Implicit responses in the judgment of attractiveness in faces with differing levels of makeup

Article

Accepted Version

7   **Implicit Responses in the Judgement of Attractiveness in Faces with Differing**
8   **Levels of Makeup**

9

10   **Abstract**

11   Makeup is a form of body art which has been used for over 7000 years and is present in

12   the great majority of human cultures, often used to enhance facial attractiveness and to

13   accentuate features that represent femininity. This study examined how cumulative

14   levels of facial makeup influenced approach and avoidance tendencies and on facial

15   muscle responses associated with emotional response obtained through facial

16   electromyography (EMG) in a passive viewing task. Experiment 1 employed the

17   joystick variant of the approach-avoidance task, where 30 subjects categorised female

18   faces by visual orientation (portrait/landscape) in 7 cumulatively-added makeup levels.

19   In Experiment 2, facial EMG was recorded from 40 subjects in the passive viewing of

20   the same images. The present study shows that makeup application modulates implicit

21   responses and reveals two distinct implicit preferences, behavioural and affective, with

22   a male behavioural preference for heavy eye cosmetics, a female behavioural preference

23   for light makeup, and an overall affective preference in both men and women for

24   makeup accentuating visual contrast in the eye and mouth regions. These results are

25   consistent with the conception that perceptual cues underlying cosmetic enhancement

26   are key determinants in aesthetic facial preferences.

27

28   *Keywords***:** cosmetics, facial attractiveness, facial electromyography, approach-

29   avoidance task.

30   *Abstract word count*: 190

31   *Text word count*: 7977

Introduction

33   Judgements of facial attractiveness have been shown to be remarkably consistent

34   between individuals and cultures (Langlois et al., 2000), in marked contradiction to the

35   commonly-held belief that "beauty is in the eye of the beholder". Recent studies have

36   shed further light on the perceptual bases of facial attractiveness, by using techniques

37   traditionally employed in studies of the recognition of facial expressions to further

38   elucidate common perceptual cues for the evaluation of facial attractiveness, moving

39   beyond the concepts laid out in human ethology and evolutionary psychology of facial

40   symmetry, averageness and skin texture as the principal determinants of facial

41   attractiveness in humans (Fink & Neave, 2005; Rhodes, 2006; Little et al., 2011).

42   One key area of interest is the extent to which the application of cosmetics enhances

43   female facial attractiveness in the face of several conflicting claims from the health and

44   beauty industry. Female faces are judged to be significantly more attractive following

45   the application of differing layers of eye, lip and full-face makeup (Mulhern et al.,

46   2003), and this effect is enhanced for female observers compared to males. Mulhern et

47   al. (2003) investigated the effect of makeup on female facial attractiveness evaluated by

48   male and female participants using five cosmetic conditions: no makeup, foundation

49   only, eye makeup only, lip makeup only, and full facial makeup. Their results showed

50   that in explicit ratings of attractiveness, faces with full makeup were rated as more

51   attractive than those with no makeup or less makeup (e.g. lipstick only). While their

52   design allowed for the evaluation of cosmetic enhancement by region, in naturalistic

53   settings women more commonly combine cosmetic products applied to the face as a

54   whole. As such, both the number and combination of makeup products used should be

55   addressed when investigating the cosmetic enhancement of facial attractiveness.

56   In addition, women judged to be more attractive when wearing makeup are also

57   perceived as healthier, more confident and more professionally successful by male and

58   female participants (Nash et al., 2006), as part of a generalised "attractiveness halo"

59   effect for attractive faces (for a review, see Zebrowitz & Montepare, 2008). However,

60   differences in judgements of facial attractiveness due to makeup are statistically

61   negligible in comparison to differences due to identity (Jones & Kramer, 2016) and thus

62   the enhancement effect of makeup on overall attractiveness appears to be slight and may

63   potentially be due to an interaction with individual physiognomy.

64   Facial attractiveness is highly correlated with femininity or sexual dimorphism (Koehler

65   et al., 2004; Little et al., 2011). Perceptual analyses of female and male faces have

66   highlighted differences in luminance contrast in facial regions such as the eyes and

67   mouth as one of the prime determinants of sexual dimorphism (judging whether a face

68   is more feminine or masculine) regardless of face gender (Russell, 2003; 2009).

69   Accordingly, makeup appears to lead to an enhancement of this contrast effect with

70   products such as eyeshadow that accentuate the femininity of the face. At the same time,

71   products such as foundation are used to mask imperfections and smooth skin texture,

72   leading to an increase in overall facial symmetry and averageness which are seen as

73   common evolutionary-defined cues for assessing attractiveness (Rhodes, 2006; Russell,

74   2010). This is consistent with attractiveness research, which has highlighted specific

75   facial characteristics as determinants of female facial attractiveness beyond facial

76   symmetry, such as high cheekbones, large eyes and lips, thin eyebrows, and small noses

77   and chins (Cunningham et al., 1995; Baudouin & Tiberghien, 2004).

78   Previous studies have shown a remarkable consistency in male preferences for female

79   facial attractiveness across different races and cultures (Cunningham et al., 1995;

80   Thornhill & Grammer, 1999; Fink & Penton-Voak, 2002). One of the main divergences

81   in attractiveness preferences across race is related to the degree of sexual dimorphism

82   that female faces display (Penton-Voak et al., 2004). White British and Japanese men

83   show significant preference for more feminine female faces of the same race, while

84   Jamaican men show less preference for sexual dimorphism, evaluating female faces

85   with greater masculinity as significantly more attractive than feminised morphs (Penton-

86   Voak et al., 2004). Overall however, increased sexual dimorphism is preferred in female

87   but not male faces (Morrison et al., 2010), primarily as an indicator of health and

88   increased fertility (Law-Smith et al., 2006).

89   The majority of studies conducted to date on the effect of makeup on judgements of

90   facial attractiveness have employed an overt rating scale, with participants assigning

91   explicit scores to faces on the basis of their conscious perception of the individual's

92   attractiveness, similar to the design by Osborn (1996). There have been relatively few

93   studies conducted to date measuring implicit responses to faces with differing levels of

94   makeup. One of the few studies to do so (Richetin et al., 2004), used the Implicit

95   Association Test to test differences in reaction time in response to pairings of female

96   faces with and without makeup and positive and negative stimuli such as personality

97   traits, pleasant/unpleasant words and professions of high-/low-social status. They found

98   that faces with makeup were associated with positive personality traits and high-status

99   professions more than faces with no makeup, similarly to the results from Nash et al.

100 (2006). As makeup had no effect on reaction time in response to pleasant and

101   unpleasant words, the implicit processing of makeup may be dependent on social

102   context, and not merely affected by the emotional valence of the stimuli.

103   Another method for measuring implicit response, traditionally in the context of

104   emotional valence, is facial electromyography (EMG). Facial EMG is capable of

105   delivering great sensitivity and accuracy in the detection of the movement of facial

106   muscles associated with emotional expressions such as the *M. corrugator supercilii*

107   (associated with frowning and negative affect) and the *M. zygomaticus major*

108   (associated with smiling and positive affect). Facial EMG can detect face muscle

109   activations that are so subtle that they are not visible in the face due to the overlaying

110   fatty tissue and skin (Rinn, 1984). In addition, recordings from facial EMG can capture

111   responses to low-intensity emotional stimuli and even in situations where the participant

112   has no conscious awareness of producing an emotional response (Cacioppo et al., 1986;

113   Dimberg et al., 2000). An example of face muscle activity that can be measured using

114   EMG despite a lack of participants' awareness of the muscle activations is the

115   phenomenon termed 'facial mimicry' (see a review by Hess & Fischer, 2014). That is,

116   presenting participants with stimuli portraying facial emotional expressions on a

117   computer screen will produce only a subliminal perception of muscle activation in the

118   participant in accordance to the observed facial expression. Facial mimicry has been

119   demonstrated in the corrugator when participants observe emotional expressions of

120   negative valence and in the zygomaticus when observing positively valenced emotional

121   expressions (Achaibou et al., 2008; Dimberg, 1982; Dimberg & Thunberg, 1998;

122   Lundqvist, 1995; Lundqvist & Dimberg, 1995), and increased levator activity when

123   observing facial expressions of disgust and greater frontalis activation when observing

124   facial expression of fear and surprise (Lundqvist, 1995; Lundqvist & Dimberg, 1995).

125   That facial muscle activity associated with emotional facial expressions can be

126 measured without participants being aware of these activations makes facial EMG an

127 ideal implicit measure. Facial EMG holds great utility as an implicit measure of

128 individual affect, since face muscle activations (as measured via EMG) can reflect

129 underlying emotional states. For example, when watching pictures related to positive

130 and negative emotions, participants' EMG activity will increase in the zygomatic and

131 corrugator regions respectively, and subjective affective valence ratings are in line with

132 these increased muscle activations (e.g. Larsen et al., 2003). Facial EMG thus allows to

133 measure participants' affective responses independent of participants awareness of their

134 emotional states or accompanying facial muscle activations.

