The influence of multiple social categories on emotion perception

Craig, Belinda M.; Lipp, Ottmar V.

Published in:
Journal of Experimental Social Psychology

DOI:
10.1016/j.jesp.2017.11.002

The Influence of Multiple Social Categories on Emotion Perception

Belinda M. Craig & Ottmar V. Lipp

School of Psychology and Speech Pathology, Curtin University, Bentley, WA, 6102, Australia

Word count: 10 189

Running head: Multiple Social Categories on Emotion

Address for correspondence:

Belinda M. Craig, Curtin University, School of Psychology and Speech Pathology, Bentley WA, 6102, Australia. E-mail: belinda.craig@curtin.edu.au

Acknowledgments:

This research was partly funded by a grant awarded to BMC and OVL from the Curtin University School of Psychology and Speech Pathology Research Allocation Fund SRAF-2015 and an Australian Research Council Discovery Project (project number DP150101540) awarded to OVL.

Abstract

Although the human face provides multiple sources of social information concurrently (race, sex, age, etc.), the majority of studies investigating how social category cues influence emotional expression perception have investigated the influence of only one social category at a time. Only a couple of studies have investigated how race and sex cues concurrently influence emotion perception and these studies have produced mixed results. In addition, the concurrent influence of age and sex cues on emotion perception has not been investigated. To address this, participants categorized happy and angry expressions on faces varying in race (Black and White) and sex (Experiments 1a and 1b) or age (older adult and young adult) and sex (Experiment 2). In Experiments 1a and 1b, results indicated that sex but not race influenced emotion categorization. Participants were, on average, faster to categorize happiness than anger on female, but not on male faces. In Experiment 2, both the age and the sex of the face independently influenced emotion categorization. Participants were faster to categorize happiness than anger on female and young adult faces, but not on male or older adult faces. Bayesian ANOVAs provided additional evidence that the sex of the face had the strongest influence on emotion categorization speeds in Experiment 1a and 1b, but both age and sex cues had an equal influence on emotion categorization in Experiment 2.

Keywords: Emotion recognition, social categorization, race, sex, age

A person's sex, race, age, and emotional expressions can be quickly and accurately recognized from a face (e.g., Karnadewi & Lipp, 2011). Although early research focused on understanding how each facial attribute is processed in isolation (see Bruce & Young, 1986), subsequent studies demonstrated that other task irrelevant facial cues influence how quickly and accurately a task relevant attribute is recognized (e.g., Craig, Mallan, & Lipp, 2012; Hugenberg, 2005; Hugenberg & Sczesny, 2006; Johnson, Freeman, & Pauker, 2012; Kloth, Damm, Schweinberger, & Wiese, 2015). For example, facial cues of race, sex, and age can influence the recognition of emotional expressions (Becker, Kenrick, Neuberg, Blackwell, & Smith, 2007; Craig & Lipp, 2017a; Hugenberg, 2005; Hugenberg & Sczesny, 2006).

One method used to probe the influence of social category information on emotion perception is speeded emotion categorization, and in particular looking at modulation of the happy categorization advantage – the faster categorization of happy than negative expressions like anger and sadness (Leppänen & Hietanen, 2003). Previous studies have demonstrated that when male and female faces are presented within a task, participants are faster to categorize happiness than anger when expressed on (Caucasian) female, but not male faces (Becker et al., 2007; Craig & Lipp, 2017b; Hugenberg & Sczesny, 2006). Similarly, White participants are faster to categorize happiness than anger on White (male) faces, but not on Black (male) faces (Craig, Koch, & Lipp, 2017; Craig et al., 2012; Hugenberg, 2005). Participants are also faster to categorize happiness than anger on young adult (male) faces, but not on older adult (male) faces (Craig & Lipp, 2017a).

The finding that social category information influences emotion perception, suggests that social category information is quickly processed before the emotion categorization response is made and a number of mechanisms have been proposed to explain why these social category cues influence emotion categorization. One possibility is that stereotypes (information based associations) held about different social groups are quickly activated and influence emotion categorization (Bijlstra, Holland, & Wigboldus, 2010). For example, implicit stereotypes of males or African Americans as threatening may facilitate recognition of anger on these faces.

Alternatively, it has been suggested that some facial features that indicate social categories (e.g., face shape, wrinkled skin etc.) overlap with features of particular emotional expressions. For example, the wrinkled skin and down turn of the mouth corners that occur with old age may facilitate recognition of sadness on older adult faces (Fölster, Hess, & Werheid, 2014; Malatesta & Izard, 1984) or the full brow and angular jaw line associated with masculinity could facilitate categorization of anger (e.g., Becker et al., 2007). Finally, it has been suggested that implicit evaluations (affective associations) of social categories influence emotion categorization. Females are evaluated as more pleasant than males (e.g., Eagly, Mladinic, & Otto, 1991), White faces as more pleasant than Black faces (e.g., Nosek, Greenwald, & Banaji, 2005), and young faces as more pleasant than older adult faces (e.g., Greenwald et al., 2002). The activation of these positive evaluations results in faster categorization of positive expressions like happiness on relatively positively evaluated faces (female, own-race, and young adult faces; Craig et al., 2017; Craig & Lipp; 2017a; Hugenberg, 2005; Hugenberg & Sczesny, 2006).

This evaluative congruence explanation has been favored over stereotype based or facial structure based explanations to account for the influence of social categories on the happy categorization advantage. This is because social category cues have been found to moderate the happy categorization advantage similarly when both stereotype congruent and structurally overlapping as well as stereotype incongruent and structurally dissimilar negative expressions are used to investigate the happy categorization advantage. For example, in previous studies the happy categorization advantage was larger for female than male faces and larger for White than for Black faces when participants categorized happiness vs. anger (a negative expression stereotypically associated with males and Black faces, and structurally overlapping with masculine facial structure) as well as sadness (a negative expression that is more stereotypically associated with females and structurally overlaps with the feminine facial structure; Craig et al., 2017; Hugenberg, 2005; Hugenberg & Sczesny, 2006). Similarly, the happy categorization advantage was larger for young adult than older adult faces when participants categorized happiness vs. both anger (stereotypically associated with young adults) and sadness (stereotypically associated with older adults and structurally overlapping with old age cues; Craig & Lipp, 2017a). It is likely that stereotypes about social groups and the structural overlap between cues indicating social categories and emotional expressions play an important role in the recognition of emotional expressions under some circumstances and within some tasks (e.g., see Becker et al., 2007; Bijlstra et al., 2010; Hess, Adams, Grammer, & Kleck, 2009; Sacco & Hugenberg, 2009); however, these explanations are less likely to account for the influence of social category information on the happy categorization advantage.

As described above, the majority of studies have focused on the influence of one social category on emotion perception at a time. Few studies have investigated how multiple social categories concurrently influence emotion categorization; however, there is a body of research describing whether and how social stimuli varying along multiple social dimensions (crossed categories) are categorized and evaluated. Many studies investigating these processes have used the 'who-said-what' procedure. In these experiments participants are presented with faces or vignettes varying across two social categories. For example, Black and White male and female faces may be presented paired with utterances attributed to each person. At a subsequent test phase, participants must indicate 'who said what' by matching the faces to the utterances they were paired with. Categorization processes are assessed by looking at how many within- and between-category misattributions are made. Within-sex-within-race, within-sex-between-race, between-sex-within-race, and between-sex-between-race misattributions are compared to determine the extent to which participants categorized the faces by their race, or sex or both (e.g., see Stangor, Lynch, Duan, & Glass, 1992 for use of this approach). This work has provided evidence for a number of ways that multiple social categories are combined and evaluated (see Urban & Miller, 1998 for a meta-analytic review). Considering these different possible combinations is important as automatic social categorization and evaluation processes underlie the influence of social category cues on emotion perception.

Findings from the crossed-categorization literature suggest a number of possible ways that multiple categories may concurrently influence emotion perception. It is possible that the influence of each relevant category is summed (*additive model*). As females are evaluated as more pleasant than males (Eagly et al., 1991) and White faces are implicitly evaluated as more pleasant than Black faces (Nosek et al., 2005), if the influence of these two facial cues on emotion perception is additive, White females would be evaluated as most positive, followed by Black females and White males together, followed by Black males. This would result in the largest happy categorization advantage for White females, a reduced advantage for Black female and White males and a further reduced advantage for Black males. This could be conceptualized as a Target race × Emotion and a Target sex × Emotion interaction in the absence of a three-way Target race × Target sex × Emotion interaction best explaining the data. Alternatively, one category may be dominant and have an influence on emotion categorization with a reduced or absent influence of the other social category (*category dominance model*). Within this model, we may observe a happy categorization advantage for both White and Black females but a reduced or absent advantage for both White and Black males indicating a stronger influence of sex than race on emotion categorization. Statistically, this could be conceptualized as a single two-way interaction (e.g., Target sex × Emotion interaction) in the absence of any other interactions (e.g., Target race × Emotion or Target race × Target sex × Emotion interactions) as the model that best fits the observed data. Finally, it is possible that two categories could interact and intersectional identities, distinctly different from the sum of their categories, could influence emotion perception (*interactive model*). For example, the largest happy categorization advantage could be observed for Black females as the evaluation of the intersectional identity (Black female) is different to the sum of the two category evaluations. Statistically this could be conceptualized as a Target race × Target sex × Emotion interaction best explaining the observed data.

