediately, perhaps even before you recognize them.

## References

- Adams, R. B., Jr., Hess, U., & Kleck, R. E. (2015). The intersection of gender-related facial appearance and facial displays of emotion. *Emotion Review*, 7, 5–13. <http://dx.doi.org/10.1177/1754073914544407>
- 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>
- Ballem, C. C., II, & Todorov, A. (2007). Predicting political elections from rapid and unreflective face judgments. In *Proceedings of the National Academy of Sciences of the United States of America*, 104, 17948–17953. <http://dx.doi.org/10.1073/pnas.0705435104>
- 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>
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255–278. <http://dx.doi.org/10.1016/j.jml.2012.11.001>
- Bascandziev, I., & Harris, P. L. (2014). In beauty we trust: Children prefer information from more attractive informants. *British Journal of Developmental Psychology*, 32, 94–99. <http://dx.doi.org/10.1111/bjdp.12022>
- Bascandziev, I., & Harris, P. L. (2016). The beautiful and the accurate: Are children's selective trust decisions biased? *Journal of Experimental Child Psychology*, 152, 92–105. <http://dx.doi.org/10.1016/j.jecp.2016.06.017>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67, 1–48. <http://dx.doi.org/10.18637/jss.v067.i01>
- Bliss-Moreau, E., Barrett, L. F., & Wright, C. I. (2008). Individual differences in learning the affective value of others under minimal conditions. *Emotion*, 8, 479–493. <http://dx.doi.org/10.1037/1528-3542.8.4.479>
- Bornstein, R. F. (1989). Exposure and affect: Overview and meta-analysis of research, 1968–1987. *Psychological Bulletin*, 106, 265–289. <http://dx.doi.org/10.1037/0033-2909.106.2.265>
- Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience*, 6, 641–651. <http://dx.doi.org/10.1038/nrn1724>
- Carlston, D. E., & Skowronski, J. J. (1994). Saving in the relearning of trait information as evidence for spontaneous inference generation. *Journal of Personality and Social Psychology*, 66, 840–856. <http://dx.doi.org/10.1037/0022-3514.66.5.840>
- Cogsdill, E. J., Todorov, A. T., Spelke, E. S., & Banaji, M. R. (2014). Inferring character from faces. *Psychological Science*, 25, 1132–1139. <http://dx.doi.org/10.1177/0956797614523297>
- Cogsdill, E. J., & Banaji, M. R. (2015). Face-trait inferences show robust child-adult agreement: evidence from three types of faces. *Journal of Experimental Social Psychology*, 60, 150–156. <http://dx.doi.org/10.1016/j.jesp.2015.05.007>
- Croft, K. E., Duff, M. C., Kovach, C. K., Anderson, S. W., Adolphs, R., & Tranel, D. (2010). Detestable or marvelous? Neuroanatomical correlates of character judgments. *Neuropsychologia*, 48, 1789–1801. <http://dx.doi.org/10.1016/j.neuropsychologia.2010.03.001>
- De Houwer, J. (2009). How do people evaluate objects? A brief review. *Social and Personality Psychology Compass*, 3, 36–48. <http://dx.doi.org/10.1111/j.1751-9004.2008.00162.x>
- Ellis, H. D., & Lewis, M. B. (2001). Capgras delusion: A window on face recognition. *Trends in Cognitive Sciences*, 5, 149–156. [http://dx.doi.org/10.1016/S1364-6613\(00\)01620-X](http://dx.doi.org/10.1016/S1364-6613(00)01620-X)
