Stereotype Associations and Emotion Recognition

Gijsbert Bijlstra, Rob W. Holland, Ron Dotsch, Kurt Hugenberg and Daniel H. J. Wigboldus

Pers Soc Psychol Bull published online 12 February 2014
DOI: 10.1177/0146167213520458

Personality and Social Psychology Bulletin
1–11

DOI: 10.1177/0146167213520458
pspb.sagepub.com
SAGE

Stereotype Associations and Emotion Recognition

Gijsbert Bijlstra[1], Rob W. Holland[1], Ron Dotsch[1], Kurt Hugenberg[2], and Daniel H. J. Wigboldus[1]

Abstract

We investigated whether stereotype associations between specific emotional expressions and social categories underlie stereotypic emotion recognition biases. Across two studies, we replicated previously documented stereotype biases in emotion recognition using both dynamic (Study 1) and static (Study 2) expression displays. Stereotype consistent expressions were more quickly decoded than stereotype inconsistent expression on Moroccan and White male faces. Importantly, we found consistent and novel evidence that participants' associations between ethnicities and emotions, as measured with a newly developed emotion Implicit Association Test (eIAT), predicted the strength of their ethnicity-based stereotype biases in expression recognition. In both studies, as perceivers' level of Moroccan-anger and Dutch-sadness associations (compared with the opposite) increased, so did perceivers' tendency to decode anger more readily on Moroccan faces and sadness on White faces. The observed stereotype effect seemed to be independent of implicit prejudice (Study 2), suggesting dissociable effects of prejudices and stereotypes in expression perception.

Keywords

emotion, stereotypes, evaluations, implicit associations

Received April 26, 2013; revision accepted December 19, 2013

To enable smooth social communication, it is crucial to quickly and accurately decode emotional expressions of others. Such communication is inherently dyadic—that is, it involves two people: an expresser and a perceiver. Expression perception can be influenced by expressers' facial characteristics, such as the type and intensity of expression, or the facial structure of the expresser (e.g., Becker, Kenrick, Neuberg, Blackwell, & Smith, 2007; Hess, Blairy, & Kleck, 1997; Leppänen, Kauppinen, Peltola, & Hietanen, 2007; Marsh, Adams, & Kleck, 2005; Sacco & Hugenberg, 2009; Zebrowitz, Kikuchi, & Fellous, 2010; see also Said, Sebe, & Todorov, 2009). However, perceivers' characteristics, such as perceivers' expectancies, stereotypes, or prejudices can also influence categorization of emotional expressions (e.g., Bijlstra, Holland, & Wigboldus, 2010; Hugenberg, 2005; Hugenberg & Bodenhausen, 2003; Hutchings & Haddock, 2008; Kang & Chasteen, 2009).

Here, we expand on existing evidence concerning how perceivers' stereotypes affect expression recognition. We aim to provide direct evidence for the influence of perceivers' stereotypic associations in decoding statically and dynamically displayed emotional expressions of Moroccan and White males.[1] To that end, we first briefly discuss the existing evidence of the effects of perceivers' stereotypes on expression recognition, and then present two studies where we find that stereotype associations between specific emotional expressions and social categories underlie stereotypic emotion recognition effects.

Perceivers' Associations and Emotion Recognition

Prior research on social categorization and emotion recognition has found evidence for an apparent stereotypic bias in the perception and categorization of facial expressions (e.g., Bijlstra et al., 2010; Hugenberg, 2005; Hugenberg & Bodenhausen, 2003; Kang & Chasteen, 2009). In an initial demonstration of this phenomenon, Hugenberg and Bodenhausen (2003) demonstrated that the race of a face interacts with perceivers' level of prejudice to influence the perception of facial anger. In this research, participants

[1]Behavioural Science Institute, Radboud University Nijmegen, The Netherlands
[2]Miami University, Oxford, OH, USA

Corresponding Author:
Gijsbert Bijlstra, Radboud University Nijmegen, P. O. Box 9104, 6500 HE Nijmegen, The Netherlands.
Email: g.bijlstra@bsi.ru.nl

watched Black and White faces morph from angry to happy expressions (Study 1) or neutral to angry expressions (Study 2). In both studies, as perceivers' level of prejudice increased, so did perceivers' tendency to see anger more readily in Black but not White faces. High-prejudice perceivers saw anger lingering for longer and appearing earlier on Black faces. However, perceivers' level of prejudice had no influence on the perception of anger in matched White faces. Subsequent work extended this Black-anger link to response latencies as well—White perceivers recognize anger more quickly on Black than White faces (Hugenberg, 2005). Despite the potential interest value of these findings, these initial studies do not clarify whether these results occur because of stereotypes (e.g., stereotypic associations between Black and anger) or prejudice (e.g., negative evaluations of Blacks or outgroups more generally; Dunham, 2011).

This is an important question because research has demonstrated that stereotypes and prejudice are often related but empirically distinct constructs (e.g., Amodio & Devine, 2006; Dotsch, Wigboldus, & van Knippenberg, 2011). For example, take the demonstration of a Black-anger link in expression recognition. Anger signals both globally evaluative and more discrete emotional information; anger can simply be interpreted as a negative emotion but also informs perceivers about the expresser's specific internal state. Consequently, it is ambiguous whether the Black-anger link is caused by specific emotion-group stereotypes or by broader negative prejudicial evaluations of Blacks (or outgroups more generally). Differentiating between these two underlying processes—stereotypes versus prejudice—is important for multiple reasons. First, this distinction is important because they have different behavioral implications. Discrete emotional expressions elicit specific behavioral tendencies (e.g., Cottrell & Neuberg, 2005). For example, although anger and sadness both hold a negative valence, an expression of anger often signals expressers' intent to approach and dominate, whereas for an expression of sadness, this is less likely.

Second, Amodio and Devine (2006) have recently emphasized the importance of distinguishing implicit stereotyping from prejudice because they predict different biases. In three studies, they provided evidence for a dissociation between implicit stereotyping and prejudice and their independent consequences. Stereotypic associations predicted stereotypic biases in participants' judgments. Participants who held strong implicit stereotypes of African Americans, for example, ascribed traits that were associated with Black stereotypes more strongly in these judgments. In contrast, implicit prejudice scores were related to interpersonal behavior, such as seating distance. Seating distance between participants and a chair with belongings ostensibly to an African American coparticipant increased as implicit prejudice increased.

More specific to the current question, stereotyping and prejudice also appear to affect emotion recognition independently. Bijlstra and colleagues (2010) recently provided indirect evidence for this idea. Bijlstra and colleagues found that biases in expression recognition were dependent on the specific comparative context in which emotional expressions were presented. When valence was made salient by the experimental context, prejudice-congruency effects were observed in expression recognition. In contrast, when valence was not made salient by the experimental context, evidence for stereotype-congruency effects was found.