Only a few studies have looked at the combined influence of two social categories on emotion perception and their results led to different conclusions about how multiple social categories influence emotion perception. Looking at the combined influence of age and race cues, Kang and Chasteen (2009) presented videos of computer generated young and old, Black and White, male faces with facial expressions morphing between happy, angry, and neutral. Non-Black participants were faster to detect the onset and slower to detect the offset of anger presented on young Black than on young White faces; however, this pattern reversed when the faces were old. These findings suggest that the influence of multiple categories on emotion perception is interactive. In this case it was suggested that, age and race cues combined to create unique intersectional identities which influenced emotion onset/offset detection.

In another study (Lipp, Craig, & Dat, 2015), participants categorized happy and angry expressions on photographic images of White male faces along with White female faces in one task, Black male faces in a second task and Black female faces in a third task. Looking across tasks, participants were faster to categorize happiness than anger on Black female faces but this happy categorization advantage was numerically smaller (25ms) than that for White females (46ms). This attenuated happy categorization advantage may have occurred as the positive female evaluation and the negative black evaluation combined to an intermediate evaluation and thus a smaller happy advantage. This may be evidence that the influence of race and sex on emotion perception is additive; however, race and sex were not fully crossed within a single task in this investigation.

Further, Smith, La France, and Dovidio (2017) asked participants to categorize happy and neutral expressions or angry and neutral expressions presented on White and Black, male and female faces. Participants who completed the angry vs. neutral categorization task were faster to categorize neutral than angry expressions presented on White female faces, no faster to categorize either neutral or angry expressions presented on Black female or White male faces and faster to categorize angry than neutral expressions on Black male faces providing some evidence for an additive influence of race and sex cues on emotion categorization. In the happy vs. neutral task, however, happy expressions were categorized faster than neutral expressions when presented on

Black female faces, Black male faces, and White female faces, but not on White male faces and the happy categorization advantage for Black females was the largest of all. The results of the happy vs. neutral categorization task suggested an interactive pattern of results where the intersectional identity (Black female) rather than the sum of the two categories (Black and female) influenced emotion categorization.

Although there were some similarities across studies (e.g., all studies used expressions of happiness and anger), these studies differed in a number of ways which may have contributed to the discrepant patterns of results observed across studies. The combinations of categories under investigation differed across studies with race and age investigated by Kang and Chasteen (2009) and race and sex investigated by Smith et al. (2017). Different types of stimuli were also used. Kang and Chasteen used computer generated stimuli and Lipp et al. (2015) and Smith et al. used photographs drawn from a range of face databases (with some overlap in the databases used across the studies). A different number of targets were used to represent each category. For example, Kang and Chasteen used two base faces of each category, Lipp and colleagues used eight faces of each category with the same individuals encountered expressing happiness and anger, and Smith et al. used eight faces of each category, but each individual only appeared expressing one emotion. Different tasks were also used, with Kang and Chasteen asking participants to detect the onset/offset of emotions in videos and Lipp and colleagues and Smith and colleagues asking participants to categorize emotional expressions as quickly as possible. Happy expressions were presented with angry expressions in the study by Lipp and colleagues and happy and neutral or angry and neutral expressions were presented together in the study by Smith and colleagues.

The Current Study

The findings of Kang and Chasteen, (2009), Lipp et al. (2015), and Smith et al. (2017), provide a mixed impression of the way multiple social categories combine to influence the processing of emotional expressions. Additionally, the combined influence of age and sex cues on emotion categorization has not yet been investigated. As such, the current study had two main aims. Firstly, we aimed to reexamine the concurrent influence of race and sex cues on emotion recognition using speeded emotion categorization tasks. Like all previous studies (Kang & Chasteen, 2009; Lipp et al., 2015; Smith et al., 2017), we investigated the influence of multiple social categories on happiness and anger perception. We decided to use only happy and angry faces in a single task rather than comparing these expressions with neutral faces as the mechanism underlying the influence of social categories on the happy categorization advantage is well established (Hugenberg, 2005; Hugenberg & Sczesny, 2006; Craig et al., 2017). Secondly, we aimed to determine how age and sex cues combine to influence emotion processing. To address these aims, participants categorized happy and angry expressions on faces varying in race (Black and White) and sex (Experiments 1a and 1b) or age (young and older adult) and sex (Experiment 2) as quickly and accurately as possible. Data were analyzed using traditional frequentist statistics as well as Bayesian statistics in order to determine the statistical model that best fit the observed data in each experiment.

Consistent with the most similar previous study (Lipp et al., 2015), we predicted that race and sex cues would have an additive influence on emotion perception. This would manifest in significant Target race × Emotion and Target sex × Emotion interactions in the absence of a higher order Target race × Target sex × Emotion interaction and the Bayesian analysis would indicate this to be the strongest model that best fit the observed data. In Experiment 2, an additive influence of age and sex cues on emotion categorization was also predicted. This prediction was made as a number of similarities have been observed in the processing of race and age cues. For example, race and age cues have a similar influence on emotion categorization when investigated separately

(Craig & Lipp, 2017a; Hugenberg, 2005), own race and own age faces are evaluated as more pleasant than other race and other age faces (Craig, Lipp, & Mallan, 2014; Greenwald et al., 2002; Nosek et al., 2005), and both other race and other age faces are categorized faster by their race or age and recognized less accurately than own race and own age faces (Johnston, Kanazawa, Kato, Oda, 1997; Levin, 2000; Meissner & Brigham, 2001; Rhodes & Anastasi, 2012).

Experiments 1a, 1b, and 2

Methods

The study was conducted in line with the American Psychological Association's Ethical Principles in the Conduct of Research with Human Participants and the procedures were approved by the Curtin University Human Research Ethics Committee. In the methods section we report all measures, manipulations, and exclusions. No further data were collected after analyses had taken place for each experiment.

Participants. Initially target sample size was based on previous studies that successfully found significant Race × Emotion, Sex × Emotion, or Age × Emotion interactions (e.g., Craig & Lipp, 2017a; Lipp et al., 2015; Hugenberg, 2005) or a Race × Sex × Emotion interaction (Smith et al., 2017). We aimed for a sample that was of similar or larger size than used previously for Experiments 1a and 2. We also conducted a power analysis using a simulation method based on the means, variances, and inter-correlations observed in Experiment 1a. It was estimated that for the Sex × Emotion interaction, around 13 participants would be required, for the Race × Emotion interaction around 65 participants would be required, and for the Sex × Race × Emotion interaction around 650 participants would be required to have an 80% chance of detecting a significant effect. In Experiments 1a and 2, we stopped recruiting once at least 36 participants had signed up to complete the study and stopped testing once all participants who had already signed up had taken part. For Experiment 1b we aimed to sample around twice this number in order to have sufficient male and female participants to investigate the influence of participant sex and because the power analysis indicated this sample size would provide a good chance of detecting significant Target race × Emotion and Target sex × Emotion interactions. For Experiment 1b we requested 80 participants on the Amazon Mechanical Turk platform assuming some data loss.

The final samples were 35 (5 Males, M = 22.69, SD = 6.98; Experiment 1a) and 37 undergraduate volunteers (11 Males, M = 20.24, SD = 2.51; Experiment 2) who received partial course credit for participation and 66 MTurk workers (32 Males, M = 35.66, SD = 9.74; Experiment 1b) who received monetary compensation in return for their participation. An additional two participants in Experiment 1a, 14 participants in Experiment 1b, and three participants in Experiment 2 took part, but were not included in analyses as they identified as a member of the racial or age outgroup (African/African American for Experiments 1a and 1b or aged over 31 for Experiment 2). These exclusions did not alter the significance or direction of the results of Experiment 1a or Experiment 2, but including the 14 African American participants in Experiment 1b led the significant Target race × Emotion interaction observed in the repeated measures ANOVA to become marginal.

Stimuli.

Experiment 1a and 1b. Photographs of eight individuals represented each category (White male, White female, Black male and Black female). These images were the same as those used in the previous investigation by Lipp and colleagues (2015). The photographs of happy and angry expressions for each individual were sourced from the Nimstim Face Stimulus Set (Tottenham et al., 2009; Poses AN_O and HA_O of models 1, 2, 3, 5, 6, 11, 12, 13, 14, 20, 21, 23, 24, 28, 38, 39, 40, 41, and 43) and the Montreal Set of Facial Displays of Emotion (Beaupré & Hess, 2005; Poses 1 and 2 of models 20, 22, 23, 25, 27, 28, 30, 32, 33, 35, 36, 37, and 38). To maintain consistency across sets, image backgrounds, necks, and hair were removed. The images were converted to grayscale where necessary, resized, and placed on a dark gray background 187 × 240 pixels in size.