135 Employing facial EMG in response to faces of varying facial attractiveness results in a

136 modulation of activity mainly in the zygomatic and corrugator regions (Hazlett &

137 Hoehn-Saric, 2000; Gerger et al., 2011) and over the *levator labii superioris* muscle

138 (Principe & Langlois, 2011), associated with emotional reactions of disgust. Hazlett and

139 Hoehn-Saric (2000) found an interesting sex difference in facial EMG response to facial

140 attractiveness, with female subjects revealing increased corrugator response when

141 presented with highly-attractive female faces and greater zygomatic response when

142 viewing highly-attractive male faces. Overall however, there appears to be a linear

143 negative correlation between facial attractiveness and corrugator and levator response

144 (Principe & Langlois, 2011), and, to a lesser extent, a positive correlation between

145 attractiveness and zygomatic response (Gerger et al., 2011), regardless of the gender of

146 the subject. Facial EMG modulation has been shown to be consistent with explicit

147 ratings of facial attractiveness (Gerger et al., 2011) and sexual arousal (Hazlett &

148 Hoehn-Saric, 2000), but has not yet been compared to other measurements of implicit

149 response, such as approach/avoidance behaviour. Though, facial EMG might reveal

150 affective responses beyond what can be assessed with explicit measures. Another more

151 recent study (Tagai et al., 2017) investigated the effect of different levels of makeup on

152 amplitude differences of ERP components associated with face processing such as the

153 N170 and VPP. The authors observed that the processing of faces with light makeup

154 was accompanied by a decrease in N170 and VPP amplitude as compared to faces with

155 heavy makeup. This result was consistent with the explicit classification of facial

156 attractiveness, with slightly softer faces being evaluated as more attractive than faces

157 with heavy makeup, possibly due to the greater fluency and ease of visual processing of

158 faces with lighter makeup.

159 Several studies have investigated trustworthiness and facial emotion through an

160 approach/avoidance paradigm, whether through measurements of amygdalar activation

161 (Todorov et al., 2008) or manipulation of a virtual manikin or physical joystick (Heuer

162 et al., 2007; Krieglmeyer & Deutsch, 2010). As facial attractiveness and

163 approachability/trustworthiness are highly correlated (Todorov, 2008; Sofer et al.,

164 2015), additional feedback from an approach/avoidance measure such as joystick

165 position may provide useful data in response to emotional modulation by both

166 attractiveness and makeup levels. Concomitant effects of facial attractiveness on

167 emotional processing have long been established (Nakamura et al., 1998); several areas

168 associated with reward and positive-valenced emotions, such as the orbitofrontal cortex

169 and amygdala, are also activated when viewing and categorising faces by level of

170 attractiveness (Winston et al., 2007). In addition, this activation frequently occurs even

171 when the task is not specific to categorisation of facial attractiveness (Chatterjee et al.,

172 2009), indicating that the reward-inducing properties of attractive faces are at least

173 partly automatised. More recently, approach-related behaviour has been directly linked

174 with the reward value of faces explicitly categorised as more attractive in both male and

175 female participants (Kramer et al., 2020), with greater physical "lean" and approach

176 response towards attractive than unattractive female faces even in the absence of active

177 task demands. As emotionally-expressive facial cues have been shown to lead to the

178 modulation of task-selective motor response, the authors argue for a similar modulatory

179 effect from facial cues signalling attractiveness (Kramer et al., 2020).

180 This close correlation between beauty and emotional and behavioural response opens up

181 several possibilities for testing the perception of facial attractiveness using implicit

182 measures more commonly used for the analysis of affective valence or intensity

183 (Chatterjee & Vartanian, 2016). Interestingly, an interaction between positive emotional

184 feedback and visual fluency may lie behind one of the main determinants of facial

185 attractiveness, prototypicality or averageness. Winkielman et al. (2006) found that more

186 prototypical random-dot patterns were categorised more quickly and consistently rated

187 as more attractive than less prototypical displays, together with increased zygomatic

188 EMG response, revealing the close association between the increased perceptual fluency

189 of prototypical stimuli and higher measures of attractiveness and positive affect.

190 Principe and Langlois (2012) investigated the effect of face prototypicality on emotional

191 response when categorising faces by attractiveness, and found that previous

192 familiarisation with human-chimpanzee morphed faces led to a shift in preferences;

193 with human-chimpanzee morphs categorised as more attractive and with a

194 correspondent increase in zygomatic activity in those participants than for participants

195 who received no previous familiarisation. This reveals that our internal prototypes for

196 facial attractiveness are both malleable and subject to previous cultural experience.

197 Taken together, these results reveal a diverse set of social and cognitive mechanisms

198 underlying the perception of makeup, ranging from a generalised "halo effect", with

199 more heavily-applied cosmetics associated with greater professional success,

200 competence, and even physical health and wellbeing (Nash et al., 2006; Richetin et al.,

201   2004), to visual cues signalling femininity, youth and attractiveness, particularly a skin-

202   smoothing effect of foundation (Russell, 2010) and an increase in luminance contrast

203   provided by eye makeup (Russell, 2003, 2009) associated with sexual dimorphism.

204   Some researchers have argued for the integration of these components as part of an

205   'extended phenotype' of cosmetic use, as a cultural tool to increase one's social and

206   sexual success (Etcoff et al., 2011; Mileva et al., 2016).

207   Present study: The implicit techniques of facial EMG and AAT response were

208   employed as dependent measures in the present study to investigate the effect of

209   cumulative levels of applied makeup on participants' implicit emotional and

210   behavioural responses to facial attractiveness and ethnicity while conducting a

211   perceptual categorisation task. We sought to investigate the implicit perception of

212   makeup by testing participants' responses to varying conditions of makeup, from a basic

213   layer of foundation to the "heavier" application of eyeshadow. As makeup is commonly

214   applied in different stages in response to social context, with foundation and lipstick

215   used in more "everyday" contexts than other products, we designed a set of facial

216   stimuli containing the cumulative addition of cosmetic products from a base of

217   foundation and lipstick to the greater visual contrast of pencilling, mascara, eyeliner and

218   eyeshadow respectively. The inclusion of gradually-applied makeup levels in the stimuli

219   also sought to differentiate participants' implicit response to the qualitative changes to

220   facial features and configuration caused by different makeup products, within a

221   naturalistic setting. Previous studies investigating implicit or physiological responses to

222   makeup differences have employed either a no-makeup/makeup design (Richetin et al.,

223   2004) or a no/light/heavy makeup design (Tagai et al., 2016; 2017), which may not have

224   shown sufficient sensitivity towards differences in intermediate levels of makeup

225   application.

226   As EMG markers associated with negative emotional response such as corrugator and

227   levator activation have previously shown to be negatively correlated with explicit face

228   attractiveness ratings (Principe & Langlois, 2011), and conversely zygomatic activation,

229   associated with positive affect, has been shown to be positively correlated with explicit

230   face attractiveness ratings (Gerger et al., 2011), we opted to include these specific

231   muscle sites in the design of the current study. Additionally, we included the frontalis

232   muscle (*M. Frontalis, pars lateralis*) as a site of EMG response, due to the previously

233   reported association of this muscle with the inducement of stress (Kukde & Neufeld,

234   1994), and negative affect (Cacioppo et al., 1986) similarly to the corrugator. We

235   hypothesised that the increased cosmetic enhancement of facial attractiveness would

236   lead to a decrease in activation at the three muscle sites associated with negative affect

237   and emotional response (corrugator, levator and frontalis), and a concurrent increase in

238   zygomatic activation, associated with positive affect. In addition, in line with studies

239   reporting a perceptual preference for increased visual fluency in faces with light

240   compared to heavy makeup (Tagai et al., 2016; 2017), we expected to observe a drop-

241   off in zygomatic activation following intermediate levels of makeup application, and an

242   increase in corrugator activation in response to heavy levels of makeup application,

243   consistent with previously-reported visual fluency effects in aesthetic preference

244   (Gerger & Leder, 2015).

245   The previously-reported effect of the cosmetic enhancement of facial attractiveness in

246   an implicit experimental design (IAT; Richetin et al., 2004) and a recent study

247   highlighting the utility of the approach-avoidance task (AAT) as an implicit measure of

248   facial attractiveness (Kramer et al., 2020), led us to include AAT response as a

249   dependent variable in our study. That is, beyond an affective preference for certain

250   makeup products or combinations of such products, we tested a behavioural preference

251   for cosmetic products in terms of participants' reaction time towards engaging more

252   closely or more distally with cosmetically-enhanced facial stimuli. Similarly to the

253   EMG response, we expected to observe a faster approach and slower avoidance time to

254   intermediate levels of makeup application, and a slower approach and faster avoidance

255   time to both no makeup and heavy levels of makeup application, driven primarily by

256   visual fluency effects.