- Engell, A. D., Haxby, J. V., & Todorov, A. (2007). Implicit trustworthiness decisions: Automatic coding of face properties in the human amygdala. *Journal of Cognitive Neuroscience*, 19, 1508–1519. <http://dx.doi.org/10.1162/jocn.2007.19.9.1508>
- Falvello, V. B., Vinson, M., Ferrari, C., & Todorov, A. (2015). The robustness of learning about the trustworthiness of other people. *Social Cognition*, 33, 368–386. <http://dx.doi.org/10.1521/soco.2015.33.5.368>
- Fuhrman, R. W., Bodenhausen, G. V., & Lichtenstein, M. (1989). On the trait implications of social behaviors: Kindness, intelligence, goodness, and normality ratings for 400 behavior statements. *Behavior Research Methods, Instruments & Computers*, 21, 587–597. <http://dx.doi.org/10.3758/BF03210581>
- Germine, L., Russell, R., Bronstad, P. M., Blokland, G. A. M., Smoller, J. W., Kwok, H., . . . Wilmer, J. B. (2015). Individual aesthetic preferences for faces are shaped mostly by environments, not genes. *Current Biology*, 25, 2684–2689. <http://dx.doi.org/10.1016/j.cub.2015.08.048>
- Goldstone, R. L., & Steyvers, M. (2001). The sensitization and differentiation of dimensions during category learning. *Journal of Experimental Psychology: General*, 130, 116–139. <http://dx.doi.org/10.1037/0096-3445.130.1.116>
- Hönekopp, J. (2006). Once more: Is beauty in the eye of the beholder? Relative contributions of private and shared taste to judgments of facial attractiveness. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 199–209. <http://dx.doi.org/10.1037/0096-1523.32.2.199>
- Johnson, M. K., Kim, J. K., & Risse, G. (1985). Do alcoholic Korsakoff's syndrome patients acquire affective reactions? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11, 22–36. <http://dx.doi.org/10.1037/0278-7393.11.1.22>
- Kayser, A. (1997). *Heads*. New York, NY: Abbeville Press.
- Koenig, M. A., Clément, F., & Harris, P. L. (2004). Trust in testimony: Children's use of true and false statements. *Psychological Science*, 15, 694–698. <http://dx.doi.org/10.1111/j.0956-7976.2004.00742.x>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). *lmerTest: Tests in linear mixed effects models*. R package version 2.0–33. Retrieved from <http://CRAN.R-project.org/package=lmerTest>
- Lenth, R. V. (2016). Least-squares means: The R package lsmeans. *Journal of Statistical Software*, 69, 1–33. <http://dx.doi.org/10.18637/jss.v069.i01>
- Mende-Siedlecki, P., Baron, S. G., & Todorov, A. (2013). Diagnostic value underlies asymmetric updating of impressions in the morality and ability domains. *The Journal of Neuroscience*, 33, 19406–19415. <http://dx.doi.org/10.1523/JNEUROSCI.2334-13.2013>
- Montepare, J. M., & Dobish, H. (2003). The contribution of emotion perceptions and their overgeneralizations to trait impressions. *Journal of Nonverbal Behavior*, 27, 236–254. <http://dx.doi.org/10.1023/A:1027332800296>
- 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>
- Oosterhof, N. N., & Todorov, A. (2009). Shared perceptual basis of emotional expressions and trustworthiness impressions from faces. *Emotion*, 9, 128–133. <http://dx.doi.org/10.1037/a0014520>
- R Core Team. (2015). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <https://www.R-project.org/>