Perceivers in the Bijlstra et al. (2010) studies completed an expression recognition task in which members of two different social categories (Moroccan and White men, Study 1; men and women, Study 2) displayed emotional expressions. The specific expressions displayed differed between conditions. In a so-called dual-valence comparative context, perceivers categorized a positive expression (happiness) and a negative expression (anger *or* sadness), whereas in a single-valence comparative context, perceivers categorized only negatively valenced expressions (anger and sadness). In line with Hugenberg and Bodenhausen (2003; see also Hugenberg, 2005; Hugenberg & Sczesny, 2006; Kang & Chasteen, 2009), results consistent with an implicit prejudice perspective were found in a dual-valence comparative context. An interaction pattern between valence of expressions and social category indicated that positive expressions were more quickly categorized on the relatively positively evaluated social category in the Netherlands (White men), whereas negative expressions were more quickly categorized on the relatively negatively evaluated social category in the Netherlands (Moroccan men, Study 1). In addition, novel evidence was provided for stereotype-congruency effects when participants completed the single-valence (sad vs. angry) expression recognition task. For example, in line with common cultural stereotypes about men and women (e.g., Plant, Hyde, Keltner, & Devine, 2000), the negative emotional expression of anger was recognized faster on male than female faces, whereas sadness was more quickly recognized on female than male faces (Study 2).

In sum, Bijlstra and colleagues (2010) provided evidence that biases in emotion recognition are context dependent—prejudice effects appear to occur when valence is salient, but stereotyping effects appear to occur when valence is not salient. This advance was important because it provides a new means by which to understand previous category-expression effects. For example, it strongly suggests that prejudice-congruency is a better explanation than stereotype-congruency for Hugenberg and Bodenhausen's (2003) and related findings. That is, the evaluative component of social categorization, or more specifically implicit prejudice, significantly predicted the perception of facial anger on Black and not White faces in, what we have called, a dual-valence condition, thereby showing direct evidence of implicit prejudice on decoding anger when valence is made salient. This is consistent with research showing that an evaluative context facilitates the processing of evaluative judgments (e.g., Spruyt, De Houwer, Hermans, & Eelen, 2007). Despite the

importance of knowing that stereotype- and prejudice-congruence effects should be treated as distinct effects, as yet, no direct evidence for stereotype associations underlying stereotype-congruency effects in emotion perception has been documented. In other words, Bijlstra and colleagues only inferred indirectly from their results that prejudice and stereotype associations underlie the demonstrated biases in expression recognition, but did not provide direct evidence for this supposed underlying mechanism.

Present Research

The present research sought to address two open questions in the research on the context dependency of stereotype associations in expression recognition. Although Bijlstra and colleagues (2010) provided preliminary evidence for the context dependency of prejudice versus stereotypic biases in expression decoding, this case is far from complete. In two studies, we investigated whether the previously observed context dependency of stereotypic biases in emotion recognition (a) is directly related to perceivers' strength of stereotype associations and (b) generalizes to dynamic facial expressions.

The primary goal of the current research is to provide direct evidence for effects of stereotypes on expression recognition biases. Past research has only indirectly demonstrated this by inferring the differential influence of group-based stereotypes and prejudice from participants' differential responses to faces based on context (Bijlstra et al., 2010). By using an individual difference measure of perceiver prejudice, Hugenberg and Bodenhausen (2003) have provided direct evidence for prejudice-congruency effects on expression recognition (Hugenberg & Bodenhausen, 2003). However, to date, no one has yet demonstrated direct relationships between individual differences in stereotypes of target groups and purported stereotype-based biases in expression recognition. If emotion recognition in a single-valence (sad vs. angry) comparative context is influenced by stereotypes (Bijlstra et al., 2010), the effect should depend on the strength of perceivers' stereotypic associations. Stronger stereotypic associations among perceivers should lead to stronger stereotype-congruency effects in expression categorization biases.

A secondary goal of the current research is to demonstrate that the predicted effects generalize to both static and dynamic faces. Indeed, past research on the context dependency of stereotypic biases in expression recognition has been conducted exclusively with static faces. There is increasing evidence that the time course of facial expressions plays a potent role in decoding of and responding to expressions (e.g., Recio, Sommer, & Schacht, 2011; Sato, Fujimura, & Suzuki, 2008; Sato, Kochiyama, Yoshikawa, Naito, & Matsumura, 2004), and that dynamic expressions are simply more ecologically valid (see Ishii, Miyamoto, Niedenthal, & Mayama, 2011; Niedenthal, Halberstadt, Margolin, & Innes-Ker, 2000). To this end, Study 1 used a modified morph movies task (see Hugenberg & Bodenhausen, 2003; Niedenthal et al., 2000) where perceivers viewed and responded to dynamic expressions (i.e., expressions changing over time from one expression to another). Using this new technique, within a single-valence (sad vs. angry) comparative context adapted from Bijlstra and colleagues (2010), Study 1 aims to provide novel evidence for stereotype application in recognizing emotions from dynamic facial displays.

In both of the current studies, we address this troubling gap in the literature, providing novel evidence that individual differences in stereotypic associations predict the strength of the stereotypic bias in emotion recognition. Furthermore, Study 2 is aimed at showing that individual differences in stereotypic associations affect expression recognition in different ways from individual differences in implicit prejudice.

Study 1

In Study 1, we sought to provide direct evidence for the earlier indirectly inferred effects of stereotypes on expression recognition (Bijlstra et al., 2010). To this end, participants completed a morph movies task (Niedenthal et al., 2000) wherein expressions on both Moroccan and White male faces changed over time between sad and angry. This sad-to-angry dimension was selected for two reasons. First, it creates a single-valence comparison in which valence remains negative, but the stereotypicality of the expressions can change. Second, because for the Dutch participants, anger is stereotypical for Moroccans, but sadness is not (Bijlstra et al., 2010), this creates an experimental context in which stereotyping effects are most likely to occur. Thus, the (negative) valence is held constant within each trial but the expressions change between a Moroccan stereotypic expression (anger) and a non-stereotypic expression (sadness). A pilot experiment in our lab revealed initial evidence that previously found prejudice- and stereotype-congruency effects could be replicated using a morph movies task.[2]

After participants finished the morph movies task, they also completed a newly developed Implicit Association Test (IAT; Greenwald, McGhee, & Schwartz, 1998) to assess their implicit stereotype associations of Dutch[3] versus Moroccan men with discrete emotions (anger; sadness). Originally the IAT was designed to index the association between target concepts (e.g., two social categories) and an attribute dimension (e.g., positive versus negative; Greenwald et al., 1998). In our new emotion IAT (eIAT), we assessed participants' individual differences in associative strengths between social categories (Moroccan and Dutch) and two discrete emotional expressions (anger and sadness). Thus, we replaced the evaluative dimension of the original IAT with the discrete emotional expression attributes of anger and sadness. Note that these discrete emotional expressions are both negatively valenced, and therefore, responses are not likely to be strongly influenced by general evaluative associations (see Amodio & Devine, 2006).