257

258   **Experiment 1 - Methods**

259   **Subjects**

260   The sample of the approach/avoidance task in Experiment 1 was composed of 15

261   women and 15 men, heterosexual, between 19 and 27 years (M: 21.77, SD: 2.743) and

262   of Caucasian ethnicity. Sample size and composition were calculated based on the effect

263   sizes reported in previous related research studies (Tagai et al., 2016; 2017; N = 38-45,

264    $\eta^2 \approx 0.3$), using G*Power 3.1.9.2 (Faul et al., 2007). All participants were informed

265   about the procedure but not informed about the specific objective in the study

266   (approach-avoidance response to different cosmetics), and signed an informed consent

267   form indicating their willingness to participate in the experiment. Participant

268   recruitment took place through digital media such as social media and scientific

269   research recruitment sites. Participants enrolled as students received course credit for

270   their participation in the experiment. The study adhered to the Declaration of Helsinki

271   guidelines and was approved by the institutional ethics committee and national ethics

272   committee.

273

274   **Procedure**

275 ***Stimuli***

276 The experimental stimuli were composed of 126 emotionally-neutral images of 18

277 female faces with 7 different levels of makeup (no makeup, added foundation, added

278 lipstick, added mascara, added pencilling, added eyeliner, added eyeshadow) from a

279 previously-constructed face database. The database used in the current study is

280 composed of facial photographs taken of 60 women aged 19 – 32, in a frontal pose, and

281 of three distinct ethnicities as identified by their self-classification on the electoral roll

282 as of Asian, Caucasian or African descent (20 faces for each ethnicity). All face models

283 were recruited through digital media such as social media and scientific research

284 recruitment sites, and models enrolled as students received course credit for their

285 participation in the study. The models were photographed in a frontal pose following the

286 application of makeup by a professional makeup artist. Makeup was applied in a

287 standardised manner for an "everyday" setting, with the first 5 levels (no makeup –

288 pencilling) corresponding to daytime use and the last 2 levels (eyeliner – eyeshadow)

289 corresponding to nighttime use, to closely mirror makeup use in naturalistic settings

290 (see Figure 1 for an example). All face images in the database were feature-aligned and

291 digitally standardised for luminance, visual contrast and visual spatial frequency. One

292 hundred and twenty-six images of 18 individuals were selected from the database

293 following image processing and standardisation. In addition, the images selected for the

294 study showed a linear increase in perceived lightness (as measured through HSV and

295 CIELAB) following each successive stage of application, with an additional increase in

296 Global Contrast Factor (GCF; Matkovic et al., 2005) on the last four levels, consistent

297 with the makeup "looks" employed in Etcoff et al. (2011).

298

299 FIGURE 1 GOES HERE

300

301   The images selected for the current study were tested in an online validation task

302   whereby each face was rated according to emotional valence (1: negative valence – 7:

303   positive valence) and facial attractiveness (1: very unattractive – 7: very attractive). In

304   addition, participants were required to indicate whether they had previously met or

305   knew the person shown in the task. Twelve participants (6 male/female, mean age:

306   23.14, SD: 2.47) completed the validation task on an online research platform (Google

307   Forms), with no participants indicating they were familiar with the identities presented.

308   A repeated-measures ANOVA revealed no significant differences in median facial

309   attractiveness or emotional valence scores between the three ethnicity groups (Asian,

310   Caucasian, African descent), with mean emotional valence scores ranging between 3

311   and 5 on the rating scale.

312   ***Approach / Avoidance Task (AAT)***

313   The approach-avoidance task used in the present study was based on the Approach-

314   Avoidance Task (AAT) used in the study by Wiers et al. (2009). The task version was

315   designed and executed using the Inquisit psychological research software (Millisecond,

316   Inc.).

317   Images of 18 facial identities and 7 cumulative makeup levels were presented in both

318   vertical (portrait) and horizontal (landscape) orientations, totalling 252 images. Initial

319   portrait resolution was 1500 x 2000 pixels while initial landscape resolution was 2000 x

320   1500 pixels, measuring approx. 168° of visual angle. All facial stimuli were unframed

321   and presented against a grey background. Participants were instructed to maintain their

322   attention in the centre of the screen and to move the joystick forwards or backwards

323   according to the image orientation, with the movement assigned to either orientation

324   counterbalanced between participants. The image size increased or decreased according

325   to the extension of the joystick (Thrustmaster® PC USB) in a backwards and forwards

326   direction respectively, up to a maximum increase or decrease of 70% percent of the

327   original image size (see Rinck & Becker, 2007, for a more detailed technical

328   description). That is, pulling the joystick towards the participant resulted in a continuous

329   increase in image size, with a maximum increase of 70% of the original image size,

330   while pushing the joystick away from the participant resulted in a continuous decrease

331   in image size, with a maximum decrease of 70% of the original image size. Each image

332   was presented four times over the course of four blocks, with a total of 1008 trials (144

333   trials per makeup condition), and an equal number of portrait/landscape presentations.

334   Each image stayed on screen until the joystick was fully extended in either direction,

335   and the next trial was initiated. The reaction time on each trial was calculated as the

336   difference between onset of stimulus presentation and the terminus of joystick

337   extension, to ensure a standardised response for all participants, as image

338   contraction/inflation is also an exteroceptive cue of approach/avoidance (Wiers et al.,

339   2009). The order of image presentation was randomised with no replacement.

340

341   *Explicit Rating Task*

342   Immediately following the completion of the AAT, all participants were instructed to

343   rate the images shown in the AAT rated according to emotional valence (1: negative

344   valence – 7: positive valence) and facial attractiveness (1: very unattractive – 7: very

345   attractive). In addition, participants were required to indicate whether they had

346   previously met or knew the person shown in the task, with no participants indicating

347   they were familiar with the identities presented.

348

349 *AAT Pre-Processing*

350 The initial phase of data analysis consisted in excluding trials containing incorrect

351 responses. Only 3.12% of trials contained an incorrect response and means comparisons

352 revealed no significant differences in error rate between the experimental conditions of

353 makeup level and participant gender. Next, AAT difference scores were calculated from

354 the subtraction of the median approach value (pulling the joystick) from the median

355 avoidance value (pushing the joystick) for each image (see Table 1 in the Supplement).

356 Thus, positive values correspond to a faster approach time and slower avoidance time,

357 values close to zero correspond to equal speeds of approach and avoidance, and a

358 negative index corresponds to faster avoidance times and slower approach times. The

359 mean AAT score for each face ethnicity was then computed for every participant. Two

360 experimental factors were examined: makeup level and gender of the participant.

361 To compare the indices between the different levels of makeup, a mixed ANOVA

362 analysis was performed, with participant gender as a between-subjects factor, and

363 makeup level as a within-subjects factor.

364

365 **Results – Experiment 1**

366 **Explicit Rating Task**

367 A repeated-measures ANOVA revealed a significant main effect of makeup on median

368 attractiveness scores ($F_{(6, 23)} = 6.476$, $p < .001$, partial $\eta^2 = .188$), with Bonferroni

369 post-hoc testing revealing significantly lower scores in response to M1 than all other

370 makeup levels ($p < .05$), with no other significant differences between makeup levels (p

371   = n.s.). No significant main effect of gender on median attractiveness (F (1, 28) = 1.679,

372   p = n.s.) or emotional valence (F (1, 28) = 1.082, p = n.s.) was observed, and no

373   significant main effect of makeup level on emotional valence (F (6, 23) = 1.585, p =

374   n.s.) was observed (see Figure 2 for details).

375

376   INSERT FIGURE 2 HERE

377

378   **AAT Difference Scores**

379   A repeated-measures ANOVA was conducted to test the effect of two factors (makeup

380   level and participant gender) on the AAT difference scores computed from the

381   subtraction of the avoidance response by the approach response. The analysis revealed a

382   significant main effect of gender (F (1, 41) = 8.233, p = .005, partial η^2 = .086), with a

383   significantly higher AAT value for the female group (M = 15.82, SD = 4.649) as

384   compared to the male group (M = -3.046, SD = 4.649), indicating a faster approach time

385   and a slower avoidance time in response to all images, and a significant interaction

386   between gender and makeup level (F (6, 41) = 2.299, p = .034, partial η^2 = .025). Simple

387   effects testing Bonferroni-corrected for multiple comparisons revealed a significantly

388   higher female response at level M2 (M = 32.98) compared to M5 (M = -0.667; p = .017)

389   and M7 (M = 8.4; p = .027), and significantly higher male responses at level M1 (M =

390   5.389) compared to M2 (M = -14.62; p = .038) and M4 (M = -18.26; p = .032), at level

391   M6 (M = 8.744) compared to M2 (p = .027) and M4 (p = .009), and at level M7 (M =

392   13.23) compared to M2 (p = .008), M3 (M = -8.033; p = .029), M4 (p = .003) and M5

393   (M = -7.778; p = .030), indicating a female preference for light compared to heavy

394   makeup and a male behavioural preference for no or heavy makeup as opposed to

395   medium makeup. In addition, there was a near-significant main effect of makeup (F (6,

396   41) = 2.108, p = .051). See Figure 3 below for more details.