Experiment 2. Again, photographs of eight individuals represented each category (young adult male, young adult female, older adult male, and older adult female). All faces were

Caucasian and images of each individual expressing happiness and anger were selected. Faces were drawn from the FACES database (Ebner, Riediger, & Lindenberger, 2010; posers 008, 013, 016, 031, 037, 057, 049, 072, 010, 020, 054, 069, 098, 115, 177, 182, 004, 015, 033, 042, 053, 059, 065, 076, 005, 024, 060, 075, 088, 096, 110, 130). They were edited in a similar manner to the stimuli used in Experiment 1; however, the hair was not removed and the stimuli were slightly larger than in Experiment 1 (304 × 390 pixels in size).

Procedures.

Laboratory studies (Experiment 1a and Experiment 2). Participants were tested in a computer laboratory in groups of no more than four. Participants were seated in front of 24in LED monitors with a screen resolution of 1920 × 1080 pixels and a refresh rate of 120 Hz. The experiment was executed using DMDX (Forster & Forster, 2003). Participants were instructed that they would see faces presented one at a time and were asked to indicate whether each face was ‘happy’ or ‘angry’ by pressing the right or left shift key as quickly and accurately as possible. Response mapping was counterbalanced across participants. Instructions were given both verbally and in writing before the commencement of the task. All participants completed only one task. In this task, faces varying in both race (Black and White) and sex (male and female) in Experiment 1a or age (older adult and young adult) and sex (male and female) in Experiment 2 were presented. On each trial, a black fixation cross was presented on a light gray background. After 500ms, the fixation cross was replaced by one of the face stimuli. The face remained on the screen until a response was made or for 2000ms. Each of the 32 individuals (eight from each category) was presented four times in a randomized order, twice with a happy expression and twice with an angry expression, resulting in 128 trials.

Online study (Experiment 1b). Participants were recruited online via the Amazon Mechanical Turk platform. The experiment was executed using Inquisit web 4 (Inquisit, 2015). As participants completed the task remotely on their own devices meaning that we could not control the monitor size or refresh rate and no experimenter was present to provide direction, images were sized to be 20% of the height of the screen, and the fixation cross was presented for 1000ms, the images remained on the screen for up to 3000ms or until a response was made, and response mapping instructions were presented at the bottom of the screen for the duration of the task. Participants also completed a short 10 trial practice with error feedback before commencing the main task. Apart from these differences, the experiment proceeded as described above.

Data preparation and analysis. As we used eight individuals per stimulus category, it was possible that our results may not generalize beyond the stimuli used in the current experiment. To address this, following the recommendation of Judd, Westfall, and Kenny (2012), we conducted linear mixed-effects analyses including stimuli and participants as random effects to account for these unwanted sources of variability. In line with recommendations, we analyzed the data with linear mixed-effects models using the lme4 package (Bates, Maechler, Bolker, & Walker, 2015) in R (R Core Team, 2017) with p-values provided by the lmerTest package (Kuznetsova, Brockhoff, Christensen, 2016). Contrast coding was used to code the fixed factors, Target sex (.5 = female, -.5 = male), Target race/age (.5 = White, young, -.5 = Black, old), and Emotion (.5 = angry, -.5 = happy). All correct response times greater than 100ms were included in the analysis. Initially, we specified participants and stimuli as random factors with random intercepts and random slopes for all main effects and interactions for participants, and random intercepts and random slopes for emotion effects for the stimuli. Where the full model failed to converge using a number of different optimizers, we simplified the models and estimated random slopes for theoretically relevant effects. For example, in Experiments 1a and 2, random intercepts and random slopes for the emotion main effect, and the Target race/age × Emotion, Target sex × Emotion, and Target race/age × Target sex × Emotion interactions were entered for participants and random intercepts and random slopes were entered for emotion effects for the stimuli.

In addition, JZS Bayes Factor ANOVAs with default priors were performed using the software package JASP (JASP Team, 2016; Morey & Rouder, 2015; Rouder, Morey, Speckman, & Province, 2012). Rather than providing information about whether each difference between means (main effect or interaction) is significant or not based on whether it is unlikely due to chance, this type of analysis compares how well different statistical models including each of the main effects, the two- and three-way interactions, and combinations of these, fit the observed data by providing a Bayes Factor (BF) for each model. The BF indicates how much more likely each model is compared to the null and the model with the largest BF is the one that best fits the observed data. In addition, BFs are transitive meaning that the BF for one model can be divided by the BF for another to provide evidence about how much more likely one model is than another given the observed data (Rouder, Engelhardt, McCabe, & Morey, 2016). The resulting BFs can be interpreted using conventions to indicate whether there is evidence for a given model (Kass & Raftery, 1995). A BF_{10} between 1 and 3 suggests only anecdotal evidence for a model, between 3 and 20 suggests positive evidence for a model, between 20 and 150 provides strong evidence for a model, and greater than 150 suggests very strong evidence for a model. Using this model fitting approach provides a better means of inferring whether the observed data best fit a category dominance, additive, or interactive pattern as Bayesian analyses not only provide evidence for or against a model but also allow us to identify the best model for the observed data.

We also conducted traditional repeated measures ANOVAs which are presented in footnotes. For this analysis, response times faster than 100ms or faster or slower than three standard deviations away from each participant's mean response time were removed as invalid. In Experiments 1a and 1b, mean response times were submitted to a 2 (Target race: Black, White) × 2 (Target sex: male, female) × 2 (Emotion: happy, angry) repeated measures ANOVA. In Experiment 2, response times were submitted to a 2 (Target age: young adult, older adult) × 2 (Target sex: male, female) × 2 (Emotion: happy, angry) repeated measures ANOVA in IBM SPSS Statistics.

Data from one participant in Experiment 1a and one participant in Experiment 1b were not included in the analyses as their error rate was over 5 standard deviations higher than the average number of errors made in each task (>48% of responses incorrect or invalid). Initial analyses were also conducted including participant sex as a factor. The results from the larger and gender balanced replication (Experiment 1b) suggested that participant sex did not moderate the influence of social category cues on emotion categorization. As such, participant sex effects that were observed are reported as footnotes for transparency and completeness, but should be interpreted with extreme caution given the small number of male participants in Experiments 1a and 2, and given the absence of a moderating influence of sex in a larger and more gender balanced sample. The pattern of results observed in accuracy data was comparable to response times and there was no evidence for a speed-accuracy trade off. These analyses are reported in Supplement 1.

Results

Experiment 1a.

Linear mixed-effects analysis. As can be seen in Figure 1, the sex of the face influenced emotion categorization, and there was also a trend towards an influence of the race of the face on emotion categorization. The results of the linear mixed-effects analysis indicated a significant main effect of emotion, $b = 31.26$, $t(33.26) = 2.97$, $p = .006$, that was moderated by a significant Target sex × Emotion interaction, $b = 58.04$, $t(27.49) = 3.34$, $p = .002$. This interaction emerged as participants were significantly faster to categorize happiness than anger on female faces, $b = 60.72$, $t(29.03) = 5.37$, $p < .001$, but not on male faces, $b = 2.61$, $t(26.03) = 0.16$, $p = .878$. The Target race × Emotion interaction was marginally significant, $b = 28.08$, $t(28.29) = 1.82$, $p = .079$. Participants were faster to categorize happiness than anger on White faces, $b = 45.57$, $t(20.14) = 3.14$, $p = .005$, but not on Black faces, $b = 17.42$, $t(18.44) = 1.25$, $p = .228$. The three way interaction was not significant, b = -15.87, t(26.29) = -0.51, p = .614, and neither were any other main effects or interactions, bs < 10.53, ts < 1.29, ps < .207[1,2].

