



Swansea University  
Prifysgol Abertawe



## Cronfa - Swansea University Open Access Repository

---

This is an author produced version of a paper published in :

*PLOS ONE*

Cronfa URL for this paper:

<http://cronfa.swan.ac.uk/Record/cronfa30525>

---

### Paper:

Jones, A. & Kramer, R. (2016). Facial cosmetics and attractiveness: Comparing the effect sizes of professionally-applied cosmetics and identity. *PLOS ONE*, 11(10), e0164218

<http://dx.doi.org/10.1371/journal.pone.0164218>

---

This article is brought to you by Swansea University. Any person downloading material is agreeing to abide by the terms of the repository licence. Authors are personally responsible for adhering to publisher restrictions or conditions. When uploading content they are required to comply with their publisher agreement and the SHERPA RoMEO database to judge whether or not it is copyright safe to add this version of the paper to this repository.

<http://www.swansea.ac.uk/iss/researchsupport/cronfa-support/>

1 **Facial Cosmetics and Attractiveness: Comparing the Effect Sizes of Professionally-Applied**  
2 **Cosmetics and Identity**

3

4 Alex L. Jones<sup>1\*</sup>, Robin S. S. Kramer<sup>2</sup>

5

6 <sup>1</sup> Department of Psychology, Swansea University, Swansea, UK

7

8 <sup>2</sup> Department of Psychology, University of York, York, UK

9

10

11 \* Corresponding author

12

13 E-mail: alexjonesphd@gmail.com (ALJ)

14

15

16 Short title: Facial Cosmetics and Attractiveness

## 17 **Abstract**

18           Forms of body decoration exist in all human cultures. However, in Western societies,  
19 females are more likely to engage in appearance modification, especially through the use of  
20 facial cosmetics. How effective are cosmetics at altering attractiveness? Previous research has  
21 hinted that the effect is not large, especially when compared to the variation in attractiveness  
22 observed between individuals due to differences in identity. In order to build a fuller  
23 understanding of how cosmetics and identity affect attractiveness, here we examine how  
24 professionally-applied cosmetics alter attractiveness and compare this effect with the variation in  
25 attractiveness observed between individuals. In Study 1, 33 YouTube models were rated for  
26 attractiveness before and after the application of professionally-applied cosmetics. Cosmetics  
27 explained a larger proportion of the variation in attractiveness compared with previous studies,  
28 but this effect remained smaller than variation caused by differences in attractiveness between  
29 individuals. Study 2 replicated the results of the first study with a sample of 45 supermodels,  
30 with the aim of examining the effect of cosmetics in a sample of faces with low variation in  
31 attractiveness between individuals. While the effect size of cosmetics was generally large,  
32 between-person variability due to identity remained larger. Both studies also found interactions  
33 between cosmetics and identity – more attractive models received smaller increases when  
34 cosmetics were worn. Overall, we show that professionally-applied cosmetics produce a larger  
35 effect than self-applied cosmetics, an important theoretical consideration for the field. However,  
36 the effect of individual differences in facial appearance is ultimately more important in  
37 perceptions of attractiveness.

38 **Keywords:** cosmetics; attractiveness; effect size; social perceptions

## 39 **Introduction**

40           Modification of the body with dyes, paints, and other pigments is among the most  
41 universal of human behaviours, present in all cultures [1–3]. However, in Western society,  
42 women perform the majority of self-adornment [4], and perhaps the most prevalent behavior of  
43 this kind is the use of facial cosmetics. This behaviour is served by the global cosmetics industry  
44 which is worth billions of pounds [5].

45           Women report using cosmetics for a variety of reasons, ranging from anxiety about facial  
46 appearance, conformity to social norms, and public self-consciousness [6–8], through to  
47 appearing more sociable and assertive to others [6]. Cosmetics are effective at improving social  
48 perceptions that the wearer may wish to modulate, with individuals appearing to be healthier and  
49 earning more [9], displaying greater competence, likeability and trustworthiness [10], as well as  
50 appearing more prestigious or dominant [11]. Cosmetics also influence the behaviour of others,  
51 especially males, who tip higher amounts and with greater frequency to waitresses wearing  
52 cosmetics [12], and are more likely to approach wearers in the environment [13]. It is likely that  
53 the effect of cosmetics on social perceptions is brought about by the increase in attractiveness it  
54 confers to faces, which is now a well documented effect [10,14–17]. Research has documented  
55 cosmetics function by altering sex-typical colouration in faces such as facial contrast [18–21], by  
56 increasing the homogeneity of facial skin [22,23], or by affecting colour cues to traits such as  
57 health [24] and age [25].

