Supplemental Material 3

S3. ANOVA's for all between-subjects factors as predictors and target emotion as outcome variable for S01 – S12.

Study 01

Study 01 had a 2 (BS, group: IG/OG) X 2 (BS, context: emotional/neutral) X 2 (BS, face: face/name). Participants were 187 Dutch students. Participants judged nine facial expressions (contempt, embarrassment, pride, surprise, fear, anger, happiness, sadness, and disgust) from stimulus set 1. An analysis of variance with group membership, context, and face as predictors and emotion (aggregated average score for all target emotions) as dependent variable showed that there were main effects of group membership ($F(1, 179) = 27.05$, $p = .0001$, $\eta_p^2 = .131$), context ($F(1, 179) = 364.52$, $p = .0001$, $\eta_p^2 = .671$), and face ($F(1, 179) = 236.61$, $p = .0001$, $\eta_p^2 = .569$). These main effects were qualified by a marginal interaction between group membership and context ($F(1, 179) = 3.90$, $p = .050$, $\eta_p^2 = .021$). The ingroup was perceived to feel more intense emotions than the outgroup in emotional contexts ($F(1, 179) = 5.32$, $p = .022$, $\eta_p^2 = .029$, $M_{ingroup} = 4.56$, $SE = .08$ vs. $M_{outgroup} = 4.31$, $SE = .09$), and in neutral contexts ($F(1, 179) = 25.20$, $p = .0001$, $\eta_p^2 = .123$, $M_{ingroup} = 3.14$, $SE = .09$ vs. $M_{outgroup} = 2.54$, $SE = .08$). The difference between emotional and neutral contexts was significant for the ingroup ($F(1, 179) = 154.74$, $p = .0001$, $\eta_p^2 = .464$, $M_{emotional} = 4.56$, $SE = .08$ vs. $M_{neutral} = 3.14$, $SE = .09$) and for the outgroup ($F(1, 179) = 210.71$, $p = .0001$, $\eta_p^2 = .541$, $M_{emotional} = 4.31$, $SE = .09$ vs. $M_{neutral} = 2.54$, $SE = .08$). Furthermore, there was an interaction between context and face $F(1, 179) = 145.55$, $p = .0001$, $\eta_p^2 = .448$). Emotional contexts were judged as more intense than neutral contexts when participants saw a face ($F(1, 179) = 24.39$, $p = .0001$, $\eta_p^2 = .120$, $M_{emotional} = 4.58$, $SE = .08$ vs. $M_{neutral} = 3.99$, $SE = .09$) and when they saw names

($F(1, 179) = 491.63$, $p = .0001$, $\eta_p^2 = .733$, $M_{emotional} = 4.30$, $SE = .08$ vs. $M_{neutral} = 1.68$, $SE = .08$). The difference between face and names was significant for emotional contexts ($F(1, 179) = 5.63$, $p = .019$, $\eta p^2 = .030$, $M_{face} = 4.58$, $SE = .08$ vs. $M_{name} = 4.30$, $SE = .08$) and neutral contexts ($F(1, 179) = 368.69$, $p = .0001$, $\eta_p^2 = .673$, $M_{face} = 3.99$, $SE = .09$ vs. $M_{name} = 1.68$, $SE = .08$). Since pairwise comparisons other than ingroup outgroup comparisons are secondary to our hypothesis they will not be discussed in the following part.

Study 02

Study 02 had a 2 (BS, group: IG/OG) X 2 (BS, context: emotional/neutral) X 2 (BS, intensity: high/low). 175 European-Americans from Mechanical Turk participated. Participants judged nine facial expressions (contempt, embarrassment, pride, surprise, fear, anger, happiness, sadness, and disgust) from stimulus set 1. An analysis of variance with group membership, context, and intensity as predictors and emotion (aggregated average score for all target emotions) revealed that there were main effects of group membership ($F(1, 167) = 16.44$, $p = .0001$, $\eta_p^2 = .090$), context ($F(1, 167) = 61.77$, $p = .0001$, $\eta_p^2 = .270$), intensity ($F(1, 167) = 5.69$, $p = .018$, $\eta p^2 = .033$). There was also an interaction between group membership and context ($F(1, 167) = 4.84$, $p = .029$, $\eta_p^2 = .028$) which was qualified by a three-way interaction between group membership, context, and intensity ($F(1, 167) = 5.49$, $p = .020$, $\eta_p^2 = .032$). In emotional contexts with low intensity displays there was no difference between perceived intensity of ingroup and outgroup expressions ($F(1, 167) = 0.15$, $p = .699$, $\eta_p^2 = .001$, $M_{ingroup} = 81.37$, $SE = 1.70$ vs. $M_{outgroup} = 82.30$, $SE = 1.70$). In emotional contexts with high intensity displays the ingroup was perceived to feel more intense emotions than the outgroup ($F(1, 167) = 5.03$, $p = .026$, $\eta_p^2 = .029$,