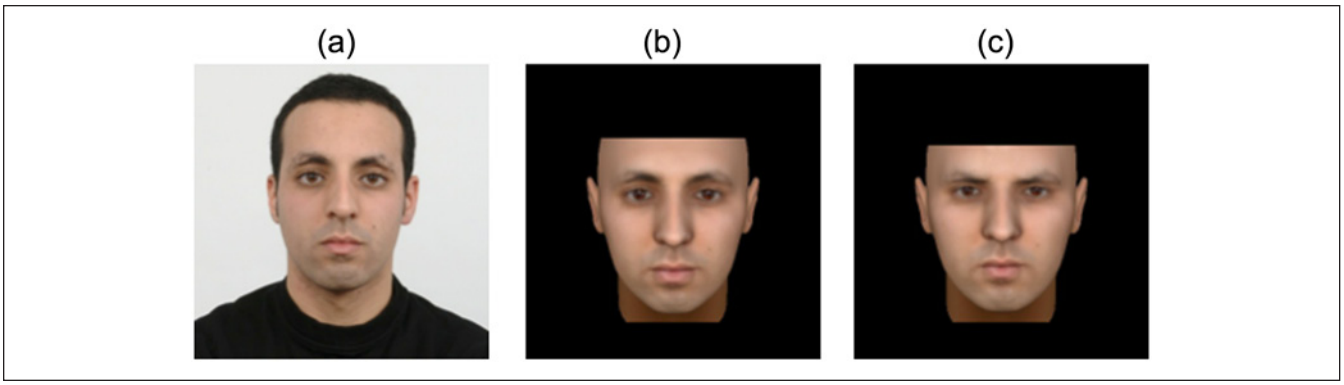

Figure 1. (a) An original Moroccan RaFD image, (b) the same neutral face in FaceGen, and (c) the created angry version of this face.
Note. RaFD = Radboud Faces Database.

Of particular interest in the current study was whether individual differences in stereotype associations predicted stereotypic emotion recognition biases. Specifically, participants' implicit stereotype associations should be reflected in participants' responses in our eIAT measurement, with greater eIAT scores reflecting stronger anger-Moroccan and sadness-Dutch (relative to sadness-Moroccan and anger-Dutch) associations. More importantly, we hypothesized that these stereotype associations would underlie emotion recognition effects obtained in previous research (Bijlstra et al., 2010). Extending this logic, we predicted that the eIAT should predict the emotion recognition stereotype effect in this emotion categorization task. That is, participants with stronger anger-Moroccan and sadness-Dutch associations should recognize anger more readily on Moroccan than White faces, and sadness more readily on White than Moroccan faces.

Method

Participants and overview. A total of 103 Radboud University Nijmegen students (83 women, M_{age} = 22) completed a single-valence (sad vs. angry) emotion morph movies task in which participants watched short film clips of Moroccan and White male faces changing from anger (sadness) into sadness (anger). Subsequently, participants completed an implicit association test to measure their discrete emotion associations (eIAT) with the two social categories central in this study.

Materials and procedure

Morph movies task. We created short film clips in which White and Moroccan male faces gradually change from one expression to another expression. To create these clips, we used FaceGen (Singular Inversions, www.facegen.com) to first import neutral expression faces (using FaceGen's PhotoFit procedure) and subsequently to create sad and angry expressions for all faces. We imported the neutral expression faces of 10 models, 5 White and 5 Moroccan males, from the Radboud Faces Database (RaFD; Langner et al., 2010). Using FaceGen, we were able to systematically alter emotional faces by changing exactly the same parameters for each face (see Figure 1) for an original Moroccan RaFD image (a), the same neutral face in FaceGen (b), and the angry face we created from this (c).

Next, for every model, we created film clips by morphing two emotional faces using MorphX (Norrkross, www.norrkross.com) with a duration of 8 s per clip, each containing 240 frames. We created 2 film clips (sad-to-angry; angry-to-sad) for each of the 10 models (5 White; 5 Moroccan), for a total of 20 film clips. In addition, two film clips were created and used to familiarize participants with the experimental task in a preliminary practice block.

Participants were seated in individual cubicles and informed that they would be presented with short film clips of faces displaying an emotional expression that changed into a second emotional expression. This morph movies task was adapted from Niedenthal and colleagues (2000) and Hugenberg and Bodenhausen (2003). We instructed participants to watch each film clip and press spacebar the moment they detected the onset of a new expression on a face. The experiment consisted of a 2-trial practice block, followed by two 20-trial test blocks. There was a 1 s intertrial interval. Each clip was shown once per test block; the presentation order of film clips was randomized for each participant.[4]

Emotion implicit association test. On finishing the morph movies task, participants completed the eIAT. In the eIAT, participants were asked to categorize words into four distinct categories: Moroccan names, Dutch names, and words associated with sadness or anger. This eIAT was designed to measure the strength of an individual's association between these two social categories and the two discrete emotional expressions.

In total, the eIAT consisted of five blocks: three practice and two test blocks. First participants performed two 20-trial practice blocks to familiarize the response mapping; in the first block, they categorized male names typical for the social

categories, Moroccan and Dutch—five Moroccan (e.g., Ibrahim, Achmed) and five Dutch (e.g., Johan, Karel) names—by pressing one of two response keys (A—left and 6—right). In a second practice block, participants categorized emotion words associated with sadness (*verdriet*) and anger (*boos*). The five words we used for the sadness category were crying (*huilen*), weeping (*wenen*), sadly (*droevig*), sad (*treurig*), and tears (*tranen*); the five words we used for the anger category were furious (*furieus*), upset (*kwaad*), outrageous (*laaiend*), rage (*woest*), and mad (*razend*).

After finishing the two practice blocks, the four categories were combined in the first 40-trial test block. For all participants, this led first to a block in which they categorized anger words and Moroccan names by pressing one key and sadness words and Dutch names by pressing the other key (stereotype-congruent block). Participants then completed a social category practice block in which the social category labels switched sides (no emotion words were presented in this block). Subsequently, all participants categorized anger words and Dutch names with one response key and sadness words and Moroccan names with the other response key (stereotype-incongruent block). Shorter response latencies in the congruent block than the incongruent block were assumed to reflect stronger automatic associations between anger with Moroccans and sadness with the Dutch social category compared with the opposite emotion-ethnicity combinations.

In this study, the eIAT was used as individual difference measures for predicting participants' performance on the morph movies task. To minimize between-participant differences in IAT scores due to order effects, blocks as well as words were presented in a fixed, pseudo-random order.