58           While the effect of cosmetics on perceived attractiveness seems clear [14,17], other  
59 research has revealed it is more nuanced than previously thought. Etcoff and colleagues [10]  
60 demonstrated that attractiveness increased linearly with the amount of cosmetics worn – simply,  
61 more cosmetics equates to appearing more attractive. Of the range of cosmetics that can be worn,

62 the quantity of cosmetics applied to the eyes and mouth have been shown to be significant  
63 predictors of attractiveness [26], with more cosmetics on these features leading to higher ratings  
64 of attractiveness. However, other evidence suggests that the typical amount of cosmetics applied  
65 by a sample of young women is excessive, with observers preferring close to half the actual  
66 amount for optimal attractiveness [16], calling into question the linear relationship between  
67 cosmetics quantity and attractiveness.

68 One concern of facial attractiveness research is that it does not compare the effects of  
69 predictors of attractiveness (e.g., symmetry, averageness, sex typicality [27]; against other  
70 sources of variation [28]. Recent work has begun to address this by examining the importance of  
71 within-person variation in attractiveness (caused by the presence or absence of makeup, for  
72 example), compared with the between-person variation in attractiveness simply due to  
73 differences between identities [29]. Specifically, it has been previously shown that the effect of  
74 cosmetics on attractiveness, a source of within-person variation, is very small, explaining just 2%  
75 of the variance in ratings [15]. This is an especially small effect when compared with differences  
76 in attractiveness between individuals, a between-person variation in attractiveness, which  
77 explained 69% of the variance in judgements. More simply, while facial cosmetics do increase  
78 attractiveness, that contribution is small and does little to change an individual's attractiveness  
79 standing in the population.

80 However, the use of cosmetics is an idiosyncratic and extremely varied practice [3], and  
81 its effect on attractiveness is more complex than previously thought. The use of a professional  
82 makeup artist is a common practice in almost all studies examining the effect of cosmetics on  
83 perceptions [9,10,12,17,30,31], and only a few utilise self-applied cosmetics [14,16,26]. An  
84 initial examination of the effect size of cosmetics on attractiveness also had models self-apply

85 their cosmetics [15]. There are good reasons for using professionally-applied cosmetics, as it  
86 provides a clearer test of how cosmetics alter facial attractiveness. The increased variability in  
87 self-applied cosmetics, due, for example, to differences in application skill or the products used,  
88 could make it more difficult to detect an effect of cosmetics on attractiveness, and previous work  
89 has indeed found the effect to be small [15]. This distinction represents a trade-off between  
90 experimental control and ecological validity – the vast majority of women, if any, do not have a  
91 professional makeup artist apply their cosmetics daily, yet the majority of studies examining  
92 cosmetics and attractiveness draw conclusions based on professionally-applied cosmetics, which  
93 may only indirectly inform as to how cosmetics affect attractiveness in the real world.

94 We seek to address important theoretical points regarding how cosmetics influence  
95 attractiveness. How large is the effect size of cosmetics on attractiveness when cosmetics have  
96 been professionally-applied? If cosmetics in psychological experiments are applied with more  
97 skill than is typically achieved, then current knowledge of cosmetics and attractiveness likely  
98 overstates the relationship, given the reliance of professionally-applied cosmetics in the  
99 literature. Moreover, how does the ability of professionally-applied cosmetics compare to  
100 previous measures of the effect of cosmetics on attractiveness? In the following study, we  
101 examine the effect size of cosmetics on attractiveness in two sets of faces that have had  
102 cosmetics applied professionally, with the prediction that the effect will be substantially larger  
103 than the previous assessment that considered self-applied cosmetics [15]. In addition, by using a  
104 similar design to previous research, we can draw direct comparisons with current knowledge of  
105 how cosmetics and identity affect attractiveness.

106 A separate but related question regarding cosmetics concerns how it affects faces of  
107 different levels of attractiveness. Many studies in the literature on cosmetics and social

108 perceptions have used models recruited from university or college [14,15,20]. How do cosmetics  
109 affect faces of a different population, specifically faces considered to be very attractive? Previous  
110 research found no interaction between cosmetics and identity [15], suggesting cosmetics affect  
111 each face's attractiveness similarly. However, the models used were of a university-aged sample  
112 of population-typical attractiveness levels. The present studies, particularly Study 2, examine the  
113 effect cosmetics have on perceived attractiveness in a sample of women typically considered to  
114 be very attractive – models. Using a sample of faces that are already constrained in attractiveness  
115 enables us to manipulate another source of variation in attractiveness, specifically between-  
116 person variability. As such, we can observe the effects of cosmetics on attractiveness in a sample  
117 with a (hypothesised) lower effect of identity (differences between individuals) than elsewhere.