Bayesian repeated measures ANOVA. According to a Bayes Factor ANOVA, the model with the largest Bayes factor indicating best fit with the data was the model including the main effects of target sex and emotion and only the Target sex × Emotion interaction, BF_{10} = 249348.70. This model is consistent with a category dominance pattern. There was still very strong evidence for the strongest model including the Target sex × Emotion and Target race × Emotion interactions, BF_{10} = 47079.51; however, comparing these two models, the simpler model including only the Target sex × Emotion interaction was 5.57 times more probable than the additive model including both the Target race × Emotion and Target sex × Emotion interactions indicating positive evidence for the category dominance model over the additive model. There was also very strong evidence for the model including the three-way Target race × Target sex × Emotion

[1] The results of the repeated measures ANOVA indicated that race and sex cues separately influenced emotion categorization. Main effects of sex, F(1, 33) = 4.18, p = .049, η_p^2 = .11, and emotion, F(1, 33) = 10.31, p = .003, η_p^2 = .24, were moderated by two significant two-way interactions. A significant Target race × Emotion interaction emerged, F(1, 33) = 5.50, p = .025, η_p^2 = .14. Participants were faster to categorize happiness than anger on White faces, t(33) = 3.71, p = .001, d_z = 0.64, but no faster to categorize happiness or anger on Black faces, t(33) = 1.85, p = .074, d_z = 0.32. Further, a significant Target sex × Emotion interaction emerged, F(1, 33) = 16.54, p < .001, η_p^2 = .33, as participants were faster to categorize happiness than anger on female faces, t(33) = 5.29, p < .001, d_z = 0.91, but not on male faces, t(33) = 0.12, p = .906, d_z = 0.02. These effects were not further moderated by a Target race × Target sex × Emotion interaction, F(1, 33) = 0.48, p = .495, η_p^2 = .01, and no other effects were significant, Fs < 1.78, p > .191, η_p^2s < .05.

[2] Participant sex significantly moderated the Target Sex × Emotion interaction, F(1, 32) = 4.54, p = .041, η_p^2 = .12. Male participants were not significantly faster to categorize happiness or anger on faces of either sex; ts < 0.96, ps > .346, d_z = 0.43, although there was a numerical trend towards faster categorization of happiness than anger for male faces (happy – M = 611.79, SD =260.43; angry – M = 637.46, SD = 290.05) and female faces (happy –M = 624.26, SD = 232.72; angry – M = 640.45, SD = 261.09). Female participants were significantly faster to categorize happiness (M = 580.73, SD = 96.63) than anger (M = 635.75, SD = 108.41) on female faces, t(32) = 5.55, p < .001, d_z = 1.03, but no faster to categorize happiness (M = 624.79, SD =108.14) or anger (M = 621.78, SD = 120.44) displayed on male faces, t(32) = 0.27, p = .790, d_z = 0.05. Participant sex also significantly moderated the Target race × Emotion interaction, F(1, 32) = 10.07, p = .003, η_p^2 = .24. Male participants were significantly faster to categorize happiness (M = 588.72, SD = 226.26) than anger (M = 648.02, SD = 277.96) expressed on White faces, t(32) = 2.41, p = .022, d_z = 1.08, but not when expressed on Black faces (happy – M = 647.23, SD = 258.86, angry – M = 629.89, SD = 271.07), t(32) = 0.82, p = .419, d_z = 0.37. On the other hand, female participants were significantly faster to categorize happiness than anger on both White (happy – M = 598.00, SD = 93.95, angry – M = 628.96, SD = 115.42), t(32) = 3.02, p = .005, d_z = 0.56, and Black faces (happy – M = 607.51, SD = 107.49, angry – M = 628.57, SD = 112.55), t(32) = 2.39, p = .023, d_z = 0.44.

interaction, BF_{10} = 2797.283, but the category dominance model was around 107 times more likely to account for the data than the interactive model.

Experiment 1b.

Linear mixed-effects analysis. The results of Experiment 1b are depicted in Figure 2. As in Experiment 1a, the linear mixed-effects analysis including random effects of stimuli and participants produced a significant Target sex × Emotion interaction, b = 33.30, t(30.90) = 2.98, p = .006. This interaction emerged as participants were significantly faster to categorize happiness than anger on female faces, b = 46.64, t(30.98) = 4.52, p < .001, but not on male faces, b = 12.99, t(22.02) = 1.56, p = .133. The main effects of race, b = -16.49, t(32.52) = -2.35, p = .025, and emotion were also significant, b = 29.94, t(54.45) = 3.79, p < .001. The Target race × Emotion interaction was not significant, b = -16.15, t(28.95) = -1.45, p = .159 and all other effects including the three-way Target race × Target sex × Emotion interaction did not reach significance, *b*s < 10.98, *t*s < 0.78, *p*s > .441[3,4].

Bayesian repeated measures ANOVA. Similar to Experiment 1a, the model with the largest BF indicating best fit with the data was the model including main effects of race, sex, and emotion and the Target sex × Emotion interaction $BF_{10} = 9.33 \times 10^9$. There was also very strong evidence for the additive model including both the Target race × Emotion and Target sex × Emotion interactions, $BF_{10} = 8.32 \times 10^9$. Comparing these two models there was only anecdotal evidence that the category dominance model was a better fit for the data than the additive model,

[3] When looking at the data using a repeated measures ANOVA, the results of Experiment 1b were similar to those of Experiment 1a. Main effects of race, F(1, 64) = 19.90, p < .001, η_p^2 = .24, and emotion, F(1, 64) = 19.11, p < .001, η_p^2 = .23, were moderated by significant Target race × Emotion, F(1, 64) = 4.24, p = .044, η_p^2 = .06, and Target sex × Emotion interactions, F(1, 64) = 14.67, p < .001, η_p^2 = .19. Following up these significant two-way interactions indicated that participants were overall faster to categorize happiness than anger on both Black and White faces, *ts* > 2.41, *p*s < .020, d_zs > 0.40, but this happy advantage was significantly larger for White faces, t(64) = 2.06, p = .044, d_z = 0.26. They were also faster to categorize happiness than anger expressed on female faces, t(64) = 5.64, p < .001, d_z = 0.70, but the happy advantage for male faces was only marginal, t(64) = 1.79, p = .079, d_z = 0.22. No other effects (including the three-way Target race × Target sex × Emotion interaction) were significant, *F*s < 0.61, p > .438, η_p^2s < .01.

[4] With a larger and more gender balanced sample size, there was evidence of a significant Target sex × Participant sex interaction, F(1, 63) = 7.93, p = .006, η_p^2 = .11. Female participants were significantly faster to categorize expressions on female (M = 596.78, SD = 112.50) than on male faces (M = 605.36, SD = 114.81), t(63) = 2.09, p = .041, d_z = .36, but males tended to be faster to categorize expressions on male (M = 592.37, SD = 114.81) than on female faces (M = 600.30, SD = 112.50), t(63) = 1.90, p = .062, d_z = 0.34. Apart from this, participant sex did not significantly moderate any other main effects or interactions, *F*s < 1.74, *p*s > .192, η_p^2 < .03.

BF_{10} = 1.12. There was also very strong evidence for the interactive model including the three-way Target race × Target sex × Emotion interaction, $BF_{10} = 2.98 \times 10^8$, but again, the strongest model including the Target sex × Emotion interaction was 3.13 times more likely than the interactive model, providing positive evidence for the category dominance over the interactive model.

Experiment 2.

Linear mixed-effects analysis. As can be seen in Figure 3, both the sex and the age of the faces influenced emotion categorization. There were significant main effects of age, $b = -22.23$, $t(27.61) = -3.35$, $p = .002$, and sex, $b = -16.16$, $t(27.61) = -2.43$, $p = .022$, which were further moderated by significant Target sex × Emotion, $b = 38.66$, $t(29.46) = 2.74$, $p = .010$, and Target age × Emotion interactions, $b = 38.56$, $t(34.33) = 2.47$, $p = .019$. These interactions emerged as participants were faster to categorize happiness than anger on female faces, $b = 24.72$, $t(20.06) = 2.35$, $p = .029$, but not on male faces, $b = -13.65$, $t(16.62) = -1.13$, $p = .277$. They were also faster to categorize happiness on young adult faces, $b = 24.60$, $t(19.19) = 2.10$, $p = .049$, but not on older adult faces, $b = -13.84$, $t(21.30) = -1.15$, $p = .265$. The Target age × Target sex interaction was marginally significant, $b = 25.55$, $t(27.61) = 1.92$, $p = .065$, but participants were no faster to categorize expressions on female or male faces when the faces were old, $b = -17.55$, $t(14.49) = -1.72$, $p = .106$, or young, $b = -1.51$, $t(15.01) = -0.21$, $p = .835$. The three-way Target Age × Target sex × Emotion interaction did not reach significance, $b = 3.57$, $t(29.59) = 0.13$, $p = .900$, and the emotion main effect was not significant, $b = 5.34$, $t(35.66) = 0.67$, $p = .509$ [5,6].