Results

The primary dependent variable in the present study was the average response time to detect the onset of the second emotional expression in the film clips. Before analyzing, data of 13 participants were excluded because the median of their response latencies on the emotion morph movies task fell outside the duration of the film clips (>8 s). Furthermore, for each of the remaining participants, we excluded response latencies above 8 s (20.6%) and those trials differing more than three standard deviations from the overall mean (0.6%).[5] Three more participants were excluded from the analyses because of an eIAT score that differed more than three *SD*s from the overall mean eIAT score.

Emotion implicit association test. Using the algorithm of Greenwald et al. (1998), we calculated a measure of associative strength between social categories and discrete emotional expressions (eIAT score) separately for each participant. Incorrect responses (7.9%), practice blocks, and the first two trials of all test blocks were excluded from analysis; response latencies smaller than 300 ms were recoded to 300 ms and those bigger than 3,000 to 3,000 ms. Analyses were performed on log-transformed response latencies, but untransformed latencies are reported for ease of interpretation.

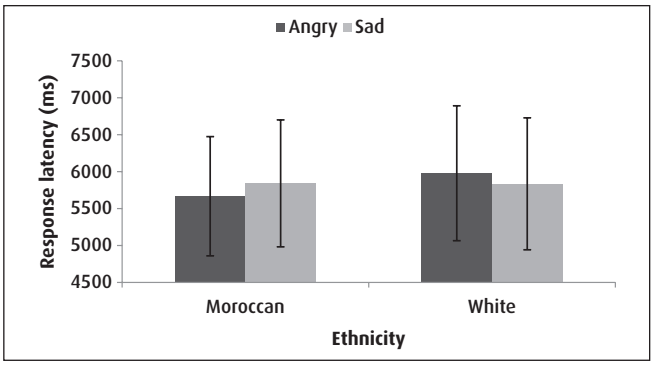

Figure 2. Mean response latencies and standard deviations in milliseconds as a function of ethnicity and expression measured with the morph movies task.

In the eIAT, people generally responded faster during the stereotype-congruent block ($M = 589$, $SD = 87$) than during the stereotype-incongruent block ($M = 666$, $SD = 109$), $F(1, 86) = 130.00$, $p < .001$, $\eta_p^2 = .60$. This indicates that, on average, our participants associated anger more with Moroccan and sadness more with Dutch males than sadness with Moroccan and anger with Dutch males.

Morph movies task. Because the design of the current study involved a conceptual replication of the single-valence condition of Bijlstra et al. (2010), we predicted a replication of the Ethnicity × Expression interaction previously observed in a single-valence condition. Of primary interest, however, was whether the eIAT scores predicted the magnitude of the ethnicity-based bias in expression recognition.

To investigate this question, we subjected the mean response latencies to a 2 (Ethnicity: White vs. Moroccan) × 2 (Expression: angry vs. sad) repeated-measures ANOVA with standardized eIAT scores as continuous factor. This analysis revealed a main effect of Ethnicity, $F(1, 85) = 12.82$, $p = .001$, $\eta_p^2 = .13$. Emotional expressions were detected more quickly on Moroccan ($M = 5{,}747$, $SD = 769$) than White faces ($M = 5{,}900$, $SD = 835$). Furthermore, the expected two-way interaction between Ethnicity and Expression was significant, $F(1, 85) = 18.99$, $p < .001$, $\eta_p^2 = .18$ (see Figure 2). In line with Bijlstra and colleagues (2010), we found that anger was detected more quickly on Moroccan ($M = 5{,}667$, $SD = 807$) than White faces ($M = 5{,}978$, $SD = 913$), $F(1, 85) = 25.31$, $p < .001$, $\eta_p^2 = .23$. No difference in detecting sadness on the two ethnicities was found ($F < 1$). Comparing response latencies of detecting Expressions within the Ethnicity groups, we found that sadness ($M = 5{,}835$, $SD = 893$) was detected more quickly than anger on White faces, $F(1, 85) = 4.62$, $p = .03$, $\eta_p^2 = .05$. In contrast, anger was detected more easily than sadness ($M = 5{,}841$, $SD = 859$) on Moroccan faces, $F(1, 85) = 6.78$, $p = .011$, $\eta_p^2 = .07$.

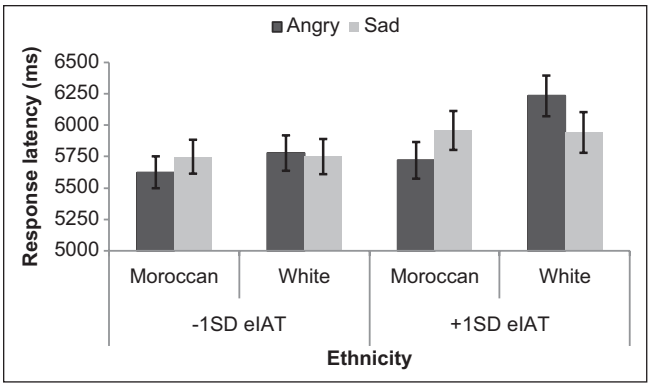

Figure 3. The relationship between participants' eIAT score and their stereotype bias in emotion recognition, represented by estimated response latencies and standard error for −1 *SD* and +1 *SD* from the eIAT mean.
Note. eIAT = emotion Implicit Association Test.

Most importantly, the predicted three-way interaction between Ethnicity, Expression, and eIAT score was also observed, $F(1, 85) = 4.13$, $p = .045$, $\eta_p^2 = .05$ (see Figure 3). The stronger the participants associated anger with Moroccan and sadness with Dutch men, compared with the opposite, the stronger was their stereotype effect in detecting emotional expressions.[6]

Discussion

The results of Study 1 conceptually replicated the stereotype effect of Bijlstra and colleagues (2010), but using dynamic facial stimuli instead of the static emotional face paradigms used thus far. More importantly, we provided novel evidence that this effect was correlated with individuals' associations between discrete emotional expressions and social categories. To our knowledge, this is the first direct evidence for stereotype associations predicting the recognition of emotional expressions: The newly developed eIAT significantly predicted participants' stereotype effect. Prior research (e.g., Hugenberg & Bodenhausen, 2003, 2004; Hutchings & Haddock, 2008), using a dual-valence context, has already demonstrated that perceivers' prejudice, as measured with IATs, can bias face processing. However, the current study expands on previous research by demonstrating that specific stereotypic content, rather than prejudice, can directly bias expression perception processing in a single-valence context.