118         The present study has several aims. First, we examine how cosmetics affect attractiveness  
119 when cosmetics have been professionally-applied. We predict that cosmetics will have a notably  
120 larger effect size in this sample compared to the previous study examining this question [15].  
121 Second, we consider the effect size of cosmetics in sets of faces that are considered highly  
122 attractive, where between-person variation (identity effect size) should be reduced. The relative  
123 effect size of cosmetics may therefore be increased, and may be more likely to overshadow the  
124 smaller between-person variation in attractiveness. Conversely, cosmetics may have less of an  
125 effect in these samples as the women are already at the higher end of attractiveness without  
126 cosmetics, leaving little room for judgements of attractiveness to increase when cosmetics are  
127 applied. Finally, by using an identical design to previous research [15], we will compare the  
128 findings obtained in these studies to those presented in previous research in order to build a fuller  
129 picture of the relative importance of cosmetics and identity in attractiveness perceptions.

130

## 131 **Study 1**

132           In the first study, we examine how cosmetics impact attractiveness when they are applied  
133 professionally. To do this, we take advantage of an Internet-based sample to acquire images of  
134 models whose cosmetics have been applied by high-profile makeup artists. Compared to  
135 previous work examining this question [15], we predict that the effect size due to cosmetics  
136 should be larger here. However, the effect size of identity may still overshadow it.

137

## 138 **Method**

### 139 **Participants**

140           Ninety North American university students (age  $M = 18.57$  years,  $SD = 0.75$ , 41 men)  
141 participated in the main study for course credit. Due to a software error, age data was not  
142 recorded for the first 50 participants, with the mean age being calculated from the remaining  
143 participants. However, all participants were within the same demographic and age range. A  
144 further 15 students (age  $M = 19.93$  years,  $SD = 1.16$ , three men) rated the quantity of cosmetics  
145 worn by the models. Informed consent was obtained from all participants included in the study.

146

### 147 **Ethics Statement**

148           Ethical approval for all studies was obtained from the Gettysburg College institutional  
149 review board (IRB). All participants gave written informed consent before beginning the study.

150

### 151 **Stimuli**



152 From the YouTube website, we collected images of White British women ( $n = 33$ , age  
153 unknown but approximately 20-35 years), who acted as models while their cosmetics were  
154 applied by high-profile professional makeup artists from the United Kingdom. Twenty-three  
155 models were obtained from one artist's channel ([www.youtube.com/user/lisaeldridgedotcom](http://www.youtube.com/user/lisaeldridgedotcom))  
156 with a further ten collected from another ([www.youtube.com/user/ctilburymakeup](http://www.youtube.com/user/ctilburymakeup)). We utilised  
157 all available videos at the time of writing that featured a model receiving a makeover where they  
158 were shown before and after an application of cosmetics. In addition, we included only videos  
159 where faces began free of cosmetics, and the artist had the intention of applying a particular  
160 cosmetics look, rather than with the aim of hiding blemishes or skin conditions (such as acne).  
161 Images were captured from video tutorials, which served to instruct viewers on a number of  
162 popular cosmetics styles for a range of scenarios. Both authors classified the cosmetics looks into  
163 categories using information provided by descriptions within the videos. Three categories were  
164 apparent – an everyday, natural look ( $n = 7$ ), a 'going out' look ( $n = 14$ ), and vintage or editorial  
165 looks based on cosmetics the makeup artist had applied during professional photo shoots in the  
166 past ( $n = 12$ ). A third researcher, with extensive experience in this field, arrived at these three  
167 categories independently, providing further confirmation.

168 We captured a high-resolution screenshot of each model at the end of each video, where  
169 images of the models were presented before and after their application of cosmetics side-by-side.  
170 Models had a neutral expression and looked directly into the camera for the comparison. In  
171 addition, the two photographs were taken under the same lighting and camera conditions. From  
172 each comparison screenshot, we cropped the 'before' and 'after' versions of each model to  
173 produce two separate images. Final images were cropped just below the chin, at the hairline (or  
174 mid-forehead based on the limitations of the original), and tight to the widest part of the face

175 (and so removing the ears). Given the variable nature of the images in terms of hairstyle, we  
176 chose models whose hair did not occlude their faces, and we masked loose hair in the lower  
177 portions of the images if it was not tied back. Images were resized to a height of 451 pixels.  
178 Given copyright restrictions, we present the average of models without cosmetics, and separately  
179 with cosmetics, in Fig 1 to illustrate. Averages were produced using JPsychomorph after  
180 landmarks were applied to the facial features in each image [32].

181

182 **Fig 1.** The average model without (left) and with cosmetics (right). These averages are cropped  
183 mid-forehead because several of the YouTube videos presented individuals in this way, resulting  
184 in insufficient information above this point for generating averages.

185

## 186 **Procedure**

187 Participants rated the attractiveness of the models using custom PsychoPy software [33].  
188 Images were presented in a random order, and each participant rated each model only once, in a  
189 randomly selected cosmetics condition (i.e., either with or without cosmetics). This design was  
190 specifically chosen to prevent carryover effects between conditions [15,29]. Participants rated  
191 the attractiveness of the models on a 1 (*very unattractive*) to 7 (*very attractive*) scale, indicating  
192 their response via mouse click. Stimuli remained onscreen until a judgement was made.