397

398   INSERT FIGURE 3 HERE

399

400   **Experiment 2 - Methods**

401   **Subjects**

402   In Experiment 2 facial EMG recordings were collected for 40 participants, 20 men and

403   20 women, heterosexual, aged 20 to 26 (M: 22.72, SD: 3.879) and of Caucasian

404   ethnicity. Sample size and composition were calculated based on the effect sizes

405   reported in previous related research studies (Tagai et al., 2016; 2017; N = 38-45, $\eta^2 \approx$

406   0.3), using G*Power 3.1.9.2 (Faul et al., 2007). All participants were informed about the

407   procedure but not informed about the specific objective in the study. Furthermore,

408   deception was employed, by instructing participants the EMG recording device

409   measured skin conductance response (SCR), not muscle activity, to prevent participants

410   from modulating their facial expressions. Participants signed an informed consent form

411   indicating their willingness to participate in the experiment. Participant recruitment took

412   place through digital media such as social media and scientific research recruitment

413   sites. Participants enrolled as students received course credit for their participation in

414   the experiment. The study adhered to the Declaration of Helsinki guidelines and was

415   approved by the institutional ethics committee and national ethics committee.

416   **Procedure**

417   *Stimuli*

418   The experimental stimuli were composed of 108 emotionally-neutral images of 18

419   female faces with 6 different levels of makeup (no makeup, added foundation, added

420   lipstick, added eyebrow pencil and mascara, added eyeliner, and added eyeshadow)

421   from a previously-constructed face database. In this experiment only 6 separate levels

422   were used: 1: no makeup; 2: added foundation; 3: added lipstick; 4: added eyebrow

423   pencil and mascara; 5: added eyeliner; 6: added eyeshadow, to reduce the total number

424   of trials, as no significant differences were reported between the M4 (eyebrow pencil)

425   and M5 (mascara) levels used in Experiment 1, and both products serve as similar

426   perceptual cues (cues signaling higher visual contrast in eye region).

427   ***Facial Electromyography (EMG) Recording***

428   Psychophysiological data was collected through surface EMG recording with four

429   shielded electrode pairs to measure voltage changes linked to muscle activity while

430   participants passively viewed images of female faces with different cumulative levels of

431   makeup, composed of the same images used in the first experiment. The passive

432   viewing task was designed and executed using the E-Prime 2.0 psychological

433   presentation software (Psychology Software Tools, Inc., Pittsburgh, PA). After placing

434   all electrodes on the left side of the face, participants were instructed to maintain their

435   attention in the centre of the screen and passively view images of female faces while

436   recording the EMG signal, with each stimulus level corresponding to a specific marker

437   in the EMG signal. In addition, instructions were given to maintain a relaxed and still

438   posture so as to minimise interference with the recording device, which participants

439   were deceptively informed was for monitoring their skin conductance response. The

440   experiment consisted of 3 blocks, with a total of 324 trials (54 trials per makeup

441   condition). Participants were instructed to take a 5-minute break in between each block,

442   with a total experiment time of approximately 50 minutes. Each image was presented

443   for 2000ms, with an inter-stimulus interval (ISI) of 2000ms, and a 500ms pre-stimulus

444   baseline containing a fixation cross. Prior to the participant debriefing, participants were

445   asked what they perceived to be the objectives of the experiment, with no participant

446   correctly identifying the objective.

447   Data was recorded using the BIOPAC MP150 system with Acqknowledge software

448   (Version 4, Biopac Systems, Inc., Goleta, CA) and a separate EMG110C unit for each

449   of the four facial muscles sires recorded with the current study. The electrodes were

450   positioned over the following muscle sites, according to the guidelines of Fridlund and

451   Cacioppo (1986): corrugator (Corrugador supercilii); zygomatic (Zygomaticus major);

452   levator (Levator labii superiors); and frontalis (Frontalis, pars lateralis). Silver-silver

453   chloride (Ag-AgCl) shielded surface electrode pairs (EL254S) filled with conductive gel

454   (Signa Gel with saline solution) with a contact area of 4mm diameter were used. EMG

455   amplifiers were set to a gain of 2000 and real-time data filtering was conducted through

456   a bandwidth with lower and upper thresholds of 10 Hz and 500 Hz, respectively.

457   Grounding was performed through an additional electrode placed in the middle of the

458   forehead. The sampling rate was held constant at 1000 Hz throughout the experiment.

459   Prior to electrode placement, the surface area of participant's face was wiped with cotton

460   wool and an ethanol solution to remove excess oils and dead skin and thus secure

461   electrode attachment with double-stick adhesive rings. During the task, the experimenter

462   observed the participant through a webcam (recording offline) placed above the

463   monitor, and documented any instances of movement such as coughing or sneezing, for

464   later removal of experimental artefacts in the EMG data.

465   ***Facial EMG Pre-Processing***

466   EMG data preparation was conducted with a custom-made MATLAB script

467   (Mathworks, Natick, MA). First, artifacts were removed according to the documentation

468   during data collection by excluding artefactitious data segments per participant in the

469   respective channel. The EMG data was then filtered with a 28 Hz high-pass filter,

470   rectified and smoothed with a moving average of 50 ms. A total of 2.179 %, 4.362 %,

471   4.341 % and 2.685 % of trials were excluded from the data recorded at the corrugator,

472   zygomatic, levator and frontalis respectively. Each trial was segmented in 100 ms bins

473   resulting in a 500 ms initial baseline period, a 2000 ms period corresponding to stimulus

474   presentation, and a 2000 ms interstimulus interval (ISI). Further data preparation was

475   conducted in Excel (Microsoft Office, Microsoft, Inc.). A spike filter was applied to the

476   EMG data defined by a deviation of +/- 3 SD of the total mean from one bin to the next.

477   All trials that exceeded this definition were winsorized, such that extreme values were

478   set to the next-highest value, as described by Field (2013). To compare the mixed

479   factors of makeup condition (within-subjects) and participant gender (between-

480   subjects), the bins from each trial were z-standardised according to the participant mean,

481   as a secondary dataset. The bins from the two datasets (EMG values and z-scores) were

482   then baseline-corrected, subtracting the mean value from the baseline period. Statistical

483   analysis of the within-subjects factor of makeup condition was conducted using the

484   means across participants for each makeup condition (M1 – M7). An R script (R-

485   Project) was used to conduct functional ANOVA (FANOVA) analyses on the EMG

486   response observed during stimulus presentation, separately for each muscle site

487   (corrugator, zygomatic, levator and frontalis).

488   Functional ANOVA applies the assumptions of analysis of variance to functional

489   observations that, while discrete to specific timepoints in the data, are sampled

490   frequently over a defined period (Ramsay & Silverman, 2005). FANOVA was

491   employed in this case for its utility in analysing the time course of facial EMG response

492   to a complex visual stimulus such as a cosmetically-enhanced face, with a Type-II sum

493 of squares for testing main effects and interactions (Langsrud, 2003). In addition, a

494 functional generalized F-test designed for electrophysiological data analysis was

495 employed whereby exact F statistics and p-values are estimated using Monte Carlo

496 simulation (Causeur et al., 2019b). Data from all time periods (baseline, stimulus

497 presentation, ISI) was included in the analysis but only data from the period of stimulus

498 presentation was included for the purpose of significance testing.

499 Following bandpass filtering and baseline correction, the mean facial EMG response at

500 the four muscle sites: corrugator, zygomatic, levator and frontalis, was calculated for

501 each 100 ms bin including the baseline, stimulus presentation time and ISI. The within-

502 subjects analysis of makeup level was plotted against mean EMG response (μV) and

503 time (ms), while the mixed between- and within-subjects analysis of participant gender

504 and makeup level was plotted against mean z-score and time (ms). Detection of extreme

505 curves was conducted for all participants defined as curves showing large variation with

506 respect to the mean curve under the same conditions of muscle site and gender. A

507 FANOVA using Type-II sum of squares was then conducted to test the effect of makeup

508 condition on EMG response, and makeup condition and gender on participant z-scores

509 (Causeur et al., 2019a; 2019b).

510

511 **Experiment 2 - Results**

512 **Facial EMG Response**

513 The Type-II functional ANOVA of the corrugator EMG response revealed a significant

514 main effect of makeup condition (F = 41.25, p = .003), with Bonferroni-corrected

515 pairwise comparisons revealing significant differences between M1 and M2 (p < .001),

516 M1 and M3 (p < .001), M2 and M3 (p < .001), M3 and M5 (p = .042), and significant

517   differences between M6 and all other makeup levels (M1: $p < .001$; M2: $p < .001$; M3: p

518     $< .001$; M4: $p = .011$; M5: $p < .001$) (see Figure 4 for details). No other significant

519   effects were observed at other muscle sites for EMG response (Zygomatic: F = 28.75, p

520     $= 0.325$; Levator: F = 22.87, $p = 0.796$; Frontalis: F = 20.19, $p = 0.859$).