Although Study 1 provides confirmatory evidence for a direct stereotype-to-expression link (in the single-valence context), it does not provide a disconfirmatory test. Indeed, from our perspective, whereas stereotypes should predict the stereotype effect within a single-valence context, prejudice should not predict this effect. In short, if this is a stereotyping and not a prejudice effect, this inference would be made much more strongly by including a disconfirmatory test as well. Study 2 was designed to address this question by including a measure of prejudice along with a measure of stereotypes. Thus, to see whether we could replicate the initial evidence of Study 1 and to test the discriminant validity of the eIAT compared with a traditional evaluative IAT, we included both measurements in our second study.

Study 2

In the present study, we used a speeded categorization task wherein participants saw a series of still-frame angry and sad Moroccan and White male faces (again, a single-valence condition, Bijlstra et al., 2010), and were tasked with categorizing each face by expression as quickly as possible. By selecting this task, we were able to test whether the results of Study 1 generalize to paradigms other than the morph movies task and to what extent the results were specific for the used dynamic face context. After this task, participants completed both the eIAT used in Study 1, as well as the traditional evaluative IAT, which was used as an index of overall evaluation of Moroccan versus Dutch men (i.e., prejudice). The goals of Study 2 were to replicate and extend the findings that individual differences in emotion stereotypes (eIAT) predict the biasing effects of stereotypes on expression recognition, even controlling for implicit prejudice (IAT).

Method

Participants and overview. A total of 74 Radboud University Nijmegen students (59 women, $M_{age} = 22$) participated in an emotion categorization task and two implicit association tests. First, during the emotion categorization task, participants classified sad and angry emotional expressions displayed by both White and Moroccan males. Subsequently, participants completed two implicit association tests to measure their implicit stereotype associations (eIAT) and implicit prejudice (IAT) with the two social categories central in this study. The eIAT used in this experiment was identical to the eIAT used in Study 1.

Materials and procedure

Pictures. A total of 36 male face models, 18 White and 18 Moroccan, were selected from the RaFD (Langner et al., 2010). Of these models, we used the pictures of two emotional expressions: sad and angry. This resulted in a total of 72 pictures: 2 expressions × 2 ethnicities × 18 models per ethnicity.

Emotion categorization task. The emotion categorization task was adapted from Bijlstra et al. (2010). Participants were asked to categorize emotional expressions on angry and sad faces presented on a computer screen as quickly and accurately as possible. Each trial consisted of a fixation cross, presented for 1,000 ms, followed by a face displaying one of the emotional expressions for 200 ms. Participants

were asked to indicate the emotional expression displayed in pictures of White and Moroccan emotional faces, one picture per trial, by pressing one of two keys ("I" or "P"), corresponding to either angry or sad. Response mappings reversed across blocks, whereas starting order of response mapping was counterbalanced between participants. The task consisted of five blocks: two practice and three experimental blocks. A 20-trial practice block to familiarize participants with the task preceded the first and second experimental blocks. No practice block was included before the third experimental block because the response mapping used here was identical to the one used in experimental Block 1. Each experimental block consisted of all 72 pictures displayed one time in random order.

Implicit association tests. After the categorization task, all participants first completed an eIAT as described in Study 1. Finally, participants completed an implicit prejudice IAT (Greenwald et al., 1998). The procedure of this IAT was identical to the procedure we described for the eIAT with the exception that the two discrete emotions were replaced by positive and negative as attributes. As stimuli, we used five positive words: love (*liefde*), peace (*vrede*), healthy (*gezond*), nice (*leuk*), and safe (*veilig*), and five negative words—hate (*haat*), war (*oorlog*), torture (*marteling*), cancer (*kanker*), and accident (*ongeluk*).

The IAT measures implicit prejudice with Moroccan relative to Dutch men. In this study, the eIAT and the IAT were used as individual difference measures for predicting participants' performance on the emotion categorization task. To minimize between participants' differences in IATs due to order effects, words were presented in a fixed, pseudo-random order.

Results

To test whether individual differences in associative strength predicted participants' response latencies in categorizing emotional displays of White and Moroccan men, we will first report analyses of both IATs and then turn to the speeded categorization task.

Implicit association tests. Using the algorithm of Greenwald et al. (1998), we calculated two measures of associative strength per participant: one for the association between social categories and discrete emotional expressions (eIAT score) and the other for the association between social categories and general valence (IAT score). Incorrect responses (6.5% eIAT, 6.7% IAT), the practice blocks, and the first two trials of all test blocks were excluded from analysis; response latencies smaller than 300 ms were recoded to 300 ms and those bigger than 3000 to 3000 ms. Analyses were performed on log-transformed response latencies, but untransformed latencies are reported. One participant was excluded from the analyses because of an IAT score that differed more than 3 *SD*s from the overall mean.

In the eIAT, people generally responded faster during the stereotype-congruent block ($M = 613$, $SD = 107$) than during the stereotype-incongruent block ($M = 700$, $SD = 128$), $F(1, 72) = 125.00$, $p < .01$, $\eta_p^2 = .64$. This indicates that our participants, on average, associated anger more with Moroccan and sadness more with Dutch men, than sadness with Moroccan and anger with Dutch men.

A similar analysis on the IAT data yielded a significant IAT effect, $F(1, 72) = 252.06$, $p < .01$ $\eta_p^2 = .78$. Again, participants were faster in the congruent block ($M = 581$, $SD = 105$) than in the incongruent block ($M = 725$, $SD = 145$). This result indicates that participants on average associated negative words with Moroccan and positive words with Dutch men more strongly than positive words with Moroccan and negative words with Dutch men. In short, overall, our participants showed a negativity bias toward Moroccan and/or a positivity bias toward Dutch men. A modest correlation between the eIAT and the IAT ($r = .26$, $p = .03$) showed that both implicit association tests largely measured different constructs.

Emotion categorization task. In line with Bijlstra et al. (2010), incorrect trials (17%)[7] and response latencies below 200 ms or above 3,000 ms (<1%) were excluded from analysis. Next, we log-transformed and standardized response latencies per participant to minimize individual differences in overall response latencies between participants. Due to the skewed distribution of the response latencies, all analyses were performed on log-transformed response latencies; untransformed latencies are presented herein for ease of interpretation. In line with Study 1, we again predicted a two-way interaction of Expression and Ethnicity, signaling a stereotype-driven bias in emotional expression recognition across Ethnicity. More importantly, to test whether the eIAT predicted the stereotype effect different from the IAT, analyses were performed with individual difference scores of both the eIAT and the IAT as continuous predictors. We hypothesized that the eIAT would predict the Expressions × Ethnicity interaction effect, even controlling for the IAT.