193 A separate sample of participants judged the quantity of cosmetics worn by the models.  
194 These participants saw the ‘without’ and ‘with cosmetics’ images onscreen next to each other,  
195 and were asked ‘how much makeup has been applied to this face?’ Participants indicated their  
196 responses via mouse click on a 1 (*very light*) to 7 (*very heavy*) scale. Trials were presented in a  
197 random order. Though this is only a perceived measure of quantity, rather than an actual quantity

198 of cosmetics, we believe it to be suitable as it is the perceived quantity that would affect the  
199 perceptions of observers. Importantly, other studies have found general agreement in the quantity  
200 of cosmetics applied by a professional makeup artist and the perceived amount of cosmetics  
201 being worn [31].

202

## 203 **Results**

204 Each image was rated an average of 45 times ( $SD = 4.45$ ). We examined agreement by  
205 calculating the pooled standard deviation for ratings in each cosmetics condition; without  $SD_p =$   
206 1.34; with cosmetics  $SD_p = 1.44$ . Responses were given on a 7-point scale, so the generally low  
207 variability indicates good agreement in ratings [15,34]. To examine effects of observer sex on  
208 ratings, the data were split by the sex of each observer before averaging. This resulted in four  
209 scores for each model – one in each cosmetics condition, as rated by men and women.

210 We also calculated the average amount of perceived cosmetics applied ( $M = 4.96$ ,  $SD =$   
211 1.09), as judged by the separate sample of raters. These judgements of quantity were collected in  
212 order to be able to control for the varying amounts of cosmetics worn by each model in our  
213 analyses. However, this measure showed no relationship with the dependent variable  
214 (attractiveness) at all levels of observer sex and cosmetics, all  $r_s < .25$ ,  $p_s > .160$ . As such, there  
215 was no reason to include quantity as a covariate, and we therefore analysed our results using a  
216 repeated measures ANOVA with model as the unit of analysis.

217 We focus here on the effect sizes of variables in order to estimate the real world effect of  
218 cosmetics on attractiveness. In particular, we utilise eta squared ( $\eta^2$ ) as a measure of effect size,  
219 which expresses how much each factor contributes to the total variance in attractiveness ratings  
220 as an interpretable percentage value, rather than partial eta squared, which does not sum across

221 factors to one. We calculated  $\eta^2$  effect sizes for both main effects (Cosmetics, Observer Sex) and  
 222 the interaction by dividing the sums of squares (SS) attributable to each effect by the total SS,  
 223 calculated by summing the SS attributable to each effect and their respective errors. We also  
 224 gave special consideration to the variance attributable to differences between items. This  
 225 variation is typically ignored in repeated measures analyses since it usually represents variation  
 226 between participants on the measured dependent variable, which is generally unimportant for  
 227 repeated measures designs (which instead focus on variation *within* participants). However, in  
 228 this case, it takes on a useful property. By using the images of the models as the unit of analysis,  
 229 the variation between models represents variation in attractiveness arising due to the fact that  
 230 models have different facial identities or appearances. We were therefore able to calculate an  
 231 effect size for this ‘identity’ measure. The full results of the ANOVA are reported in Table 1,  
 232 illustrating the effect sizes, their associated SS, and other statistics.

233

234 **Table 1.** Results of the analysis of variance from Study 1.

Source	<i>df</i>	<i>SS</i>	$\eta^2$	<i>F</i>	<i>p</i>
Identity (I)	32	61.27	0.45		
Observer Sex	1	0.85	0.01	10.03	.003
Observer Sex $\times$ I	32	2.70	0.02		
Cosmetics	1	44.83	0.33	76.33	< .001
Cosmetics $\times$ I	32	18.79	0.14		
Observer Sex $\times$ Cosmetics	1	1.29	0.01	8.17	.007
Observer Sex $\times$ Cosmetics $\times$ I	32	5.05	0.04		
Total	131	134.78			

235

236 There is no error term for conducting an  $F$  test on differences between models, and as such, no ratio is calculated.  $df$   
237 = degrees of freedom,  $SS$  = sums of squares. Interactions with the Identity measure can be interpreted as an error  
238 term for that variable [35].

239

240 Men assigned lower ratings of attractiveness ( $M = 3.74$ , 95% CI [3.47, 4.00]) than women  
241 ( $M = 3.89$ , [3.66, 4.13]), a result consistent with previous literature [15,36,37] which we do not  
242 pursue further here. Importantly, models were rated as more attractive with cosmetics ( $M = 4.39$ ,  
243 [4.11, 4.68]) than without ( $M = 3.23$ , [2.95, 3.51]). The Observer Sex x Cosmetics interaction  
244 was driven by men rating faces without cosmetics as less attractive than women rating those  
245 same faces,  $t(32) = 4.32$ ,  $p < .001$ ,  $d = 0.75$ , but both sexes assigned similar ratings for models  
246 with cosmetics,  $t(32) = 0.42$ ,  $p = .676$ ,  $d = 0.07$ , indicating a larger influence of cosmetics on  
247 attractiveness for men. However, the effect size of this interaction was very small ( $\eta^2 = 0.01$ ),  
248 suggesting a relatively unimportant result.