[5] Repeated measures ANOVAs revealed significant main effects of target sex, $F(1, 36) = 9.65$, $p = .004$, $\eta_p^2 = .21$, and target age, $F(1, 36) = 31.68$, $p < .001$, $\eta_p^2 = .47$, that were moderated by three two-way interactions. Firstly, there was a significant Target age × Emotion interaction, $F(1, 36) = 14.21$, $p = .001$, $\eta_p^2 = .28$. Follow-up analyses indicated that participants were faster to categorize happiness than anger on young adult faces, $t(36) = 3.64$, $p = .001$, $d_z = 0.60$, but tended to be faster to categorize anger than happiness on older adult faces, $t(36) = 1.97$, $p = .056$, $d_z = 0.33$. There was also a significant Target sex × Emotion interaction, $F(1, 36) = 21.94$, $p < .001$, $\eta_p^2 = .38$. Participants were faster to categorize happiness than anger on female faces, $t(36) = 3.66$, $p = .001$, $d_z = 0.60$, but faster to categorize expressions of anger than happiness on male faces, $t(36) = 2.42$, $p = .021$, $d_z = 0.40$. A significant Target sex × Target age interaction also emerged, $F(1, 36) = 6.30$, $p = .017$, $\eta_p^2 = .15$, as participants were faster to categorize expressions on older adult female faces than on older adult male faces, $t(36) = 3.42$, $p = .002$, $d_z = 0.56$, but no faster or slower to categorize emotional expressions on young adult female and young adult male faces, $t(36) = 0.41$, $p = .686$, $d_z = 0.07$. No other effects (including the three-way Target sex × Target age × Emotion interaction) were significant, Fs < 0.73, $p > .396$, $\eta_p^2 < .03$.

Bayesian repeated measures ANOVA. The results of the Bayes Factor ANOVA indicated that the model including the three two-way interactions (Target age × Emotion, Target sex × Emotion, and Target age × Target sex) best fit the observed data, $BF_{10} = 3.25 \times 10^9$. Comparing this additive model to the interactive model including the three-way Target age × Target sex × Emotion interaction, $BF_{10} = 1.02 \times 10^9$, indicated that this additive model was 3.25 times more probable than the interactive model providing positive evidence for the additive over the interactive model. Further, analyses indicated that this additive model was also 3544 times more probable given the observed data, than the strongest model including only the Target age × Emotion interaction and 3618 times more probable given the observed data, than the strongest model including only the Target sex × Emotion interaction providing very strong evidence for the additive model over either sex or age category dominance models.

General Discussion

The aim of the current study was to determine how race and sex cues and age and sex cues concurrently influence emotion perception using speeded emotion categorization tasks. The repeated measures ANOVAs reported in the footnotes suggested an additive influence of race and sex and of age and sex on emotion categorization as there were two significant two-way interactions between social category cues and emotion (Target race/age × Emotion interaction and Target sex × Emotion interaction) in the absence of any higher order interactions in all experiments. However, linear mixed-effects analyses suggested that only the influences of sex and age on emotion categorization were likely to be generalizable beyond the stimuli and participants used in the current investigation. Additional Bayesian ANOVAs allowed us to compare the

---

[6] There was a marginally significant influence of participant sex on the Target sex × Emotion interaction, $F(1, 35) = 4.06$, $p = .052$, $\eta_p^2 = .10$. Female participants were significantly faster to categorize anger ($M = 525.14$, $SD = 103.23$) than happiness ($M = 540.04$, $SD = 92.71$) on male faces, $t(35) = 2.34$, $p = .025$, $d_z = 0.46$, but significantly faster to categorize happiness ($M = 506.17$, $SD = 85.68$) than anger ($M = 534.94$, $SD = 92.53$) on female faces, $t(35) = 4.24$, $p < .001$, $d_z = 0.83$. Male participants were not significantly faster to categorize happiness or anger on male (happy – $M = 552.49$, $SD = 142.54$; angry – $M = 544.66$, $SD = 158.71$) or female faces (happy – $M = 537.40$, $SD = 131.73$; angry – $M = 542.14$, $SD = 142.26$), *t*s < 0.80, $p = .429$, $d_z = 0.24$, though the response time pattern was in the same direction as for female participants. Participant sex did not moderate the interaction of target age and emotion or any of the other effects reported above, *F*s < 0.98, $p > .329$, η_p^2s < .03.

likelihood of the different statistical models and determine the model that best fit the observed data. The model that best fit the observed data differed between Experiments 1a and 1b, and Experiment 2. In Experiment 2, an additive model including all three two-way interactions best fit the data, but in Experiments 1a and 1b, the simpler model including only the Target sex × Emotion interaction was the strongest model best fitting the observed data. This pattern suggests dominance of the sex category over the race category but not the age category. This difference between the experiments is also observable in the patterns of response times, despite seemingly comparable statistical outcomes using traditional repeated measures ANOVAs. Although both Black females and Older adult females represent one positively evaluated (female) and one negatively evaluated category (Black or old), a happy advantage was present for Black female faces but not for old female faces supporting dominance of the sex category over race in Experiments 1a and 1b but not over age in Experiment 2.

Finding an additive influence of age and sex cues on emotion categorization is comparable to the findings of Lipp et al., (2015), who reported a significant happy categorization advantage for Black female faces that was smaller than the happy categorization advantage for White female faces. This finding is also consistent with the additive influence of race and sex cues on the categorization of angry and neutral faces reported by Smith and colleagues (2017). However, finding a dominant influence of sex over race cues (Experiments 1a and 1b) and an additive influence of age and sex on emotion categorization (Experiment 2) is inconsistent with the interactive influence of race and sex cues observed in the happy vs. neutral categorization task in the Smith et al. study, or the emotion onset/offset detection tasks in the Kang and Chasteen (2009) study.

Implications for Theory

Although category dominance of sex over race but not age has not yet been reported in the context of emotion perception, finding that sex has a stronger influence than race, but not age is consistent with findings from the social categorization literature. Cosmides, Tooby, and Kurzban

(2003) propose that sex and age, but not race, are ‘primary’ dimensions of person perception. They argue that encountering others who visually differ by their race is so evolutionarily recent that it is unlikely that there is a mechanism for specifically detecting race. On the other hand, people varying in age and sex would have been encountered throughout our evolutionary history making perception of these categories more ‘basic’. Cosmides and colleagues propose that race is processed by a more general mechanism sensitive to cues of coalition or kinship meaning that categorization by race (but not age or sex) is more flexible and sensitive to changes in context and can be easily overridden in the presence of additional coalitional information.

Evidence for this theory comes from the observation that participants tend to make social categorizations based on sex more so than on race but not age (Lieberman, Oum, & Kurzban, 2008; Kurzban, Tooby, & Cosmides, 2001; Pietraszewski, Cosmides, & Tooby, 2014; Stangor et al., 1992). For example, in ‘who said what’ tasks described in the introduction, participants made more within-sex-between-race misattributions than within-race-between-sex misattributions suggesting more of a reliance on sex cues (Stangor et al., 1992). In another study, participants made no more within-age-between-sex errors than within-sex-between-age errors suggesting equal reliance on age and sex information (Lieberman et al., 2008). Further, providing additional coalitional or kinship information about the faces encountered in the task had a strong influence on the degree to which faces were categorized by race but had less of an influence on categorization by sex (Kurzban et al. 2001; Pietraszewski et al., 2014). Considering the results of the current study within this framework, the primacy of sex and age cues over race cues seems to have an influence on the way that multiple social category cues combine to alter the speed of emotion perception.

Although the findings of the current study are consistent with this theory of social perception, there are other possible explanations for the results as there were a couple of differences between the experiments. For example, investigating the influence of old age cues and other race cues required the use of different face databases. There could have been differences between the databases in lighting or the presence of task irrelevant low level visual cues that may have influenced results. We attempted to mitigate this influence by editing the faces to make them as similar as possible. The results also suggest that this explanation is unlikely as the same pattern of results was observed for the White young adult faces drawn from the different databases across the three experiments. The results of the random effects analyses also suggested that the Target sex × Emotion interaction observed across tasks was robust and generalizable despite the faces coming from different sources. It was also possible that the difference between Experiment 1 and Experiment 2 was driven by something not specifically related to race or age that happens to systematically vary between race and age. For example, young adults tend to rate older adults as less attractive than young adults (Ebner, 2008), so differences in the attractiveness of the other-race and other-age faces could have played a role in the different patterns of results across experiments. Given that face age and attractiveness do co-vary for young adult observers, controlling for attractiveness would require selecting particularly attractive older adults or unattractive young adults limiting the generalizability of the findings, but future research may investigate the role of attractiveness in the interaction of multiple social categories with emotion.

Considering these findings together with the results of previous studies suggests that substantial variability exists in the way that multiple social category cues combine to influence emotion categorization. Some studies have found that the influence of multiple cues was additive, others have found an interactive influence and others have found evidence for category dominance. Variability in the interaction of multiple facial cues is not unique to studies focusing on the interaction of multiple social categories. Even the influence of a single category (like race) on emotion perception can change depending on seemingly small methodological decisions. For example, stimulus type (computer generated vs. photographic), presentation duration, and stimulus set size have been found to influence how race and emotion cues interact (Craig et al., 2012). Comparing the results of the repeated measures ANOVA and the linear mixed-effects analysis also indicated that the influence of race on emotion categorization in the current study was at least partly due to the properties of the stimuli used. Finding that multiple social category cues combine in different ways to influence emotion perception across tasks is also consistent with findings from the crossed-categorization literature which have produced additive patterns, interactive patterns, and various kinds of category dominance patterns depending on the nature of the task, which categories were crossed, and the outcome variables measured (e.g., Lieberman et al., 2008; Stangor et al., 1992; Urban & Miller, 1998).