521

522   FIGURE 4 GOES HERE

523

524   The Type-II functional ANOVA conducted on the z-scores again revealed a significant

525   main effect of makeup condition for the corrugator muscle (F = 45.61, $p < .001$), with

526   Bonferroni-corrected pairwise comparisons revealing a significant difference between

527   M1 and M3 ($p = .048$) and between M2 and M3 ($p = .006$), with no significant main

528   effects of makeup condition observed for the other muscle sites (Zygomatic: F = 23.61,

529     $p = 0.785$; Levator: F = 23.71, $p = 0.768$; Frontalis: F = 19.86, $p = 0.918$) (see Figure 5

530   for details). In addition, significant main effects of participant gender were observed for

531   the corrugator (F = 6.873, $p < .001$) and levator (F = 2.845, $p < .001$), revealing

532   significantly higher z-scores in both instances for female participants compared to males

533   (see Figure 6 for details). No significant main effects of gender were observed for the

534   zygomatic (F = 1.509, $p = 0.197$) or frontalis (F = 0.974, $p = 0.529$) sites. No significant

535   interactions were observed between gender and makeup condition during the time

536   period of stimulus presentation (Corrugator: F = 24.37, $p = 0.586$; Zygomatic: F =

537   40.43, $p = 0.138$; Levator: F = 27.22, $p = 0.488$; Frontalis: F = 25.42, $p = 0.503$).

538

539   FIGURE 5 GOES HERE

540   FIGURE 6 GOES HERE

541

General Discussion

543   Overall, the results from this study indicate a greater behavioural tendency for greater

544   approach and lower avoidance to no and light makeup in female participants, and a

545   specifically male behavioural tendency for greater approach and lower avoidance for

546   heavy over light makeup, and reveal differences in corrugator response (indicating

547   negative affect) towards varying levels of makeup in both men and women, as well as a

548   higher corrugator and levator response in women than men towards all makeup levels.

549   Given the linear increase in reported attractiveness observed over all makeup levels in

550   the explicit rating task (although insignificant from M2 – M7), the results from our

551   EMG analysis are in line with past research showing a negative linear relationship

552   between facial attractiveness and corrugator response (Hazlett & Hoehn-Saric, 2000;

553   Principe & Langlois, 2011). Furthermore, the present study extends past research in

554   both social psychology (Mulhern et al., 2003; Nash et al., 2006), visual perception

555   (Koehler et al., 2004; Russell, 2009), and aesthetic neuroscience (Chatterjee et al., 2009;

556   Tagai et al., 2017), by revealing the interplay between gender, aesthetic preference and

557   visual fluency through the use of implicit responses, as well as the contribution of visual

558   cues linked to female facial attractiveness to implicit emotional response. For the first

559   time, the present study shows changes in electromyographical activity linked to facial

560   affect in response to different levels of makeup in face stimuli.

561   Overall, the AAT task revealed a main effect of gender on behavioural response.

562   However, this effect appears to be in part due to the faster reaction times of the male

563   group as compared to the female group across all images, as confirmed by separate

564 analyses of the median approach and avoidance RTs (see Tables 2 and 3 in

565 Supplement), revealing that male participants were faster in their response to both

566 approach and avoidance of the images. Thus, differences in the behavioural response to

567 the distinct makeup levels should be considered with respect to a separate baseline for

568 each gender. Interestingly, simple effect analyses conducted on the significant

569 interaction between gender and makeup level showed that female participants responded

570 more positively to faces with light makeup (foundation) than heavy makeup applied to

571 the eye regions, with AAT scores for M2 higher than M5 (pencil) and M7 (eyeshadow),

572 consistent with a previously-reported "visual fluency" effect of light makeup in female

573 participants (Tagai et al., 2016; 2017). In contrast, males showed an "all-or-nothing"

574 effect of behavioural preference towards makeup, with simple effects analysis showing

575 significantly higher AAT scores in response to no makeup (M1) and heavy eye makeup

576 (M6 and M7) as opposed to light and intermediate (M2 – M5) levels of makeup. These

577 two distinct patterns of results appear to correspond to separate mechanisms of visual

578 expertise and sexual preference, as described below.

579 With regards to a previously-reported light makeup advantage for visual fluency, the

580 behavioural AAT response indicated a partial preference of female participants for faces

581 with little makeup as compared to medium and high makeup faces, primarily due to a

582 slower avoidance response to these faces (see Table 1 in the Supplement). As a previous

583 study by Tagai et al. (2016) found a recognition bias for light makeup faces in female

584 participants, this effect is in line with past research, and indicates an additional bias in

585 terms of approach-avoidance behaviour for light makeup. However, given the present

586 study does not systematically vary the information content within each makeup level

587 (i.e. visual spatial frequency, skin tone), these results do not offer widespread support

588 for or against a visual fluency account of the cosmetic enhancement of facial

589   attractiveness. Instead, the higher response to foundation (M2) in female participants

590   may simply be due to increased sensitivity to the visual cues provided by foundation, as

591   all female participants reported regularly using facial cosmetics. By contrast, the

592   specifically-male preference for heavy makeup over medium makeup appears to

593   indicate the presence of a secondary effect of sexual propensity towards heavy makeup

594   primarily accentuating greater visual contrast in the eye regions, consistent with the

595   corrugator response recorded in Experiment 2, discussed later.

596   The surprising result of higher AAT response for no makeup compared to light makeup

597   in male participants, appears to have no previous correspondence in the literature, as

598   "no makeup" conditions have been consistently rated as the least attractive faces

599   according to past studies utilising explicit ratings of attractiveness in response to

600   cosmetic enhancement (Mileva et al., 2016; Mulhern et al., 2003; Osborn, 1996), a

601   result confirmed by the explicit rating of the attractiveness of the faces in the present

602   study, revealing that "no makeup" was judged as significantly less attractive than all

603   other makeup levels. However, at an implicit behavioural level, positive male responses

604   to faces with no makeup may be due to the evaluation of such faces as neutral, non-

605   sexualised stimuli, as opposed to cosmetically-enhanced female faces viewed as

606   potential mates and rejected at the initial phase of makeup application. Similarly, female

607   participants may show more positive approach-avoidance behaviour towards faces with

608   no makeup due to being viewed as neutral non-threatening competitors (Stockley &

609   Campbell, 2013).

610   With regards to the facial EMG response in Experiment 2, activity at the corrugator

611   muscle site over the course of stimulus presentation was significantly higher in response

612   to no makeup (M1) than light makeup (M2 and M3) and eyeshadow (M6), revealing a

613   more relaxed corrugator pose in response to addition of these levels of cosmetic

614   application, indicative of decreased negative affect (Principe & Langlois, 2011). In

615   addition, lipstick (M3) was found to play a key role in the attenuation of corrugator

616   response, with the lowest corrugator response recorded and significantly lower than all

617   levels with the exception of mascara and pencilling (M4). Finally, eyeshadow (M6)

618   displayed the next lowest corrugator response, with significantly lower values to all

619   other makeup levels with the exception of lipstick. Overall, these results indicate a

620   significant effect of makeup application on the attenuation of corrugator response.

621   While corrugator response has been associated with increased cognitive load (Lishner et

622   al., 2008) which may have contributed to the smoothing effect of foundation (M2) on

623   corrugator activity, these results primarily indicate lower negative affect in response to

624   an increase in the stages of makeup application. Interestingly, this effect was observed

625   in both genders, indicating a similar affective response to the visual cues of makeup.

626   Notably however, the addition of eye makeup such as mascara, pencilling and eyeliner

627   (associated with higher visual contrast) did not produce a significant decrease in

628   corrugator response as compared to the no makeup condition. Thus, we found no

629   evidence for the effect of these products on facial attractiveness as gauged by corrugator

630   response. Instead, significant attenuation of corrugator response was observed only in

631   response to the addition of eyeshadow. Visual contrast accentuating the eye and lip

632   regions has been proposed as one of the major determinants of facial attractiveness and

633   femininity in female faces (Russell, 2009; 2010). Our results suggest that the visual

634   contrast in the eye regions must be sufficiently intense to produce a change in affective

635   response contributing to perceived attractiveness. This is consistent with the observation

636   of a linear increase in global contrast factor (Matkovic et al., 2005) on levels M4 – M7

637   of the images employed in the AAT task as well as a gradual, but non-significant,

638   increase in explicitly-rated attractiveness of these makeup levels. The significant

639   decrease in corrugator activity observed for the addition of lipstick highlights the

640   importance of this region in providing a visual cue likely associated with luminance

641   contrast (Russell, 2003; 2009), to determine an appropriate affective response for

642   guiding the evaluation of facial attractiveness. The marked reduction in corrugator

643   response towards the presence of lipstick and eyeshadow does not support the role of

644   visual fluency in reducing cognitive load as the sole determinant of perceived

645   attractiveness, at least on an affective level, instead indicating that the evaluation of

646   facial attractiveness as enhanced by makeup relies on a wide set of visual cues eliciting

647   distinct behavioural and affective reactions.