249 Of more importance was the Cosmetics x Identity interaction ( $\eta^2 = 0.14$ ), which indicates  
250 that the application of cosmetics altered the attractiveness of individual models differently. To  
251 examine this further, we computed a difference score for each model between their attractiveness  
252 with and without cosmetics, as rated by men and women. This difference illustrates the boost in  
253 attractiveness conferred by cosmetics, and we carried out a correlation between these values and  
254 the attractiveness of the models without cosmetics. Ratings assigned by both women and men  
255 showed a negative correlation between these values,  $r(31) = -.53$ , 95% CI [-.73, -.23],  $p = .001$ ,  
256 and  $r(31) = -.48$ , [-.71, -.16],  $p = .005$ , respectively, indicating that the more attractive a model  
257 was, the less of an increase in attractiveness cosmetics conferred, a pattern which did not change  
258 when combining ratings given by men and women,  $r(31) = -.46$ , [-.69, -.14],  $p = .007$  (see Fig 2).

259

260 Fig 2. An illustration of the average attractiveness (combining ratings made by men and women)  
261 of each model, both without cosmetics and with cosmetics. Models are ordered in terms of  
262 increasing attractiveness without cosmetics. An upward pointing arrow indicates an increase in  
263 attractiveness with cosmetics, while a downward arrow indicates a decrease.

264

265 Table 1 illustrates that the Identity effect size ( $\eta^2 = 0.45$ ) is 1.36 times larger than the  
266 effect size attributed to Cosmetics ( $\eta^2 = 0.33$ ). The differences in attractiveness between  
267 individuals explains more variance than an application of cosmetics, but the ratio of these effect  
268 sizes is much smaller than in previous accounts [15]. This suggests that a professional  
269 application of cosmetics (in comparison with self-application) is capable of producing a larger  
270 effect on attractiveness perceptions, although this remains smaller than the effect due to identity  
271 differences between women.

272 We conducted a final analysis to examine whether the cosmetics ‘look’ ascribed by the  
273 artist affected perceptions of attractiveness differently for men and women. The above analysis  
274 was repeated, but with the addition of ‘look’ as a source of variation between models. The three-  
275 way mixed model ANOVA revealed no significant main effects of cosmetics look or interactions  
276 with this factor, all  $F_s < 1.18$ ,  $p_s > .320$ . However, it is worth noting that the ‘cosmetics look’  
277 variable had low power (ranging from .076 to .242 across main effects and interactions), so  
278 further study is required to investigate the role of cosmetics look in perceived attractiveness.

279

## 280 **Study 2**

281           The models used in Study 1 were women who had agreed to participate for the purposes  
282 of demonstration in a makeup tutorial. We have shown that the effect of cosmetics, when  
283 professionally-applied, results in a larger effect size compared with previous research [15]. Next,  
284 we investigate how cosmetics alter the attractiveness of a sample of women who are generally  
285 regarded as very attractive and earn a living based on their appearance – supermodels. We  
286 examine how much variation in attractiveness can be explained by cosmetics, and compare it  
287 with the effect size of identity, the differences in attractiveness between supermodels. Here, the  
288 effect size of identity should be smaller, given the potentially homogenous nature of the women  
289 in terms of attractiveness. How much of a benefit do cosmetics confer to highly attractive  
290 women, and in turn, do cosmetics overcome the differences in attractiveness between  
291 individuals?

292

## 293 **Method**

### 294 **Participants**

295           One hundred new participants completed the study for course credit (age  $M = 19.28$   
296 years,  $SD = 1.46$ , 46 men), 14 of which were students at a Scottish university (age  $M = 19.28$   
297 years,  $SD = 1.68$ , one man), while the rest were students at a North American university (age  $M$   
298 = 19.28 years,  $SD = 1.05$ , 45 men). A further sample of 14 North American students from the  
299 same university (age  $M = 20.50$  years,  $SD = 1.28$ , 2 men) rated the quantity of cosmetics worn by  
300 the models.

301           The removal of the 14 participants from the Scottish university (who live in the UK rather  
302 than the US) did not change the pattern of results described below, aside from producing a  
303 significant main effect of Observer Sex,  $F(1, 44) = 18.64$ ,  $p < .001$ ,  $\eta^2 = .02$ . As in Study 1, men

304 provided lower ratings of attractiveness ( $M = 4.07$ , [3.91, 4.23]) than women ( $M = 4.32$ , [4.12,  
305 4.53]). However, as this is a well-demonstrated effect and did not alter the presence of the  
306 interaction between cosmetics and observer sex, we include these extra participants for the  
307 additional validity they confer.