Mean log-transformed response latencies were subjected to a 2 (Ethnicity: White vs. Moroccan) × 2 (Expression: angry vs. sad) repeated-measures ANOVA with standardized eIAT and IAT scores entered as continuous predictors. This analysis yielded a two-way interaction between Ethnicity and Expression, $F(1, 69) = 12.98$, $p < .01$, $\eta_p^2 = .16$ (see Figure 4). With this, we replicated the stereotype effect of Study 1 and Bijlstra et al. (2010), demonstrating faster recognition of angry Moroccan ($M = 682$, $SD = 130$) than angry White ($M = 695$, $SD = 129$) faces, $F(1, 69) = 7.42$, $p < .01$, $\eta_p^2 = .10$, and slower recognition of sad Moroccan ($M = 694$, $SD = 120$) than sad White ($M = 681$, $SD = 121$) faces, $F(1, 69) = 9.05$, $p < .01$, $\eta_p^2 = .12$.

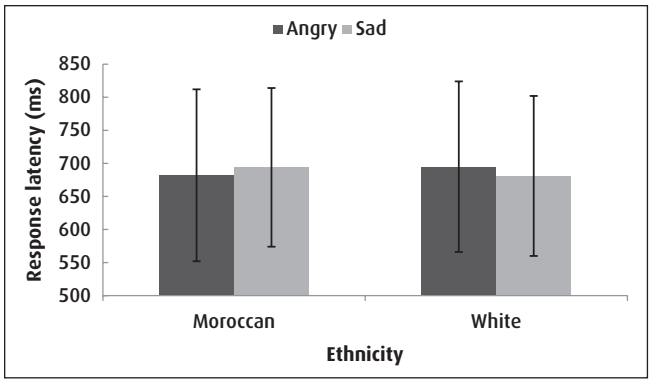

Figure 4. Mean response latencies and standard deviations in milliseconds as a function of ethnicity and expression measured with a still-frame emotion recognition task.

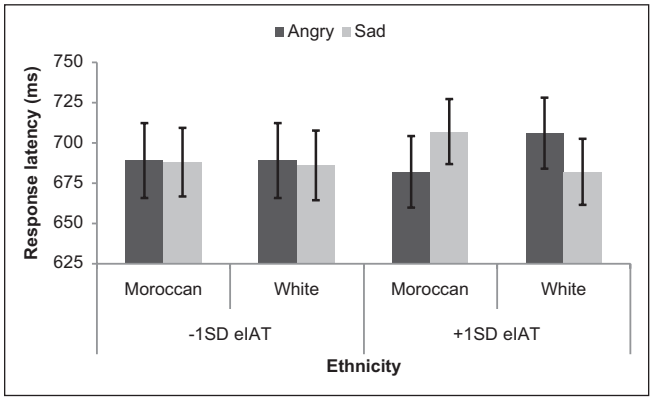

Figure 5. The relationship between participants' eIAT score and their stereotype bias in emotion recognition, represented by estimated response latencies and standard error for −1 *SD* and +1 *SD* from the eIAT mean.
Note. eIAT = emotion Implicit Association Test.

More importantly, the ANOVA also yielded the predicted three-way interaction between Ethnicity, Expression, and eIAT, $F(1, 69) = 5.30$, $p = .02$, $\eta_p^2 = .07$ (see Figure 5). Stronger anger-Moroccan and sadness-Dutch associations (compared with the opposite) were related to a stronger emotion recognition stereotype effect, conceptually replicating the eIAT moderation of Study 1.

In addition, a non-predicted three-way interaction with Ethnicity, Expression, and IAT scores was also observed, $F(1, 69) = 4.16$, $p = .04$, $\eta_p^2 = .06$ (see Figure 6). Importantly, this interaction pattern differed from the previous one. Participants low in prejudice showed stronger emotion recognition stereotype effects than highly prejudiced participants. No four-way interaction or main effects were found.[8]

General Discussion

The goal of the present research was to investigate whether stereotype associations between specific emotional expressions

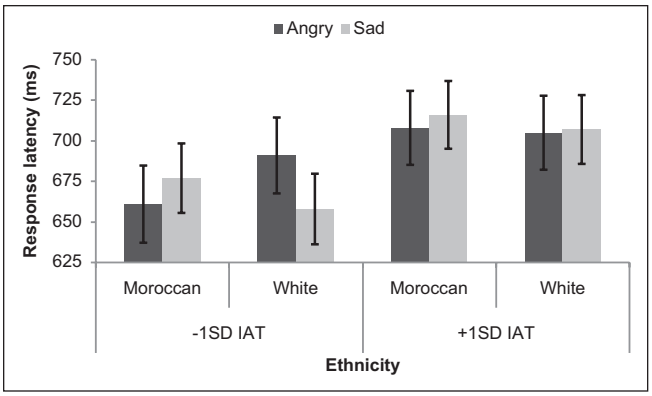

Figure 6. The relationship between participants' IAT score and their stereotype bias in emotion recognition, represented by estimated response latencies and standard error for −1 *SD* and +1 *SD* from the IAT mean.
Note. IAT = Implicit Association Test.

and social categories underlie stereotypic emotion recognition biases. Across two studies, we replicated a recently reported stereotyping effect in emotion recognition (Bijlstra et al., 2010) using research paradigms with dynamically (Study 1) and statically (Study 2) displayed expressions. In general, participants decoded anger more quickly on Moroccan than White male faces, whereas participants more quickly decoded sadness on White than Moroccan male faces.

More importantly, the current work extends previous findings by providing novel evidence that an individual difference measure of stereotype associations (eIAT) predicts the purported biasing effects of stereotypes on expression recognition. We found consistent evidence that participants' stereotype associations predicted the strength of their stereotype effects in recognizing expressions. In both studies, as perceivers' level of Moroccan-anger and Dutch-sadness associations (compared with the opposite) increased, so did perceivers' tendency to decode anger more readily on Moroccan male faces and sadness on White male faces. This finding provides direct evidence that stereotype associations underlie the effects found in earlier research by Bijlstra and colleagues (2010). Within a single-valence (sad versus angry) context, recognition of emotions that are stereotypically related to a social category is facilitated, but only to the extent that these stereotype associations are present in the perceiver.

Furthermore, the current work also offers a novel methodological innovation. To show the moderation of stereotype association on emotion recognition, we developed the emotion IAT. Based on the data from Study 2, it seems clear that this measure taps into distinctive emotion related associations, while yielding only modest overlap with the evaluative IAT (in Study 2, $r = .26$). In addition, the two implicit measures were differentially predictive of the stereotype effect, demonstrating discriminant validity of the two implicit association measures. To our knowledge, only

Becker and colleagues (2007) have used an IAT with emotional expressions as categories. However, they reported an eIAT using happiness and anger as emotion concepts, which, in our view, may be confounded with general valence effects. The purpose of the eIAT in the current research was to measure specific stereotypic associations between categories and emotions. Measuring implicit associations within a single-valence domain (or neutral valence, see Amodio & Devine, 2006) might be better suited to tapping into these specific stereotype associations. We should, however, not draw too strong conclusions about the exact constructs measured by the eIAT. For example, the eIAT involves the relative strength of the association between each discrete emotion and each of the two social categories used in the present study. Therefore, from this measure, we cannot infer whether a specific anger-Moroccan association drives our results or whether one or a combination of the other associations caused the demonstrated effects. Future research would benefit from investigating whether the rationale behind and findings demonstrated by this new eIAT generalize to other target categories and expressions as well, and further investigate its predictive value.