At this point, one can only speculate as to why multiple social categories influence emotion perception in different ways in different tasks. Only a few studies have investigated these influences and the methods adopted in these studies varied along many dimensions. Examining the similarities and differences between the current study and past investigations, two potentially interesting sources of variability to consider are the emotional expressions presented within the task/s and the type of task selected. These methodological differences have previously been identified to promote a stronger influence of either social category based evaluations or stereotypes on emotion perception. As such, these methodological differences may have had flow on effects for how multiple social categories influenced emotion perception. For example, stereotypes exist for intersectional identities (e.g., Black woman) that are different from the single categories (Black and woman, Shields, 2008). When stereotypes more strongly drive the influence of social category cues on emotion, this may potentially lead to an interactive pattern. Evaluations of multiple social dimensions may be more easily summed resulting in additive or category dominance patterns depending on the strength and direction of the evaluations. Consistent with this suggestion, the influence of social category cues when categorizing happy vs. angry and neutral vs. angry faces have been attributed to the influence of evaluations (Craig & Lipp, 2017b; Craig et al., 2017; Hugenberg, 2005; Hugenberg & Sczesny, 2006) and an additive influence was found in previous studies using these methods (Lipp et al., 2015; Smith et al., 2017). On the other hand, in tasks where participants categorized happy vs. neutral expressions, or detected emotion onsets/offsets, stereotypes rather than evaluations have been suggested as the strongest source of influence (Craig

& Lipp, 2017b; Hugenberg & Bodenhausen, 2003) and an interactive pattern was observed when tasks of this nature were used to investigate the influence of multiple social categories on emotion perception (Kang & Chasteen, 2009; Smith et al., 2017). Further research will be required to investigate the contexts in which these different mechanisms are most likely to influence emotion perception to better understand why multiple social category cues combine in different ways in different circumstances.

Implications for Experimental Design

The current findings suggest the importance of investigating the concurrent influence of multiple category cues on emotion perception. Results demonstrate that the influence of two social categories on emotion perception is not necessarily the same as the combination of these cues when examined in isolation. Although there are now more studies investigating the combined influence of multiple social categories in social perception (e.g., Kang & Chasteen, 2009; Krumhuber & Manstead, 2011; Navarrete, Olsson, Ho, Mendes, Thomsen, & Sidanius, 2009; Smith et al., 2017), the findings of the current study reinforce the need for more research like this.

The contrast between the conclusions that could be drawn based on the frequentist repeated measures ANOVAs, the linear mixed-effects analyses, and the Bayesian ANOVAs was also interesting. In Experiments 1a and 1b, both the Target race × Emotion and Target sex × Emotion interactions were significant when analyzing the results using repeated measures ANOVAs, the Target race × Emotion interaction was no longer significant when incorporating both stimuli and participants as random effects, and Bayesian analyses suggested the model including only the Target sex × Emotion interaction was the best fit for the observed data. As such, one statistical method (repeated measures ANOVA) provided evidence for an additive pattern and the others (Bayesian ANOVA and linear mixed-effects analyses) provided evidence for a category dominance pattern. Finding these differences in the current study suggests that such a discrepancy may also exist in other studies that reported only traditional repeated measures ANOVAs. This study highlights the utility of Bayesian ANOVAs for future studies in social perception. Not only can they provide evidence either for or against the null hypothesis, but Bayesian analyses can be used in investigations interested in comparing the likelihood of different models to account for a pattern of results.

Conclusion

Together, the current results along with the previous literature demonstrate that multiple facial cues concurrently influence emotion processing, though the nature of the influence can vary. Using dual valenced emotion categorization tasks, the combined influence of two concurrently varying social categories was additive with some evidence a stronger influence of sex than race, but not age on emotion categorization. Future research will be required to better understand the factors that lead to different patterns of influence as well as to extend our understanding of whether these effects generalize to other racial and other age groups and beyond two categories.

Open Practices

Raw data from Experiment 1b are available from https://osf.io/rsmxb/. Experiments 1a and 2 were conducted earlier and at the time we did not seek permission from participants to make their data publicly available. Data for Experiment 1a and 2 are available on request from the first author.

Figures

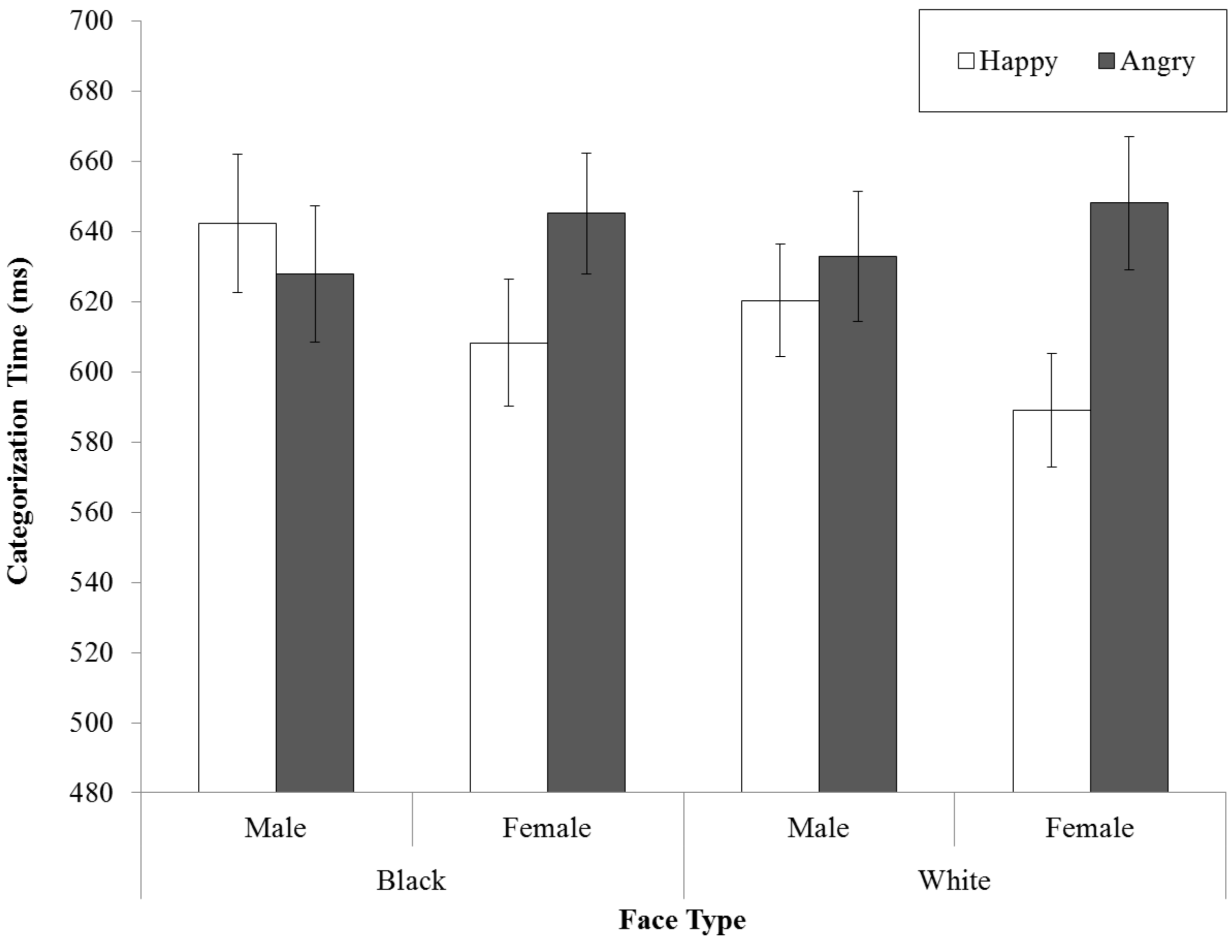

Figure 1. Mean response times from Experiment 1a for categorizing happy and angry expressions presented as a function of the race and sex of the face. Error bars represent one SEM.