648   An analysis of the participants' z-scores over the course of stimulus presentation

649   revealed higher corrugator and levator activity for female subjects than male subjects.

650   Interestingly, this gender difference occurred at an early peak of EMG response, likely

651   corresponding to an orienting response towards novel facial stimuli (Achaibou et al.,

652   2008; Dimberg, 1982). While corrugator activity has specifically been correlated with

653   early visual processing (approximately 200 ms after stimulus onset; Achaibou et al.,

654   2008), we cannot discount the possibility that this difference may be due to enhanced

655   attention towards facial stimuli containing makeup in female participants. Alternatively,

656   this result may reflect an initially adverse negative emotional response in women

657   towards female faces prior to subsequent modulation, indicating increased female

658   intrasexual competition (Stockley & Campbell, 2013).

659   Similarly, the increase in corrugator and levator response in women than men in

660   response to all images is unlikely to be due to increased cognitive load due to the greater

661   familiarisation the women had with the makeup products applied (all female

662   participants reported regularly using makeup at least once per week). While a recent

663   large-scale study examining facial muscles according to the Facial Action Coding

664   System (FACS; Ekman et al, 2002) has shown greater expressiveness in female facial

665   actions associated with positive valence, a corresponding difference in negative facial

666   affect between male and female faces was not reported (McDuff et al., 2017). Given

667   these factors and the well-established link between corrugator response and negative

668   affect (Larsen et al., 2003; Neta et al., 2009), the corrugator response observed in the

669   present study can reasonably be attributed to an affective index of aesthetic preference

670   for facial cosmetics, revealing a more negative affective response to female faces with

671   and without makeup in women than men.

672   An important caveat must be made with respect to the limitations of utilising facial

673   EMG in the measurement of differences in aesthetic judgement. We were unable to

674   compile a complete affective 'profile' of the valence and intensity of participants'

675   emotional response due to the variability and lack of significant voltage changes to the

676   different stimulus types used in Experiment 2, at all muscle sites with the exception of

677   the corrugator supercilii and levator (Figure 3.). While EMG measurement was more

678   sensitive than AAT response to intermediate differences in makeup application, for

679   example in the distinct perceptual cues associated with lipstick and eye makeup, as a

680   whole facial EMG may be insufficiently sensitive towards the effect of relatively subtle

681   physiological cues on facial attractiveness, and future studies investigating the affective

682   responses underlying aesthetic experience should consider pairing the technique with an

683   explicit attractiveness rating task, for example. The inclusion of eye-tracking measures

684   to monitor which precise face regions the participant attends to while rating

685   attractiveness, may also provide a useful 'attentional' index of aesthetic preference.

686   Previous studies have tested the role of eye gaze in evaluating female facial

687   attractiveness, indicating both that attractive faces receive longer gaze durations and a

688   greater number of directed saccades than unattractive faces (Leder et al., 2016) and that

689 participants attend longer to the nose than other facial regions during the evaluation of

690 facial attractiveness (Zhang et al., 2017). Future research may be directed at the role of

691 cosmetic enhancement in guiding attention during attractiveness judgments.

692 A further caveat is the limited support this study found for the smoothing, texturing and

693 colour distribution effects of foundation on the evaluation of facial attractiveness,

694 commonly associated with signals of youth and individual health (Fink & Matts, 2008;

695 Jones et al., 2015; Porcheron et al., 2013). While there was a slight (but non-significant)

696 increase in mean AAT score from M1 (no makeup) to M2 (foundation) for female

697 participants, this effect was inverted in male subjects, showing significantly greater

698 behavioural preference for no makeup than foundation. However, EMG recording of the

699 corrugator site revealed a significantly lower response to faces with foundation than

700 faces with no makeup in both genders, suggesting that this cue of facial attractiveness is

701 more dependent on one's affective response than the enhancement of visual contrast in

702 the eye regions for example, which was reflected in both EMG and AAT response.

703 Overall, the present study found two clear indices of the implicit evaluation of facial

704 attractiveness as modulated by changes in facial cosmetics; a behavioural index,

705 characterised in female participants by a preference for faces with light makeup, and in

706 males by an all-or-nothing preference for faces with no makeup or heavy eye cosmetics.

707 The second index corresponds to the individual's negative affective response, reflected

708 primarily in terms of reduced electromyographical response at the corrugator muscle

709 site to facial cosmetics accentuating visual contrast in the mouth and eye regions. Given

710 the counterintuitive results reported, particularly with regards to a male implicit

711 behavioural preference for no makeup over light makeup, the evaluation of female

712 facial attractiveness appears to rely on a complex set of perceptual and behavioural

713 cues, highlighting the importance of implicit measures in further investigations.

714

715

References

716

Achaibou, A., Pourtois, G., Schwartz, S., & Vuilleumier, P. (2008). Simultaneous
recording of EEG and facial muscle reactions during spontaneous emotional
mimicry. *Neuropsychologia*, *46*(4), 1104–1113.
https://doi.org/10.1016/j.neuropsychologia.2007.10.019

Baudouin, J. Y., & Tiberghien, G. (2004). Symmetry, averageness, and feature size in
the facial attractiveness of women. *Acta Psychologica*, *117*(3), 313-332.
https://doi.org/10.1016/j.actpsy.2004.07.002

Cacioppo, J. T., Petty, R. E., Losch, M. E., & Kim, H. S. (1986). Electromyographic
activity over facial muscle regions can differentiate the valence and intensity of
affective reactions. *Journal of Personality and Social Psychology*, *50*, 260–268.
https://doi.org/10.1037/0022-3514.50.2.260

Causeur, D., Sheu, C-F., Chu, M-C. and Rufini, F. (2019a). *ERP: Significance Analysis
of Event-Related Potentials Data*. R package version 2.2. Retrieved 14/02/2020
from: https://CRAN.R-project.org/package=ERP.

Causeur, D, Sheu, C.-F., Perthame, E, Rufini, F. (2019b). A functional generalized F-
test for signal detection with applications to event-related potentials significance
analysis. *Biometrics*, 1– 11. https://doi.org/10.1111/biom.13118

Chatterjee, A., Thomas, A., Smith, S. E., & Aguirre, G. K. (2009). The neural response
to facial attractiveness. *Neuropsychology*, *23*(2), 135-143.
https://doi.org/10.1037/a0014430

737 Chatterjee, A., & Vartanian, O. (2016). Neuroscience of aesthetics. *Annals of the New*

738   *York Academy of Sciences*, *1369*(1), 172-194. https://doi.org/10.1111/nyas.13035

739 Cunningham, M. R., Roberts, A. R., Barbee, A. P., Druen, P. B., & Wu, C. H. (1995).

740   "Their ideas of beauty are, on the whole, the same as ours": Consistency and

741   variability in the cross-cultural perception of female physical

742   attractiveness. *Journal of personality and social psychology*, *68*(2), 261.

743   https://doi.org/10.1037/0022-3514.68.2.261

744 Dimberg, U. (1982). Facial Reactions to Facial Expressions. *Psychophysiology*, *19*(6),

745   643–647. https://doi.org/cmpnb7

746 Dimberg, U., & Thunberg, M. (1998). Rapid facial reactions to emotional facial

747   expressions. *Scandinavian Journal of Psychology*, *39*(1), 39–45.

748   https://doi.org/10.1111/1467-9450.00054

749 Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to

750   emotional facial expressions. *Psychological Science*, *11*(1), 86-89.

751   https://doi.org/10.1111/1467-9280.00221

752 Ekman, P., Friesen, W. V., & Hager, J. C. (2002). Facial action coding system: The

753   manual on CD ROM. *A Human Face, Salt Lake City*, 77-254.

754 Etcoff, N. L., Stock, S., Haley, L. E., Vickery, S. A., & House, D. M. (2011). Cosmetics

755   as a feature of the extended human phenotype: Modulation of the perception of

756   biologically important facial signals. *PloS one*, *6*(10), e25656.

757   https://doi.org/10.1371/journal.pone.0025656

758 Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible

759   statistical power analysis program for the social, behavioral, and biomedical

760   sciences. *Behavior Research Methods, 39*, 175-191.

761   https://doi.org/10.3758/BF03193146

762   Fridlund, A. J., & Cacioppo, J. T. (1986). Guidelines for human electromyographic

763   research. *Psychophysiology*, *23*(5), 567-589. https://doi.org/btrbt5

764   Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. SAGE Publications.