$M_{ingroup}$ = 87.07, SE = 1.63 vs. $M_{outgroup}$ =81.66, SE = 1.78). For low intensity displays the ingroup was perceived to feel more intense emotions than the outgroup in neutral contexts ($F(1, 167)$ = 16.29, p = .0001, η_p^2 = .089, $M_{ingroup}$ = 77.01, SE = 1.74 vs. $M_{outgroup}$ =66.97, SE = 1.78) and emotional contexts ($F(1, 167)$ = 4.61, p = .033, η_p^2 = .027, $M_{ingroup}$ = 77.76, SE = 1.59 vs. $M_{outgroup}$ =72.69, SE = 1.74).

Study 03

　　Study 03 had a 2 (BS, group: IG/OG) X 2 (BS, context: emotional/neutral) X 2 (BS, gender: female/male) design. 229 European-Americans from Mechanical Turk participated. Participants judged nine facial expressions (contempt, embarrassment, pride, surprise, fear, anger, happiness, sadness, and disgust) from stimulus set 1. An analysis of variance with group membership, context, and gender model as predictors and target emotion (aggregated average score for all target emotions) as outcome revealed that there were main effects of group membership ($F(1, 218)$ = 20.60, p = .0001, η_p^2 = .086), and context ($F(1, 218)$ = 64.68, p = .0001, η_p^2 = .229). These main effects were qualified by a marginal interaction between group membership and context ($F(1, 218)$ = 3.77 p = .053, η_p^2 = .017). Pairwise comparison suggests that ingroup and outgroup ratings did not differ in emotional contexts ($F(1, 218)$ = 3.44, p = .065, η_p^2 = .016, $M_{ingroup}$ = 83.74, SE = 1.26 vs. $M_{outgroup}$ =80.44, SE = 1.26). In neutral contexts the ingroup was perceived to feel more intense emotions than the outgroup ($F(1, 218)$ = 20.60, p = .0001, η_p^2 = .086, $M_{ingroup}$ = 75.98, SE = 1.28 vs. $M_{outgroup}$ =67.73, SE = 1.29). Furthermore the main effects were qualified by an interaction between group membership and gender of the model ($F(1, 218)$ = 4.60 p = .033, η_p^2 = .021). Pairwise comparison suggests that there was no difference between in- and outgroup ratings for male models ($F(1, 218)$ = 2.81, p = .095, η_p^2 = .013,

$M_{ingroup}$ = 79.12, SE = 1.27 vs. $M_{outgroup}$ 76.07, SE = 1.30). For female models the ingroup was perceived to feel more intense emotions than the outgroup ($F(1, 218)$ = 22.83, p = .0001, η_p^2 = .095, $M_{ingroup}$ = 80.61, SE = 1.27 vs. $M_{outgroup}$ =72.10, SE = 1.24).

Study04

Study 04 had a 2 (BS, group: IG/OG) X 2 (BS, gender: female/male) design. 107 European-Americans from Mechanical Turk participated. Participants judged eight facial expressions (embarrassment, pride, surprise, fear, anger, sadness, disgust, and neutral) from stimulus set 1. An analysis of variance with group membership, and gender model as predictors and target emotion as outcome (aggregated average score for all target emotions) revealed that there were as a main effect of group membership in which the ingroup was perceived to feel more intense emotions than the outgroup ($F(1, 103)$ = 25.67, p = .0001, η_p^2 = .199, $M_{ingroup}$ = 76.65, SE = 1.63 vs. $M_{outgroup}$ =64.69, SE = 1.71). Furthermore, there was a main effect of gender in which men were perceived to feel more intense emotions than women ($F(1, 103)$ = 7.01, p = .009, η_p^2 = .064, $M_{ingroup}$ = 73.79, SE = 1.58 vs. $M_{outgroup}$ =67.54, SE = 1.75).