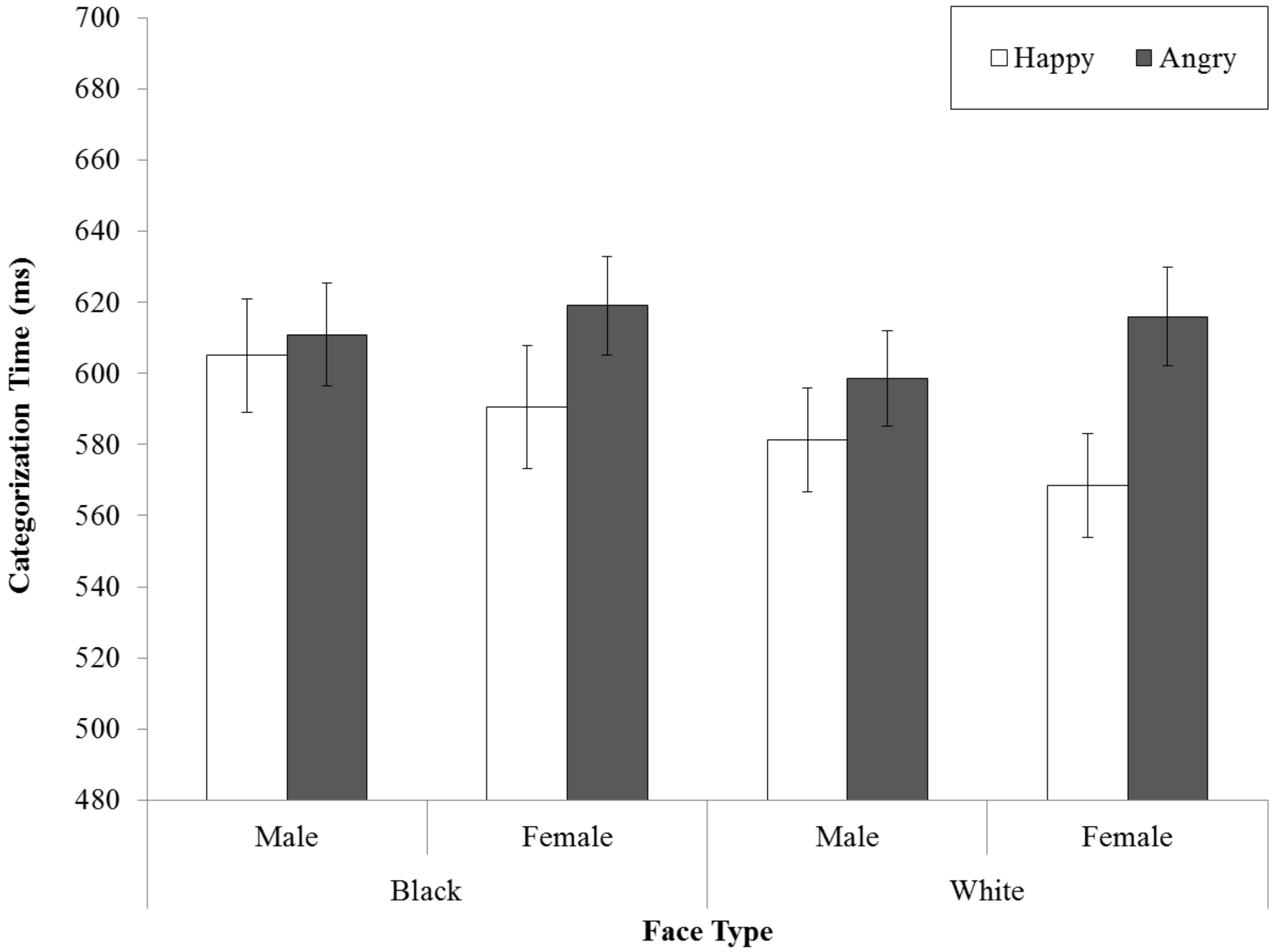

Figure 2. Mean response times from Experiment 1b for categorizing happy and angry expressions presented as a function of the race and sex of the face. Error bars represent one SEM.

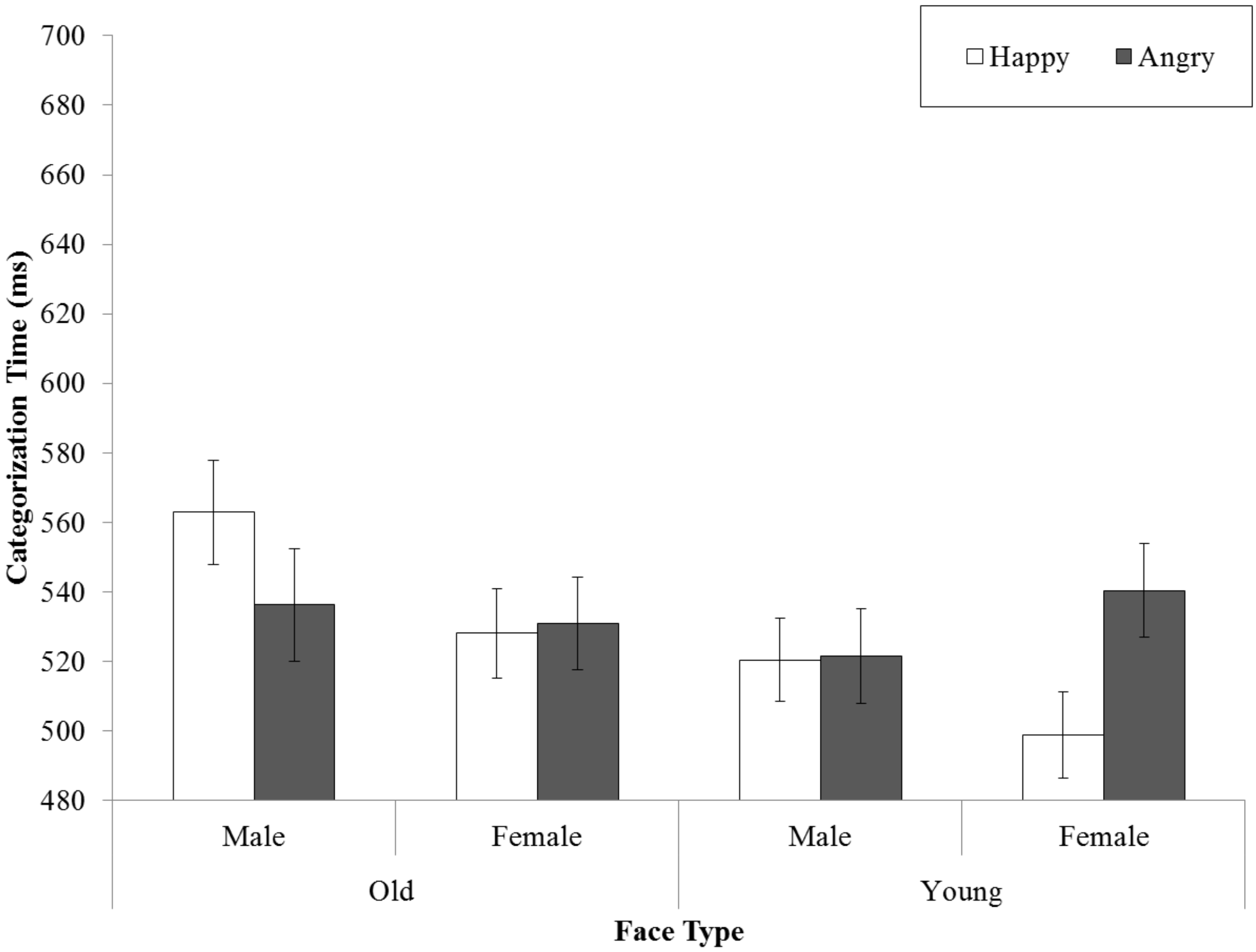

Figure 3. Mean response times from Experiment 2 for categorizing happy and angry expressions presented as a function of the age and sex of the face. Error bars represent one SEM.

References

Bates, D., Maechler, M., Ben Bolker, & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1-48. doi:10.18637/jss.v067.i01

Beaupré, M. G., & Hess, U. (2005). Cross-cultural emotion recognition among Canadian ethnic groups. *Journal of Cross-Cultural Psychology, 36*(3), 355-370. doi:10.1177/0022022104273656

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*(2), 179-190. doi:10.1037/0022-3514.92.2.179

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*(4), 657-663. doi:10.1016/j.jesp.2010.03.006

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology, 77*(3), 305-327. doi:10.1111/j.2044-8295.1986.tb02199.x

Cosmides, L., Tooby, J., & Kurzban, R. (2003). Perceptions of race. *Trends in Cognitive Sciences, 7*(4), 173-179.