765   Fink, B., & Matts, P. J. (2008). The effects of skin colour distribution and topography

766   cues on the perception of female facial age and health. *Journal of the European*

767   *Academy of Dermatology and Venereology*, *22*(4), 493-498.

768   https://doi.org/10.1111/j.1468-3083.2007.02512.x

769   Fink, B., & Neave, N. (2005). The biology of facial beauty. *International Journal of*

770   *Cosmetic Science*, *27*(6), 317-325. https://doi.org/c978s9

771   Fink, B., & Penton-Voak, I. (2002). Evolutionary psychology of facial

772   attractiveness. *Current Directions in Psychological Science*, *11*(5), 154-158.

773   https://doi.org/10.1111/1467-8721.00190

774   Gerger, G., & Leder, H. (2015). Titles change the esthetic appreciations of

775   paintings. *Frontiers in Human Neuroscience*, *9*, 464.

776   https://doi.org/10.3389/fnhum.2015.00464

777   Gerger, G., Leder, H., Tinio, P. P., & Schacht, A. (2011). Faces versus patterns:

778   Exploring aesthetic reactions using facial EMG. *Psychology of Aesthetics,*

779   *Creativity, and the Arts*, *5*(3), 241. https://doi.org/10.1037/a0024154

780   Langsrud, Ø. (2003). ANOVA for unbalanced data: Use Type II instead of Type III

781   sums of squares. *Statistics and Computing*, *13*(2), 163-167.

782   Hazlett, R. L., & Hoehn-Saric, R. (2000). Effects of perceived physical attractiveness on

783   females' facial displays and affect. *Evolution and Human Behavior*, *21*(1), 49-57.

784   https://doi.org/10.1016/S1090-5138(99)00036-7

785   Heuer, K., Rinck, M., & Becker, E. S. (2007). Avoidance of emotional facial

786   expressions in social anxiety: The approach–avoidance task. *Behaviour Research*

787   *and Therapy*, *45*(12), 2990-3001. https://doi.org/10.1016/j.brat.2007.08.010

788   Hess, U., & Fischer, A. (2014). Emotional mimicry: why and when we mimic emotions.

789   *Social and Personality Psychology Compass*, *8*(2), 45–57.

790   https://doi.org/10.1111/spc3.12083

791   Jones, A. L., & Kramer, R. S. (2016). Facial Cosmetics and Attractiveness: Comparing

792   the Effect Sizes of Professionally-Applied Cosmetics and Identity. *PloS*

793   *One*, *11*(10), e0164218. https://doi.org/10.1371/journal.pone.0164218

794   Jones, B. C., Hahn, A. C., Fisher, C. I., Wincenciak, J., Kandrik, M., Roberts, S. C., ...

795   & DeBruine, L. M. (2015). Facial coloration tracks changes in women's

796   estradiol. *Psychoneuroendocrinology*, *56*, 29-34.

797   https://doi.org/10.1016/j.psyneuen.2015.02.021

798   Koehler, N., Simmons, L. W., Rhodes, G., & Peters, M. (2004). The relationship

799   between sexual dimorphism in human faces and fluctuating asymmetry.

800   *Proceedings of the Royal Society of London B: Biological Sciences*, *271*(Suppl 4),

801   S233-S236. https://doi.org/10.1098/rsbl.2003.0146

802   Kramer, R. S., Mulgrew, J., Anderson, N. C., Vasilyev, D., Kingstone, A., Reynolds, M.

803   G., & Ward, R. (2020). Physically attractive faces attract us physically. *Cognition*,

804   *198*, 104193. https://doi.org/10.1016/j.cognition.2020.104193

805 Krieglmeyer, R., & Deutsch, R. (2010). Comparing measures of approach–avoidance

806   behaviour: The manikin task vs. two versions of the joystick task. *Cognition and*

807   *Emotion*, *24*(5), 810-828. https://doi.org/10.1080/02699930903047298

808 Kukde, M. P., & Neufeld, R. W. (1994). Facial electromyographic measures distinguish

809   covert coping from stress response to stimulus threat. *Personality and Individual*

810   *Differences*, *16*(2), 211-228. https://doi.org/10.1016/0191-8869(94)90160-0

811 Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M.

812   (2000). Maxims or myths of beauty? A meta-analytic and theoretical

813   review. *Psychological Bulletin*, *126*(3), 390. https://doi.org/cfb9cd

814 Larsen, J. T., Norris, C. J., & Cacioppo, J. T. (2003). Effects of positive and negative

815   affect on electromyographic activity over zygomaticus major and corrugator

816   supercilii. *Psychophysiology*, *40*(5), 776-785. https://doi.org/dkc6vh

817 Law-Smith, M., Perrett, D. I., Jones, B. C., Cornwell, R. E., Moore, F. R., Feinberg, D.

818   R., Boothroyd, L. G., Durrani, S. J., Stirrat, M. R., Whiten, S., Pitman, R. M., &

819   Hillier, S. G. (2006). Facial appearance is a cue to oestrogen levels in women.

820   *Proceedings of the Royal Society of London B: Biological Sciences*, *273*(1583),

821   135-140. https://doi.org/10.1098/rspb.2005.3296

822 Leder, H., Mitrovic, A., & Goller, J. (2016). How beauty determines gaze! Facial

823   attractiveness and gaze duration in images of real world scenes. *i-Perception*, *7*(4),

824   2041669516664355. https://doi.org/10.1177/2041669516664355

825 Lishner, D. A., Cooter, A. B., & Zald, D. H. (2008). Rapid emotional contagion and

826   expressive congruence under strong test conditions. *Journal of Nonverbal*

827   *Behavior*, *32*(4), 225-239. https://doi.org/10.1007/s10919-008-0053-y

828 Little, A. C., Jones, B. C., & DeBruine, L. M. (2011). Facial attractiveness: evolutionary

829   based research. *Philosophical Transactions of the Royal Society of London B:*

830   *Biological Sciences*, *366*(1571), 1638-1659. https://doi.org/10.1098/rstb.2010.0404

831 Lundqvist, L.-O. (1995). Facial EMG reactions to facial expressions: A case of facial

832   emotional contagion? *Scandinavian Journal of Psychology*, *36*(2), 130–141.

833   https://doi.org/bs85wt

834 Lundqvist, L.-O., & Dimberg, U. (1995). Facial expressions are contagious. *Journal of*

835   *Psychophysiology*, *9*(3), 203–211.

836 Matkovic, K., Neumann, L., Neumann, A., Psik, T., & Purgathofer, W. (2005). Global

837   contrast factor-a new approach to image contrast. *Computational Aesthetics in*

838   *Graphics, Visualization and Imaging*, *2005*(159-168), 1. https://doi.org/fhxp

839 McDuff, D., Kodra, E., el Kaliouby, R., & LaFrance, M. (2017). A large-scale analysis

840   of sex differences in facial expressions. *PloS one*, *12*(4), e0173942.

841   https://doi.org/10.1371/journal.pone.0173942

842 Mileva, V. R., Jones, A. L., Russell, R., & Little, A. C. (2016). Sex differences in the

843   perceived dominance and prestige of women with and without

844   cosmetics. *Perception*, *45*(10), 1166-1183. https://doi.org/f86f6v

845 Morrison, E. R., Clark, A. P., Tiddeman, B. P., & Penton-Voak, I. S. (2010).

846   Manipulating shape cues in dynamic human faces: Sexual dimorphism is preferred

847   in female but not male faces. *Ethology*, *116*(12), 1234-1243. https://doi.org/bgg6dd

848 Mulhern, R., Fieldman, G., Hussey, T., Lévêque, J. L., & Pineau, P. (2003). Do

849   cosmetics enhance female Caucasian facial attractiveness?. *International Journal*

850   *of Cosmetic Science*, *25*(4), 199-205. https://doi.org/frmj9p

851   Nakamura, K., Kawashima, R., Nagumo, S., Ito, K., Sugiura, M., Kato, T., ... & Kojima,

852   S. (1998). Neuroanatomical correlates of the assessment of facial

853   attractiveness. *Neuroreport*, *9*(4), 753-757. https://doi.org/c7hrsm

854   Nash, R., Fieldman, G., Hussey, T., Lévêque, J. L., & Pineau, P. (2006). Cosmetics:

855   They influence more than Caucasian female facial attractiveness. *Journal of*

856   *Applied Social Psychology*, *36*(2), 493-504. https://doi.org/c34j72

857   Neta, M., Norris, C. J., & Whalen, P. J. (2009). Corrugator muscle responses are

858   associated with individual differences in positivity-negativity bias. *Emotion*, *9*(5),

859   640-648. https://doi.org/10.1037/a0016819

860   Osborn, D. R. (1996). Beauty Is as Beauty Does?: Makeup and Posture Effects on

861   Physical Attractiveness Judgments. *Journal of Applied Social Psychology*, *26*(1),

862   31-51. https://doi.org/10.1111/j.1559-1816.1996.tb01837.x

863   Penton-Voak, I. S., Jacobson, A., & Trivers, R. (2004). Populational differences in

864   attractiveness judgements of male and female faces: Comparing British and

865   Jamaican samples. *Evolution and Human Behavior*, *25*(6), 355-370.

866   https://doi.org/10.1016/j.evolhumbehav.2004.06.002

867   Porcheron, A., Mauger, E., & Russell, R. (2013). Aspects of facial contrast decrease

868   with age and are cues for age perception. *PloS one*, *8*(3), e57985.