Study 05

Study 05 had a 2 (BS, group: IG/OG) X 2 (WS, gender: female/male) design. 59 European Americans from Mechanical Turk participated. Participants judged six facial expressions (contempt, embarrassment, pride, surprise, fear, happiness) from stimulus set 2. An analysis of variance with group membership as predictor and target emotion as outcome revealed a main effect of group membership in which the

outgroup was perceived to feel more intense emotions than the ingroup ($F(1, 57) =$ 8.24, $p = .006$, $\eta_p^2 = .126$, $M_{ingroup} = 55.56$, $SE = 2.11$ vs. $M_{outgroup} = 64.19$, $SE = 2.14$).

Study 06

Study 06 had a 2 (BS, group: IG/OG) X 2 (WS, gender: female/male) design. 60 European Americans from Mechanical Turk participated. Participants judged six facial expressions (contempt, embarrassment, pride, surprise, fear, happiness) from stimulus set 2. An analysis of variance with group membership as predictor and target emotion as outcome revealed that there was no main effect of group membership on perceived intensity of emotions ($F(1, 57) = 2.33$, $p = .133$, $\eta_p^2 = .039$, $M_{ingroup} = 57.81$, $SE = 2.06$ vs. $M_{outgroup} = 62.12$, $SE = 1.93$).

Study 07

Study 07 had a 2 (BS, group: IG/OG) X 2 (BS, gender: female/male) design. Participants were 108 Dutch students. Participants judged nine facial expressions (contempt, embarrassment, pride, surprise, fear, anger, happiness, sadness, and disgust) from stimulus set 1. An analysis of variance with group membership, and gender as predictors and emotion (aggregated average score for all target emotions) as dependent variable showed that there were main effects of group membership ($F(1, 104) = 15.65$, $p = .0001$, $\eta_p^2 = .131$). This main effect as qualified by an interaction between group membership and gender of the model ($F(1, 104) = 18.52$, $p = .0001$, $\eta_p^2 = .151$). Pairwise comparison suggests that there were no differences between in- and outgroup for male models ($F(1, 104) = 0.56$, $p = .813$, $\eta_p^2 = .001$, $M_{ingroup} = 69.05$, $SE = 1.78$ vs. $M_{outgroup} = 69.60$, $SE = 1.49$). For female models, the ingroup was

perceived to feel more intense emotions than the outgroup ($F(1, 104) = 36.67$, $p = .0001$, $\eta_p^2 = .261$, $M_{ingroup} = 75.27$, $SE = 1.49$ vs. $M_{outgroup} = 62.16$, $SE = 1.57$).

Study 08

Study 08 had a 2 (BS, group: IG/OG) X 2 (BS, intensity: high/low) design. 193 European-Americans from Mechanical Turk participated. Participants judged nine facial expressions (contempt, embarrassment, pride, surprise, fear, anger, happiness, sadness, and disgust) from stimulus set 1. An analysis of variance with group membership, and intensity as predictors and emotion (aggregated average score for all target emotions) as dependent variable showed that there was main effect of group membership in which ingroup members were perceived to feel more intense emotions than outgroup members ($F(1, 189) = 37.72$, $p = .0001$, $\eta_p^2 = .166$, $M_{ingroup} = 73.60$, $SE = 1.17$ vs. $M_{outgroup} = 63.35$, $SE = 1.20$). Furthermore there was a main effect of intensity in which low intensity displays were perceived as less intense than high intensity displays ($F(1, 189) = 10.39$, $p = .001$, $\eta_p^2 = .052$, $M_{low} = 65.79$, $SE = 1.17$ vs. $M_{high} = 71.17$, $SE = 1.20$).