Furthermore, it is worth noting that the evaluative IAT score did predict the emotion recognition effect, albeit in an unexpected direction. At first blush, this seems somewhat surprising because the emotion recognition task involved two negative emotions: anger and sadness. Therefore, implicit prejudice was not expected to play a role in this specific comparative context. Moreover, the effect indicated that the stronger the IAT effect (i.e., the stronger the negative associations with Moroccan men and positive associations with Dutch men), the *weaker* the emotion recognition effect. The data suggested that relatively highly prejudiced participants experience difficulties categorizing negative emotional expressions in general, at least within the studied comparative context. Notably, because both the lack of statistical evidence for a four-way interaction and the reversed pattern of what one might predict indicate that these data cannot explain the eIAT effect. However, this pattern of data is intriguing and worthy of consideration in its own right. One possibility might be that these participants ascribe negativity in general to the Moroccan social category and therefore do not distinguish one negative expression from another (i.e., increased heterogeneity among negative expressions). For relatively low-prejudice participants, there seems to be more room for distinctive, discrete negative emotion effects to be perceived. Of course, this potential explanation needs additional investigation in future research but if supported could be a valuable insight into how perceptions of group homogeneity versus heterogeneity might be driven at least in part by our evaluative associations with target categories.

With the present research, we contribute to a growing body of expression recognition and social categorization research. However, thus far, research has only demonstrated clear evidence for prejudice in biasing the decoding of facial expressions (e.g., Hugenberg, 2005; Kang & Chasteen, 2009), showing that perceivers' implicit prejudice predicted the strength of prejudice-congruency effects (in a dual-valence context; Hugenberg & Bodenhausen, 2003, 2004). With the present studies, we extend this knowledge by showing that perceivers' stereotype associations predicted recognition of stereotypic expressions when decoded in a context where only stereotyping effects are likely (i.e., a single-valence context). The observed stereotype effect seems to be independent of general evaluations of social categories (Study 2), suggesting dissociable effects of prejudices and stereotypes in expression perception. A dissociation between these processes has previously been demonstrated by Amodio and Devine (2006), and was recently suggested in emotion recognition research (Bijlstra et al., 2010). However, as yet, no direct test of stereotype biases in expression recognition was carried out. With the present research, we are the first to show that stereotypic group-expression associations affect expression decoding in a stereotype-congruent way.

Together with prior prejudice-congruency findings (e.g., Hugenberg & Bodenhausen, 2003), the present results provide direct evidence for distinct roles of stereotypes and prejudice in generating distinct emotion recognition biases. Ideally, in future research, these effects will be investigated within a full design, testing both prejudice- and stereotype-congruency at once. Furthermore, it is important to note that the results of our research apply to White participants decoding anger and sadness of two social categories present in the Netherlands. Future research should investigate whether these effects generalize to other social categories and emotional expressions as well. We would, for example, predict that similar effects could be obtained by using two other social categories in a single-valence comparative context, as long as perceivers stereotypically associate the expressions with the categories used. The results of our studies support the idea that participants' stereotype associations are indicative of their stereotype effects in recognizing emotional expressions.

Finally, the current work was focused on the effects of stereotype associations in expression perception, whereas the observed recognition results could also contribute to the way in which perceivers maintain their stereotype associations. For example, those people holding strong stereotype associations between Moroccans and anger more easily observe an angry Moroccan and potentially therefore see more angry Moroccans around them. By this manner, stereotypes may be self-perpetuating by influencing emotion recognition in a stereotype-confirming way.

In sum, we aimed at providing direct evidence for the influence of perceivers' stereotypic associations in decoding emotional expressions of Moroccan and White males. In two studies, we replicated Bijlstra and colleagues (2010) and extended these findings to dynamic facial displays. Most importantly, we showed that stereotype associations underlie emotion recognition effects obtained when people decode

stereotype expressions and valence is not made salient. Given the fact that discrete emotional expressions elicit specific behavioral tendencies, it is especially important to differentiate one negative expression from another.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Notes

1. In the current article, we use the terms *Moroccan* and *White* to refer to men from the Moroccan Dutch and White Dutch community in the Netherlands.
2. A pilot study was conducted to test whether we could conceptually replicate the dissociation between prejudice and stereotype processes in expression decoding (Bijlstra et al., 2010) using a morph movie task. Sixty-nine participants were randomly assigned to one of two dual-valence conditions (n = 46, happy and angry *or* happy and sad) or the single-valence condition (n = 23, angry *and* sad) in which Moroccan and White faces gradually change from one expression into a second expression. Before analyzing, data of three participants were excluded because the median of their response latencies fell outside the duration of the film clips (>8 s.). To test whether prior findings of prejudice and stereotype associations were replicated, we analyzed both dual-valence conditions together, followed by the single-valence condition. In line with previous work, we predicted and found a two-way interaction of expression valence and ethnicity of models in the dual-valence conditions, $F(1, 42) = 33.03$, $p < .001$, $\eta_p^2 = .44$. Replicating Bijlstra et al., this interaction was not qualified by the three-way interaction including context (happy and angry vs. happy and sad; $F < 1$), reflecting that the two-way interaction did not differ between dual-valence conditions. Furthermore, in the single-valence condition, we predicted and observed a two-way interaction of emotional expression and ethnicity of models, $F(1, 21) = 9.00$, $p = .007$, $\eta_p^2 = .30$. These findings confirm our predictions that both prejudice- and stereotype-congruency effects in emotion recognition can be discriminated when processing dynamically displayed expressions.
3. In the Implicit Association Tests (IATs), we presented Dutch and Moroccan names. We chose to use the label Dutch instead of White in the IAT instructions because the names are typical Dutch and therefore better reflect this response category than White names in general.
4. In between the morphed movies task and the emotion Implicit Association Test (eIAT), we instructed participants to rate emotional intensity of the initial states of the used film clips, using a 7-point Likert-type scale anchored from (1) *not intense* to (7) *very intense*. Subsequently, they were asked to judge an image for purposes unrelated to this Study. Per image, we calculated an average intensity score and tested whether the five Moroccan images differed from the five White images. No significant differences were found between the four groups, $F < 1$ ($M_{\text{Moroccan-angry}} = 4.71$, $SD = 0.88$, $M_{\text{Moroccan-sad}} = 4.48$, $SD = 1.21$, $M_{\text{White-angry}} = 4.38$, $SD = 0.89$, $M_{\text{White-sad}} = 4.65$, $SD = 0.54$).
5. A different, non-significant pattern of results was obtained when leaving these trials and/or participants in our data.
6. After calculating participants' stereotype effect, one participant scored more extremely than other participants (+3 *SD*). However, none of the mean group-emotion response latencies of this participant could be detected as an outlier (all mean response latencies fell inside the ±3 *SD* range). Cook's distance of the stereotype effect for this particular participant is 0.19, scoring below the cut-off of 1 and above the cut-off of 4 / n. Excluding this participant from analysis revealed a non-significant three-way interaction, $F(1, 84) = 2.525$, $p = .116$, $\eta_p^2 = .029$.
7. We subjected the percentage correctly categorized expressions to a 2 (Ethnicity: White vs. Moroccan) × 2 (Expression: angry vs. sad) repeated-measures ANOVA with standardized eIAT and IAT scores as continuous predictors. This analysis revealed only a significant two-way interaction in line with our stereotype effect, $F(1, 69) = 59.57$, $p < .001$, $\eta_p^2 = .46$.
8. To facilitate the comparison between studies, we also subjected the log-transformed response latencies to two separate 2 (ethnicity) × 2 (expression) repeated-measures ANOVAs with either the eIAT or IAT as continuous factor. These analyses revealed the expected two-way interactions (ps < .001) and a marginal significant three-way interaction when including the eIAT, $F(1, 71) = 3.276$, $p = .075$, $\eta_p^2 = .044$, whereas the three-way interaction including the IAT no longer reached significance, $F(1,71) = 2.358$, $p = .129$.