869   https://doi.org/10.1371/journal.pone.0057985

870   Principe, C. P., & Langlois, J. H. (2011). Faces differing in attractiveness elicit

871   corresponding affective responses. *Cognition and Emotion*, *25*(1), 140-148.

872   https://doi.org/10.1080/02699931003612098

873 Principe, C. P., & Langlois, J. H. (2012). Shifting the prototype: Experience with faces

874   influences affective and attractiveness preferences. *Social Cognition*, *30*(1), 109-

875   120. https://doi.org/10.1521/soco.2012.30.1.109

876 Ramsay, J. O., & Silverman, B. W. (2005). *Functional Data Analysis (2nd Edition)*.

877   Springer Series in Statistics. Springer. https://doi.org/10.1007/b98888

878 Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of*

879   *Psychology*, *57*, 199-226. https://doi.org/ddg2hh

880 Richetin, J., Croizet, J. C., & Huguet, P. (2004). Facial make-up elicits positive attitudes

881   at the implicit level: Evidence from the implicit association test. *Current Research*

882   *in Social Psychology*, *9*(11), 145-164.

883 Rinck, M., & Becker, E. S. (2007). Approach and avoidance in fear of spiders. *Journal*

884   *of Behavior Therapy and Experimental Psychiatry*, *38*(2), 105-120.

885   https://doi.org/10.1016/j.jbtep.2006.10.001

886 Rinn, W. E. (1984). The neuropsychology of facial expression: a review of the

887   neurological and psychological mechanisms for producing facial expressions.

888   *Psychological Bulletin*, *95*(1), 52–77. https://doi.org/10.1037/0033-2909.95.1.52

889 Russell, R. (2003). Sex, beauty, and the relative luminance of facial features.

890   *Perception*, *32*(9), 1093-1107. https://doi.org/10.1068/p5101

891 Russell, R. (2009). A sex difference in facial contrast and its exaggeration by cosmetics.

892   *Perception*, *38*(8), 1211-1219. https://doi.org/10.1068/p6331

893 Russell, R. (2010). Why cosmetics work. In R.B. Adams, N. Ambady, K. Nakayama,

894   and S. Shimojo, (Eds.) *The Science of Social Vision* (pp. 186-204). Oxford Series

895   in Social Cognition vol. 7. Oxford University Press.

896   https://doi.org/10.1093/acprof:oso/9780195333176.003.0011

897   Sofer, C., Dotsch, R., Wigboldus, D. H., & Todorov, A. (2015). What is typical is good:

898   The influence of face typicality on perceived trustworthiness. *Psychological*

899   *Science*, *26*(1), 39-47. https://doi.org/10.1177/0956797614554955

900   Stockley, P., & Campbell, A. (2013). Female competition and aggression:

901   interdisciplinary perspectives. https://doi.org/10.1098/rstb.2013.0073

902   Tagai, K., Ohtaka, H., & Nittono, H. (2016). Faces with light makeup are better

903   recognized than faces with heavy makeup. *Frontiers in Psychology*, *7*, 226.

904   https://doi.org/10.3389/fpsyg.2016.00226

905   Tagai, K., Shimakura, H., Isobe, H., & Nittono, H. (2017). The light-makeup advantage

906   in facial processing: Evidence from event-related potentials. *PloS One*, *12*(2),

907   e0172489. https://doi.org/10.1371/journal.pone.0172489

908   Thornhill, R., & Grammer, K. (1999). The body and face of woman: One ornament that

909   signals quality? *Evolution and Human Behavior*, *20*(2), 105-120.

910   https://doi.org/10.1016/S1090-5138(98)00044-0

911   Todorov, A. (2008). Evaluating faces on trustworthiness: an extension of systems for

912   recognition of emotions signaling approach/avoidance behaviors. *Annals of the*

913   *New York Academy of Sciences*, *1124*(1), 208-224. https://doi.org/fk35vn

914   Todorov, A., Baron, S. G., & Oosterhof, N. N. (2008). Evaluating face trustworthiness:

915   a model based approach. *Social Cognitive and Affective Neuroscience*, *3*(2), 119-

916   127. https://doi.org/10.1093/scan/nsn009

917   Wiers, R. W., Rinck, M., Dictus, M., & Van den Wildenberg, E. (2009). Relatively

918   strong automatic appetitive action-tendencies in male carriers of the OPRM1 G-

919   allele. *Genes, Brain and Behavior*, *8*(1), 101-106. https://doi.org/cbvrc7

920   Winkielman, P., Halberstadt, J., Fazendeiro, T., & Catty, S. (2006). Prototypes are

921   attractive because they are easy on the mind. *Psychological Science*, *17*(9), 799-

922   806. https://doi.org/dt4c3m

923   Winston, J. S., O'Doherty, J., Kilner, J. M., Perrett, D. I., & Dolan, R. J. (2007). Brain

924   systems for assessing facial attractiveness. *Neuropsychologia*, *45*(1), 195-206.

925   https://doi.org/d38755

926   Zebrowitz, L. A., & Montepare, J. M. (2008). Social psychological face perception:

927   Why appearance matters. *Social and Personality Psychology Compass*, *2*(3), 1497-

928   1517. https://doi.org/ff3s4w

929   Zhang, Y., Wang, X., Wang, J., Zhang, L., & Xiang, Y. (2017). Patterns of eye

930   movements when observers judge female facial attractiveness. *Frontiers in*

931   *Psychology*, *8*, 1909. https://doi.org/10.3389/fpsyg.2017.01909

932

933   **Figure Legends**

934   Figure 1. Examples of the different makeup levels used in Experiment 1 (a) and

935   Experiment 2 (b). For illustrative purposes, this figure shows the same individual at

936   different stages of makeup application. A total of 18 individuals were shown in all

937   makeup levels. The individual shown gave explicit written consent for the publication

938   of her face images.

939   Figure 2. Graphs of median response scores of a) attractiveness ("How attractive is this

940   face from 1 to 7?") and b) emotion ("How emotional is this face from 1 to 7?), as

941   measured on the explicit rating task in Experiment 1. Scores from 1 to 7 reflect faces

942   judged as a) 1: not at all attractive to 7: very attractive; and b) 1: very emotionally

943   negative, to 4: emotionally neutral, to 7: very emotionally positive. Error bars show $\pm$ 1

944   standard error of the mean.

945   Figure 3. Graph of AAT difference scores for each makeup level (no makeup,

946   foundation, lipstick, mascara, pencil, eyeliner, eyeshadow), with separate lines for

947   gender. AAT scores refer to median avoidance RT – median approach RT, with higher

948   scores reflecting faster approach and slower avoidance of the image. Error bars show $\pm$

949   1 standard error of the mean.

950   Figure 4. Graphs of EMG values (μv) recorded at the corrugator, zygomatic, levator,

951   and frontalis sites in response to viewing of 6 makeup levels (M1: no makeup; M2:

952   foundation; M3: lipstick; M4: mascara + pencil; M5: eyeliner; M6: eyeshadow). Curves

953   show average EMG voltage change across participants for each 100 ms bin over the

954   periods of pre-stimulus baseline (-500 to 0 ms), stimulus presentation (0 to 2000 ms)

955   and interstimulus interval (ISI; 2000 to 4000 ms).

956   Figure 5. Graphs of Z-transformed EMG values at the corrugator, zygomatic, frontalis

957   and levator sites in response to viewing of 6 makeup levels (M1: no makeup; M2:

958   foundation; M3: lipstick; M4: mascara + pencil; M5: eyeliner; M6: eyeshadow). Solid

959   curves show average Z-scores of all participants and shaded areas show confidence

960   intervals for each 100 ms bin over the periods of pre-stimulus baseline (-500 to 0 ms),

961   stimulus presentation (0 to 2000 ms) and interstimulus interval (ISI; 2000 to 4000 ms).

962   Figure 6. Graphs of Z-transformed EMG values at the corrugator and levator sites in

963   male and female participants. Solid curves show average Z-scores of male and female

964   participants and shaded areas show confidence intervals for each 100 ms bin over the

965   periods of pre-stimulus baseline (-500 to 0 ms), stimulus presentation (0 to 2000 ms)

966   and interstimulus interval (ISI; 2000 to 4000 ms).