308

### 309 **Ethics Statement**

310 Ethical approval for all studies was obtained from the Gettysburg College institutional  
311 review board (IRB). All participants gave written informed consent before beginning the study.  
312 The Ethical Governance and Approval System at the University of Aberdeen granted approval  
313 for the study conducted there. Again, all participants gave written informed consent before  
314 beginning the study.

315

### 316 **Stimuli**

317 We collected images ( $n = 45$ ) of supermodels without their makeup from the Internet.  
318 These images were casting photographs for Louis Vuitton's Fall-Winter 2010 runway show. All  
319 pictures were taken with the models looking directly into the camera, with a neutral expression.  
320 We then collected images of the same women wearing cosmetics from professional photo shoots,  
321 and selected images where they had a neutral expression and were looking directly into the  
322 camera in order to match the casting photographs as closely as possible. However, these  
323 cosmetics photos were considerably less constrained in that the lighting varied between images,  
324 as did the amount of time between the two photos for each model. Therefore, while every care  
325 was taken to ensure similarity between these images and those of Study 1, we note that such  
326 limitations mean that any conclusions drawn from this study are necessarily more tentative.



327 Final images were cropped as in Study 1 to just below the chin, at the hairline, and tight  
328 to the widest part of the face (and so removing the ears). Hair was masked at the bottom of the  
329 images as before, and images were resized to a height of 250 pixels. Given copyright restrictions,  
330 we present the average of supermodels without cosmetics, and separately with cosmetics, in Fig  
331 3 to illustrate.

332

333 **Fig 3.** The average supermodel without (left) and with cosmetics (right).

334

### 335 **Procedure**

336 We used the same procedure as in Study 1. However, given that the photographs were of  
337 supermodels, there was a chance they would be recognised by observers. As such, we added a  
338 ‘recognise’ option onscreen where participants could indicate their recognition of the model  
339 rather than providing a rating of attractiveness. Familiarity with the models may result in  
340 unwanted influences on ratings. Across all images, an average of 4.95 trials were skipped ( $SD =$   
341  $2.97$ ). Ratings of attractiveness were therefore only collected for models that were not recognised  
342 by the raters.

343

### 344 **Results**

345 Each image received an average of 50 ratings ( $SD = 4.68$ ). Agreement was calculated as  
346 before, using a pooled standard deviation for ratings within each cosmetics condition, and  
347 showed generally higher levels (lower variation) than Study 1; without cosmetics  $SD_p = 1.11$ ;  
348 with cosmetics  $SD_p = 1.32$ . We split the data by the sex of the observer as before, and computed  
349 the average rating for each model in both cosmetics conditions as assigned by men and women.

350 We then averaged the ratings of quantity assigned by the separate sample of raters ( $M =$   
351  $4.29$ ,  $SD = 1.17$ ) for use as a covariate in subsequent analyses. However, as in Study 1, the  
352 quantity measure showed no relationship with the dependent variable at any levels of each  
353 independent variable, all  $r_s < .11$ ,  $p_s > .476$ . As such, analyses were carried out without inclusion  
354 of this covariate using repeated measures ANOVA, the results of which are summarised in Table  
355 2. We also compared the perceived quantity ratings of the faces in Study 1 to the faces here,  
356 finding that the sample of supermodels ( $M = 4.29$ ,  $[3.95, 4.63]$ ) were perceived as wearing less  
357 cosmetics than the YouTube models ( $M = 4.96$ ,  $[4.57, 5.36]$ ),  $t(76) = 2.56$ ,  $p = .012$ ,  $d = 0.59$  .

358

359 **Table 2.** Results of the analysis of variance from Study 2.

Source	<i>df</i>	<i>SS</i>	$\eta^2$	<i>F</i>	<i>p</i>
Identity (I)	44	58.08	0.43		
Observer Sex	1	0.37	0.00	2.79	.102
Observer Sex $\times$ I	44	5.85	0.04		
Cosmetics	1	33.35	0.25	47.89	< .001
Cosmetics $\times$ I	44	30.64	0.23		
Observer Sex $\times$ Cosmetics	1	1.73	0.01	20.66	< .001
Observer Sex $\times$ Cosmetics $\times$ I	44	3.68	0.03		
Total	179	133.70			

360

361 There is no error term for conducting an *F* test on differences between models, and as such, no ratio is calculated. *df*  
362 = degrees of freedom, *SS* = sums of squares. Interactions with the Identity measure can be interpreted as an error  
363 term for that variable [35].