References

Amodio, D. M., & Devine, P. G. (2006). Stereotyping and evaluation in implicit race bias: Evidence for independent constructs and unique effects on behavior. *Journal of Personality and Social Psychology*, *91*, 652-661.

Becker, D. V., Kenrick, D. T., Neuberg, S. L., Blackwell, K. C., & Smith, D. M. (2007). The confounded nature of angry men and happy women. *Journal of Personality and Social Psychology*, *92*, 179-190.

Bijlstra, G., Holland, R. W., & Wigboldus, D. H. J. (2010). The social face of emotion recognition: Evaluations versus stereotypes. *Journal of Experimental Social Psychology*, *46*, 657-663.

Cottrell, C. A., & Neuberg, S. L. (2005). Different emotional reactions to different groups: A sociofunctional threat-based approach to "prejudice." *Journal of Personality and Social Psychology*, *88*, 770-789.

Dotsch, R., Wigboldus, D. H. J., & van Knippenberg, A. (2011). Biased allocation of faces to social categories. *Journal of Personality and Social Psychology*, *100*, 999-1014.

Dunham, Y. (2011). An angry=outgroup effect. *Journal of Experimental Psychology*, *47*, 668-671.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The implicit association test. *Journal of Personality and Social Psychology*, *74*, 1464-1480.

Hess, U., Blairy, S., & Kleck, R. E. (1997). The intensity of emotional facial expressions and decoding accuracy. *Journal of Nonverbal Behavior*, *21*, 241-257.

Hugenberg, K. (2005). Social categorization and the perception of facial threat: Target race moderates the response latency advantage for happy faces. *Emotion*, *5*, 267-276.

Hugenberg, K., & Bodenhausen, G. V. (2003). Facing prejudice: Implicit prejudice and the perception of facial threat. *Psychological Science*, *14*, 640-643.

Hugenberg, K., & Bodenhausen, G. V. (2004). Ambiguity in social categorization: The role of prejudice and facial affect in racial categorization. *Psychological Science*, *15*, 342-345.

Hugenberg, K., & Sczesny, S. (2006). On wonderful women and seeing smiles: Social categorization moderates the happy face response latency advantage. *Social Cognition*, *24*, 516-539.

Hutchings, P. B., & Haddock, G. (2008). Look Black in anger: The role of implicit prejudice in the categorization and perceived emotional intensity of racially ambiguous faces. *Journal of Experimental Social Psychology*, *44*, 1418-1420.

Ishii, K., Miyamoto, Y., Niedenthal, P. M., & Mayama, K. (2011). When your smile fades away: Cultural differences in sensitivity to the disappearance of smiles. *Social Psychology and Personality Science*, *2*, 516-522.

Kang, S. K., & Chasteen, A. L. (2009). Beyond the double jeopardy hypothesis: Assessing emotion on the faces of multiple-categorizable targets of prejudice. *Journal of Experimental Social Psychology*, *45*, 1281-1285.

Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion*, *24*, 1377-1388.

Leppänen, J. M., Kauppinen, P., Peltola, M. J., & Hietanen, J. K. (2007). Differential electrocortical responses to increasing intensities of fearful and happy emotional expressions. *Brain Research*, *1166*, 103-109.

Marsh, A. A., Adams, R. B., Jr., & Kleck, R. E. (2005). Why do fear and anger look the way they do? Form and social function in facial expressions. *Personality and Social Psychology Bulletin*, *31*, 73-86.

Niedenthal, P. M., Halberstadt, J. B., Margolin, J., & Innes-Ker, A. H. (2000). Emotional state and the detection of change in facial expression of emotion. *European Journal of Social Psychology*, *30*, 211-222.

Plant, E. A., Hyde, J. S., Keltner, D., & Devine, P. G. (2000). The gender stereotyping of emotions. *Psychology of Women Quarterly*, *24*, 81-92.

Recio, G., Sommer, W., & Schacht, A. (2011). Electrophysiological correlates of perceiving and evaluating static and dynamic facial emotional expressions. *Brain Research*, *1376*, 66-75.

Sacco, D. F., & Hugenberg, K. (2009). The look of fear and anger: Facial maturity modulates recognition of fearful and angry expressions. *Emotion*, *9*, 39-49.

Said, C. P., Sebe, N., & Todorov, A. (2009). Structural resemblance to emotional expressions predicts evaluation of emotionally neutral faces. *Emotion*, *9*, 260-264.

Sato, W., Fujimura, T., & Suzuki, N. (2008). Enhanced facial EMG activity in response to dynamic facial expressions. *International Journal of Psychophysiology*, *70*, 70-74.

Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., & Matsumura, M. (2004). Enhanced neural activity in response to dynamic facial expressions of emotion: An fMRI study. *Cognitive Brain Research*, *20*, 81-91.

Spruyt, A., De Houwer, J., Hermans, D., & Eelen, P. (2007). Affective priming of nonaffective semantic categorization responses. *Experimental Psychology*, *54*, 44-53.

Zebrowitz, L. A., Kikuchi, M., & Fellous, J. M. (2010). Facial resemblance to emotions: Group differences, impression effects, and race stereotypes. *Journal of Personality Social Psychology*, *98*, 175-189.