Study 09

Study 09 had a 2 (BS, group: IG/OG) X 2 (BS, gender: female/male) X 2 (BS, context: emotional/neutral). 409 European Americans from Mechanical Turk participated. Participants judged nine facial expressions (contempt, embarrassment, pride, surprise, fear, anger, happiness, sadness, and disgust) from stimulus set 1. An analysis of variance with group membership, context, and gender model as predictors and target emotion (aggregated average score for all target emotions) as outcome variable showed that there was main effects of group membership ($F(1, 401) = 36.96$,

$p = .0001$, $\eta_p^2 = .084$) and context ($F(1, 401) = 91.68$, $p = .0001$, $\eta_p^2 = .186$). These main effects were qualified by an interaction between group membership and context ($F(1, 401) = 7.92$, $p = .005$, $\eta_p^2 = .019$). Participants attributed more intense emotions to the ingroup than to the outgroup in both emotional contexts ($F(1, 401) = 5.54$, $p = .019$, $\eta_p^2 = .014$, $M_{ingroup} = 80.22$, $SE = 1.04$ vs. $M_{outgroup} = 76.74$, $SE = 1.05$) and in neutral contexts ($F(1, 401) = 38.12$, $p = .0001$, $\eta_p^2 = .087$, $M_{ingroup} = 73.02$, $SE = 1.08$ vs. $M_{outgroup} = 63.55$, $SE = 1.09$).

Study 10

Study 10 had a 2 (BS, group: IG/OG) x 2 (BS, nationality: IG/OG) x 2 (BS, set: 1/2) design. Contrary to earlier studies we did not only use European and Arab models as either in-or outgroup but either presented white models as ingroup (US) or outgroup (European) or Arab models as ingroup (US) or outgroup (Arab). 502 European-Americans from Mechanical Turk participated. Participants judged six facial expressions (contempt, embarrassment, pride, surprise, fear, happiness) from stimulus set 1 and set 2. An analysis of variance with ethnic group, national group, and stimulus set as predictors and target emotion (aggregated average score for all target emotions) revealed main effects of ethnic group ($F(1, 494) = 4.21$, $p = .041$, $\eta_p^2 = .008$) and stimulus set ($F(1, 494) = 7.52$, $p = .006$, $\eta_p^2 = .015$). These main effects were qualified by an interaction between ethnic group and stimulus set ($F(1, 494) = 49.83$, $p = .0001$, $\eta_p^2 = .092$). Pairwise comparison suggests that ingroup members were perceived to have more intense emotions than outgroup members in set 1 ($F(1, 494) = 42.01$, $p = .0001$, $\eta p^2 = .078$, $M_{ingroup} = 59.96$, $SE = 1.39$ vs. $M_{outgroup} = 47.20$, $SE = 1.41$). In set 2, outgroup members were perceived to have more intense emotions than ingroup members ($F(1, 494) = 12.39$, $p = .0001$, $\eta_p^2 = .024$, $M_{ingroup} =$

46.23, $SE = 1.41$ vs. $M_{outgroup} = 53.25$, $SE = 1.41$). Furthermore, there was an interaction between stimulus set and nationality ($F(1, 494) = 5.95$, $p = .015$, $\eta_p^2 = .012$). Pairwise comparison suggests that there is no difference between national in- and outgroups for set 1 ($F(1, 494) = 0.74$, $p = .392$, $\eta_p^2 = .001$, $M_{ingroup} = 52.74$, $SE = 1.38$ vs. $M_{outgroup} = 54.43$, $SE = 1.41$). For set 2 national outgroups were perceived as less intense than national ingroups ($F(1, 494) = 6.67$, $p = .010$, $\eta_p^2 = .013$, $M_{ingroup} = 52.31$, $SE = 1.39$ vs. $M_{outgroup} = 47.16$, $SE = 1.43$).

Study 11

Study 11 had a 2 (BS, group: IG/OG) x 2 (BS, nationality: UK/Arab) x 2 (BS, set: 1/2) design. As in Study 12 we did not only use European and Arab models as either in-or outgroup but either presented white models as ingroup (UK) or outgroup (Arab). 441 British crowdflower workers participated. Participants judged six facial expressions (contempt, embarrassment, pride, surprise, fear, happiness) from stimulus set 1 and set 2. An analysis of variance with ethnic group, national group, and stimulus set as predictors and target emotion (aggregated average score for all target emotions) as outcome revealed main effects of ethnic group ($F(1, 433) = 27.84$, $p = .0001$, $\eta_p^2 = .060$) and stimulus set ($F(1, 433) = 16.45$, $p = .0001$, $\eta_p^2 = .037$). These main effects were qualified by an interaction between ethnic group and set ($F(1, 433) = 13.47$, $p = .0001$, $\eta_p^2 = .030$). Pairwise comparison suggested that people perceive more intense emotions in ethnic ingroup members than in ethnic outgroup members in set 1 ($F(1, 433) = 40.87$, $p = .0001$, $\eta_p^2 = .086$, $M_{ingroup} = 52.31$, $SE = 1.39$ vs. $M_{outgroup} = 47.16$, $SE = 1.43$). In set 2, there were no differences between in- and outgroup members ($F(1, 433) = 1.26$, $p = .262$, $\eta_p^2 = .003$, $M_{ingroup} = 50.88$, $SE = 1.43$ vs. $M_{outgroup} = 48.62$, $SE = 1.41$).