364

365 As before, models were rated as more attractive with cosmetics ( $M = 4.53$ , [4.28, 4.77])  
366 than without ( $M = 3.67$ , [3.49, 3.85]). The Observer Sex x Cosmetics interaction was again  
367 driven by men rating faces without cosmetics as less attractive than women rating those same  
368 faces,  $t(44) = 4.65$ ,  $p < .001$ ,  $d = 0.69$ , with both sexes perceiving the models as similarly  
369 attractive with cosmetics,  $t(44) = 1.37$ ,  $p = .176$ ,  $d = 0.21$ . However, as before, the effect size of  
370 this interaction was small ( $\eta^2 = 0.01$ ).

371 The effect size of the Cosmetics x Identity interaction ( $\eta^2 = 0.23$ ) was almost as large as  
372 the effect of cosmetics itself ( $\eta^2 = 0.25$ ), indicating the application of cosmetics affected the  
373 attractiveness of the supermodels differently. As before, we computed a difference score (for  
374 men and women's ratings separately) between cosmetics conditions, and correlated this score  
375 with the attractiveness of the supermodels without cosmetics. Again, there was a negative  
376 correlation between the boost in attractiveness with cosmetics and the attractiveness of the  
377 models without cosmetics, for both women  $r(43) = -.40$ , [-.62, -.12],  $p = .006$ , and men  $r(43) = -$   
378  $.42$ , [-.64, -.14],  $p = .004$ , as well as when ratings given by both sexes were combined,  $r(43) = -$   
379  $.40$ , [-.62, -.12],  $p = .004$ . As before, this indicates that the more attractive the supermodel is  
380 perceived to be, the less of a boost in attractiveness cosmetics confer. This is illustrated in Figure  
381 4.

382  
383 **Fig 4.** An illustration of the average attractiveness (combining ratings made by men and women)  
384 of each model, both without cosmetics and with cosmetics. Models are ordered in terms of  
385 increasing attractiveness without cosmetics. An upward pointing arrow indicates an increase in  
386 attractiveness with cosmetics, while a downward arrow indicates a decrease.

387

388           The effect size of Identity ( $\eta^2 = 0.43$ ), due to differences in attractiveness between  
389 supermodels, was 1.72 times greater than that of cosmetics ( $\eta^2 = 0.25$ ), a ratio slightly larger than  
390 that observed in Study 1. Differences in attractiveness between individuals explained more  
391 variance than cosmetics, even among a sample of women regarded as highly attractive. The  
392 effect size of Cosmetics in this study was smaller than that found in Study 1, suggesting that  
393 cosmetics may produce a smaller increase in attractiveness for women who are already at the top  
394 end of the attractiveness scale, although the mean ratings for faces do not suggest a ceiling effect.

395           It is also worth noting that the effect size of Identity in this study was very close to the  
396 value reported in Study 1, which goes against our prediction that a sample of supermodels should  
397 have smaller between-person variability in attractiveness. However, this value is still notably  
398 smaller than the variation between individuals in a sample of university students [15], supporting  
399 our prediction of reduced between-person variability.

400

## 401 **Combined Analyses**

402           We have shown that professionally-applied cosmetics increase the attractiveness of both  
403 models and supermodels, with generally larger effect sizes than have been observed elsewhere  
404 [15]. Here, we combine the data from Study 1 with the data reported in previous work that  
405 provided an estimate of the effect size of cosmetics when self-applied to a student population  
406 [15]. This will allow a comparison of both model sets without and with cosmetics, and an overall  
407 comparison of the effect size of cosmetics and identity in a pooled setting of cosmetics use. We  
408 included only the models from Study 1 as these images were captured under more controlled  
409 conditions, similar to the images used in the previous work. In the initial study [15], there were  
410 44 self-reported White women acting as models (age  $M = 21.18$ ,  $SD = 1.94$ ). Models applied

411 their own cosmetics from a range of provided products, and were rated using the same procedure  
412 used here. See [15] for full details.

413 To conduct this analysis, we employed a three-way mixed ANOVA: Set (Students,  
414 YouTube)  $\times$  Cosmetics (With, Without)  $\times$  Observer Sex (Female, Male). Set represented a  
415 between-subjects factor, while the remaining factors were both within-subjects. As before, the  
416 model was the unit of analysis. Since a factorial ANOVA produces several statistical tests, we  
417 focus on the theoretically important outcomes. In this case, an interaction between Set and  
418 Cosmetics indicates that an application of cosmetics affects the model sets differently. We would  
419 predict models that received an application of professional cosmetics would appear more  
420 attractive.