- Rudoy, J. D., & Paller, K. A. (2009). Who can you trust? Behavioral and neural differences between perceptual and memory-based influences. *Frontiers in Human Neuroscience*, 3, 16. <http://dx.doi.org/10.3389/neuro.09.016.2009>
- Rule, N. O., Ambady, N., & Adams, R. B., Jr. (2009). Personality in perspective: Judgmental consistency across orientations of the face. *Perception*, 38, 1688–1699. <http://dx.doi.org/10.1068/p6384>
- 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>
- Skowronski, J. J., & Carlston, D. E. (1987). Social judgment and social memory: The role of cue diagnosticity in negativity, positivity, and extremity biases. *Journal of Personality and Social Psychology*, 52, 689–699. <http://dx.doi.org/10.1037/0022-3514.52.4.689>
- Skowronski, J. J., & Carlston, D. E. (1989). Negativity and extremity biases in impression formation: A review of explanations. *Psychological Bulletin*, 105, 131–142. <http://dx.doi.org/10.1037/0033-2909.105.1.131>
- Sutherland, C. A. M., Oldmeadow, J. A., Santos, I. M., Towler, J., Michael Burt, D., & Young, A. W. (2013). Social inferences from faces: Ambient images generate a three-dimensional model. *Cognition*, 127, 105–118. <http://dx.doi.org/10.1016/j.cognition.2012.12.001>
- Todorov, A. (2017). *Face value: The irresistible influence of first impressions*. Princeton, Oxford: Princeton University Press.
- Todorov, A., Loehr, V., & Oosterhof, N. N. (2010). The obligatory nature of holistic processing of faces in social judgments. *Perception*, 39, 514–532. <http://dx.doi.org/10.1068/p6501>
- Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology*, 66, 519–545. <http://dx.doi.org/10.1146/annurev-psych-113011-143831>
- Todorov, A., & Olson, I. R. (2008). Robust learning of affective trait associations with faces when the hippocampus is damaged, but not when the amygdala and temporal pole are damaged. *Social Cognitive and Affective Neuroscience*, 3, 195–203. <http://dx.doi.org/10.1093/scan/nsn013>
- 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., & Uleman, J. S. (2002). Spontaneous trait inferences are bound to actors' faces: Evidence from a false recognition paradigm. *Journal of Personality and Social Psychology*, 83, 1051–1065. <http://dx.doi.org/10.1037/0022-3514.83.5.1051>
- Todorov, A., & Uleman, J. S. (2003). The efficiency of binding spontaneous trait inferences to actors' faces. *Journal of Experimental Social Psychology*, 39, 549–562. [http://dx.doi.org/10.1016/S0022-1031\(03\)00059-3](http://dx.doi.org/10.1016/S0022-1031(03)00059-3)
- Todorov, A., & Uleman, J. S. (2004). The person reference process in spontaneous trait inferences. *Journal of Personality and Social Psychology*, 87, 482–493. <http://dx.doi.org/10.1037/0022-3514.87.4.482>
- Tranel, D., & Damasio, A. R. (1993). The covert learning of affective valence does not require structures in hippocampal system or amygdala. *Journal of Cognitive Neuroscience*, 5, 79–88. <http://dx.doi.org/10.1162/jocn.1993.5.1.79>
- Verosky, S. C., & Todorov, A. (2010). Generalization of affective learning about faces to perceptually similar faces. *Psychological Science*, 21, 779–785. <http://dx.doi.org/10.1177/0956797610371965>
- Verosky, S. C., & Todorov, A. (2013). When physical similarity matters: Mechanisms underlying affective learning generalization to the evaluation of novel faces. *Journal of Experimental Social Psychology*, 49, 661–669. <http://dx.doi.org/10.1016/j.jesp.2013.02.004>
- Walker, M., & Vetter, T. (2009). Portraits made to measure: Manipulating social judgments about individuals with a statistical face model. *Journal of Vision*, 9, 12. <http://dx.doi.org/10.1167/9.11.12>
- Walker, M., & Vetter, T. (2016). Changing the personality of a face: Perceived Big Two and Big Five personality factors modeled in real photographs. *Journal of Personality and Social Psychology*, 110, 609–624. <http://dx.doi.org/10.1037/pspp0000064>
- 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>
- Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35, 151–175. <http://dx.doi.org/10.1037/0003-066X.35.2.151>
- Zajonc, R. B. (2000). Feeling and thinking: Closing the debate over the independence of affect. In J. P. Forgas (Ed.), *Feeling and thinking: The role of affect in social cognition* (pp. 31–58). Cambridge, England: Cambridge University Press.
- Zebrowitz, L. A. (2004). The origin of first impressions. *Journal of Cultural and Evolutionary Psychology*, 2, 93–108. <http://dx.doi.org/10.1556/JCEP.2.2004.1-2.6>
- 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](http://dx.doi.org/10.1207/S15327957PSPR0703_01)
- Zebrowitz, L. A., Kikuchi, M., & Fellous, J. M. (2010). Facial resemblance to emotions: Group differences, impression effects, and race stereotypes. *Journal of Personality and Social Psychology*, 98, 175–189. <http://dx.doi.org/10.1037/a0017990>

Received March 29, 2017

Revision received September 6, 2017

Accepted October 22, 2017 ■