Craig, B. M., Koch, S., & Lipp, O. V. (2017). The influence of social category cues on the happy categorisation advantage depends on expression valence. *Cognition and Emotion,* Advance online publication. doi:10.1080/02699931.2016.1215293

Craig, B. M., & Lipp, O. V. (2017a). Facial age cues and emotional expression interact asymmetrically: Age cues moderate emotion recognition. *Cognition and Emotion*, Advance online publication. doi:10.1080/02699931.2017.1310087

Craig, B. M., & Lipp, O. V. (2017b). The influence of facial sex cues on emotional expression recognition is not fixed. *Emotion, 17*(1), 28-39. doi:10.1037/emo0000208

Craig, B. M., Lipp, O. V., & Mallan, K. M. (2014). Emotional expressions preferentially elicit implicit evaluations of faces also varying in race or age. *Emotion, 14*(5), 865-877. doi:10.1037/a0037270

Craig, B. M., Mallan, K. M., & Lipp, O. V. (2012). The effect of poser race on the happy categorization advantage depends on stimulus type, set size, and presentation duration. *Emotion, 12*(6), 1303-1014. doi:10.1037/a0028622

Eagly, A. H., Mladinic, A., & Otto, S. (1991). Are women evaluated more favorably than men? An analysis of attitudes, beliefs, and emotions. *Psychology of Women Quarterly, 15*(2), 203-216. doi:10.1111/j.1471-6402.1991.tb00792.x

Ebner, N. C. (2008). Age of face matters: Age-group differences in ratings of young and old faces. *Behavior Research Methods, 40*(1), 130-136. doi:10.3758/BRM.40.1.130

Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES—A database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods, 42*(1), 351-362. doi:10.3758/brm.42.1.351

Fölster, M., Hess, U., & Werheid, K. (2014). Facial age affects emotional expression decoding. *Frontiers in Psychology, 5*(30). doi:10.3389/fpsyg.2014.00030

Forster, K. I., & Forster, J. C. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments & Computers, 35*(1), 116-124. doi:10.3758/bf03195503

Greenwald, A. G., Banaji, M. R., Rudman, L. A., Farnham, S. D., Nosek, B. A., & Mellott, D. S. (2002). A unified theory of implicit attitudes, stereotypes, self-esteem, and self-concept. *Psychological Review, 109*, 3-25. doi:10.1037//0033-295X.109.1.3

Hess, U., Adams Jr, R. B., Grammer, K., & Kleck, R. E. (2009). Face gender and emotion expression: Are angry women more like men? *Journal of Vision, 9*(12), 1-8. doi:10.1167/9.12.19

Hugenberg, K. (2005). Social categorization and the perception of facial affect: Target race moderates the response latency advantage for happy faces. *Emotion*, *5*(3), 267-276. doi:10.1037/1528-3542.5.3.267

Hugenberg, K., & Bodenhausen, G. V. (2003). Facing prejudice: Implicit prejudice and the perception of facial threat. *Psychological Science*, *14*(6), 640-643. doi:10.1046/j.0956-7976.2003.psci_1478.x

Hugenberg, K., & Sczesny, S. (2006). On wonderful women and seeing smiles: Social categorization moderates the happy face response latency advantage. *Social Cognition, 24*(5), 516-539. doi:10.1521/soco.2006.24.5.516

Inqusiti 4 [Computer Software]. (2015). Retrieved from http://wwww.millisecond.com

JASP Team. (2016). JASP (0.8.0.0) [Computer Software]. Retrieved from https://jasp-stats.org/

Johnson, K. L., Freeman, J. B., & Pauker, K. (2012). Race is gendered: How covarying phenotypes and stereotypes bias sex categorization. *Journal of Personality and Social Psychology, 102*(1), 116-131. doi:10.1037/a0025335

Johnston, R. A., Kanazawa, M., Kato, T., & Oda, M. (1997). Exploring the structure of multidimensional face-space: The effects of age and gender. *Visual Cognition*, *4*(1), 39-57. doi:10.1080/713756750

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychololgy*, *103*(1), 54-69. doi:10.1037/a0028347

Kang, S. K., & Chasteen, A. L. (2009). Beyond the double-jeopardy hypothesis: Assessing emotion on the faces of multiply-categorizable targets of prejudice. *Journal of Experimental Social Psychology*, *45*(6), 1281-1285. doi:10.1016/j.jesp.2009.07.002

Karnadewi, F., & Lipp, O. V. (2011). The processing of invariant and variant face cues in the Garner Paradigm. *Emotion*, *11*(3), 563-571. doi:10.1037/a0021333

Kass, R. E., & Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association, 90*(430), 773-795. doi:10.1080/01621459.1995.10476572

Kloth, N., Damm, M., Schweinberger, S. R., & Wiese, H. (2015). Aging affects sex categorization of male and female faces in opposite ways. *Acta psychologica, 158,* 78-86. doi:10.1016/j.actpsy.2015.04.005

Krumhuber, E. G., & Manstead, A. S. R. (2011). When memory is better for out-group faces: On negative emotions and gender roles. *Journal of Nonverbal Behavior, 35*(1), 51-61. doi:10.1007/s10919-010-0096-8

Kurzban, R., Tooby, J., & Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences, USA, 98*(26), 15387-15392. doi:10.1073/pnas.251541498

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2016). Package 'lmerTest' version 2.0-33 [Computer Software]. Retrieved from https://cran.r-project.org/web/packages/lmerTest/index.html

Leppänen, J. M., & Hietanen, J. K. (2003). Affect and face perception: Odors modulate the recognition advantage of happy faces. *Emotion, 3*(4), 315-326. doi:10.1037/1528-3542.3.4.315

Levin, D. T. (2000). Race as a visual feature: Using visual search and perceptual discrimination tasks to understand face categories and the cross-race recognition deficit. *Journal of Experimental Psychology: General, 129*(4), 559-574. doi:10.1037/0096-3445.129.4.559

Lieberman, D., Oum, R., & Kurzban, R. (2008). The family of fundamental social categories includes kinship: evidence from the memory confusion paradigm. *European Journal of Social Psychology, 38*(6), 998-1012. doi:10.1002/ejsp.528

Lipp, O. V., Craig, B. M., & Dat, M. C. (2015). A happy face advantage with male caucasian faces: It depends on the company you keep. *Social Psychological and Personality Science, 6*(1), 109-115. doi:10.1177/1948550614546047

Malatesta, C., & Izard, C. (1984). The facial expression of emotion: Young, middle-aged, and older adult expressions *Emotion in adult development* (pp. 253-273). CA: Sage Publications.

Meissner, C. A., & Brigham, J. C. (2001). Thirty years of investigating the own-race bias in memory for faces: A meta-analytic review. *Psychology, Public Policy, and Law, 7*(1), 3-35. doi:10.1037/1076-8971.7.1.3

Morey, R. D., & Rouder, J. N. (2015). BayesFactor (Version 0.9.11-3) [Computer Software].

Navarrete, C. D., Olsson, A., Ho, A. K., Mendes, W. B., Thomsen, L., & Sidanius, J. (2009). Fear extinction to an out-group face: The role of target gender. *Psychological Science, 20*(2), 155-158. doi:10.1111/j.1467-9280.2009.02273.x

Nosek, B. A., Greenwald, A. G., & Banaji, M. R. (2005). Understanding and using the Implicit Association Test: II. Method variables and construct validity. *Personality and Social Psychology Bulletin, 31*(2), 166-180. doi:10.1177/0146167204271418

Pietraszewski, D., Cosmides, L., & Tooby, J. (2014). The content of our cooperation, not the color of our skin: An alliance detection system regulates categorization by coalition and race, but not sex. *PLoS ONE, 9*(2), e88534. doi:10.1371/journal.pone.0088534

R Core Team. (2017). R: A language and environment for statistical computing. Retrieved from https://www.R-project.org

Rhodes, M. G., & Anastasi, J. S. (2012). The own-age bias in face recognition: A meta-analytic and theoretical review. *Psychological Bulletin, 138*(1), 146-174. doi:10.1037/a0025750

Rouder, J. N., Engelhardt, C. R., McCabe, S., & Morey, R. D. (2016). Model comparison in ANOVA. *Psychonomic Bulletin & Review, 23*(6), 1779-1786. doi:10.3758/s13423-016-1026-5

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology, 56*, 356-374. doi:10.1016/j.jmp.2012.08.001

Sacco, D. F., & Hugenberg, K. (2009). The look of fear and anger: Facial maturity modulates recognition of fearful and angry expressions. *Emotion*, *9*(1), 39-49. doi:10.1037/a0014081

Shields, S. A. (2008). Gender: An intersectionality perspective. *Sex Roles, 59*(5), 301-311. doi:10.1007/s11199-008-9501-8

Smith, J. S., LaFrance, M., & Dovidio, J. F. (2017). Categorising intersectional targets: An “either/and” approach to race-and gender-emotion congruity. *Cognition and Emotion, 31*(1), 83-97. doi:10.1080/02699931.2015.1081875

Stangor, C., Lynch, L., Duan, C., & Glas, B. (1992). Categorization of individuals on the basis of multiple social features. *Journal of Personality and Social Psychology, 62*(2), 207-218. doi:10.1037/0022-3514.62.2.207

Tottenham, N., Tanaka, J. W., Leon, A. C., McCarry, T., Nurse, M., Hare, T. A., . . . Nelson, C. (2009). The NimStim set of facial expressions: Judgments from untrained research participants. *Psychiatry Research*, *168*(3), 242-249. doi:10.1016/j.psychres.2008.05.006

Urban, L. M., & Miller, N. (1998). A theoretical analysis of crossed categorization effects: A meta-analysis. *Journal of Personality and Social Psychology, 74*(4), 894-908. doi:10.1037/0022-3514.74.4.894