Study 12

Method

Participants and design. A total of 729 white British participants completed the study (375 females, 342 males, 1 other). The average age of this sample was 35.54 years (range 18 – 72). Participants were recruited via Prolific (www.prolific.ac). The study had a 2 (Group: Ingroup/Outgroup) x 2 (Gender: Female/male) x 3 (Intensity: High/Medium/Low) between subjects design. We presented nine emotions (contempt, embarrassment, pride, surprise, fear, anger, happiness, sadness, and disgust).

Stimuli, measures and procedure. We made use of a newly developed stimulus set (ADFES-BIV, Wingenbach, Ashwin, & Broswin, 2016) that provides dynamic stimuli at three different levels of intensity (low, medium, and high). The most important advantage of this set is that it provides standardized versions of different levels of intensity, which was achieved by extracting consecutive frame sequences that all start with a neutral expression. Every participant saw a total of 27 trials (9 emotions x 3 models). For each trial participants had to indicate to what extent four proposed emotions were present in the face, on a scale from 0 (not at all) to 100 (very much). Each video contained 26 frames with a frame rate of 25/sec, resulting in videos that had a 1040 ms length. All stimuli were created by Tanja Wingenbach and were based on our own stimulus set, used in most of the previous studies (ADFES). Procedure.

Results

We executed an analysis of variance with Group Membership (European vs. Arab models), Gender (female vs. male), and Intensity (low, medium, high) as predictors and emotion intensity (aggregated average of all target emotions) as outcome variable. This analysis showed a main effect of Group Membership (F(1,

706) = 26.15, p = .0001, η_p^2 = .036), with Europeans attributing more intense emotions to ingroup members ($M_{ingroup}$ = 67.34, SE =.602) than to outgroup members ($M_{outgroup}$ = 62.98, SE =.603). Furthermore, we found a main effect of Intensity ($F(2, 706)$ = 88.93, p = .0001, η_p^2 = .201). As expected, high intensity displays were perceived as most intense (M_{high} = 72.41, SE =.745), followed by intermediate displays (M_{medium} = 64.63, SE =.729), and low intensity displays (M_{low} = 58.43, SE =.740). All differences between intensity levels were statistically significant based on pairwise comparisons (low vs. medium M = -6.19, SE = 1.039, p = .0001; low vs. high M = -13.97, SE = 1.050, p = .0001; medium vs. high M = -7.78, SE = 1.042, p = .0001). However, the two way interactions between Group Membership and Intensity ($F(2, 706)$ = .812, p = .444, η_p^2 = .002) did not yield significant effects.

Furthermore, there was a significant interaction between Group Membership and Gender ($F(1, 706)$ = 4.04, p = .045, η_p^2 = .006). The difference between ingroup and outgroup was significant for both female models ($M_{ingroup}$ = 66.50, SE = .851 vs. $M_{outgroup}$ = 63.86, SE = .856; M_{diff} = 2.64, SE = 1.207, p = .029) and male models ($M_{ingroup}$ = 68.17, SE = .851 vs. $M_{outgroup}$ = 62.10, SE = .851; M_{diff} = 6.07, SE = 1.204, p = .0001). The differences between female (M_{diff} = 1.67, SE = 1.203, p = .167) and male models (M_{diff} = 1.76, SE = 1.207, p = .145) within each ethnic group were non-significant. There was no main effect of Gender ($F(1, 706)$ = 0.003, p = .955, η_p^2 = .000) and all remaining 2- and 3-way interaction effects were non-significant.