421 As observed across the original data [15] and the two studies presented here, there was a  
422 main effect of Observer Sex,  $F(1, 75) = 122.45, p < .001, \eta^2 = .04$ , following the usual pattern of  
423 men ( $M = 3.21, [3.05, 3.37]$ ) assigning lower ratings than women ( $M = 3.61, [3.44, 3.78]$ ).  
424 Models also received higher ratings of attractiveness when viewed with cosmetics ( $M = 3.76,$   
425  $[3.58, 3.95]$ ) compared to when they were viewed without ( $M = 3.05, [2.88, 3.23]$ ),  $F(1, 75) =$   
426  $97.35, p < .001, \eta^2 = .12$ . Models from the YouTube set were also rated as more attractive ( $M =$   
427  $3.82, [3.57, 4.06]$ ) than those in the student set ( $M = 3.00, [2.79, 3.21]$ ),  $F(1, 75) = 24.95, p <$   
428  $.001, \eta^2 = .16$ .

429 The predicted interaction between Set and Cosmetics was present,  $F(1, 75) = 40.59, p <$   
430  $.001, \eta^2 = .05$ . Bonferroni adjusted post-hoc tests revealed that without cosmetics, the YouTube  
431 models ( $M = 3.23, [2.97, 3.49]$ ) were rated as slightly more attractive than models from the  
432 student set ( $M = 2.87, [2.65, 3.09]$ ),  $p = .041, d = 0.24$ . However, with cosmetics, YouTube  
433 models ( $M = 4.39, [4.12, 4.67]$ ) received significantly higher ratings of attractiveness than the

434 student models ( $M = 3.12$ , [2.88, 3.37]),  $p < .001$ ,  $d = 0.79$ , indicating a larger change in  
435 attractiveness with professionally-applied cosmetics than with self-applied cosmetics.

436 We can also draw comparisons between the sizes of our effects across all three studies  
437 (the two presented here and the student set). While  $\eta^2$  is ideal for comparing effect sizes within a  
438 study (the total always sums to 100%), comparison between studies is generally not  
439 recommended because the total variability depends on the study design and the number of  
440 independent variables [38]. However, the two studies reported here, as well as earlier data [15],  
441 use identical study designs, and the total variability is very similar in all cases (Study 1  $SS$  total =  
442 134.76, Study 2  $SS$  total = 133.70, [15]  $SS$  total = 129.23). The main differences were the models  
443 used and the type of cosmetics applied. As such, we can justifiably make some comparisons  
444 between the effect sizes of cosmetics and identity across these studies.

445 While the effect size due to identity was similar in Studies 1 and 2 ( $\eta^2 = .45$  and  $\eta^2 = .43$ ,  
446 respectively), the earlier study using students showed a much larger effect ( $\eta^2 = .69$ ). The effect  
447 size of cosmetics in Studies 1 and 2 ( $\eta^2 = .33$  and  $\eta^2 = .25$ , respectively), in contrast, were much  
448 larger than in the student study ( $\eta^2 = .02$ ). Therefore, while variation in attractiveness between  
449 individuals was somewhat greater among a sample of university students as compared to models  
450 and supermodels (as we would expect), the effect size of professionally-applied cosmetics was  
451 much larger than self-applied cosmetics. It is also important to note that the effect sizes obtained  
452 for the data in Study 2 are to be interpreted cautiously, given the more unconstrained nature of  
453 the images.

454

## 455 **General Discussion**

456           Across several studies, we find that using cosmetics increases perceptions of  
457 attractiveness compared to no cosmetics, with several novel findings and caveats. First, we show  
458 that the effect size of cosmetics on attractiveness is large when those cosmetics have been  
459 professionally-applied, though the effect of identity is still greater. However, the difference  
460 between identity and cosmetics effects is much smaller than in a student sample of faces with  
461 self-applied cosmetics [15]. Second, we show that in a sample of supermodels with a smaller,  
462 more constrained effect size of identity (i.e., reduced between-person variance in attractiveness),  
463 identity is still more important than cosmetics, though the effect size of cosmetics is still larger  
464 than in previous cases. In both cases, but particularly the set of supermodels, we found evidence  
465 of an interaction between facial identity and cosmetics, indicating a differential effect of  
466 cosmetics on attractiveness. Further analysis revealed that the more attractive a face was without  
467 cosmetics, the less of an increase in attractiveness cosmetics conferred.

468           Across all studies, we observed that the effect of facial identity was larger than the effect  
469 of cosmetics. This finding extends previous research demonstrating that between-person  
470 variation is consistently larger than within-person manipulations of attractiveness [14, 28].  
471 Interestingly, the ratio between the effect sizes of identity and cosmetics in these studies (i.e.,  
472 how much more variation identity explained than cosmetics in attractiveness judgements) is  
473 smaller than the comparison observed with emotional expression [29], suggesting that  
474 professionally-applied cosmetics might be more effective at modulating attractiveness  
475 perceptions than facial expression, at least in female faces. Additionally, the finding that identity  
476 might be more important than within-person variation should perhaps be interpreted with  
477 caution. We refer to ‘identity’ in the current paper but use single, passport-style images of each

478 model. However, individuals appear differently across different photographs, and this within-  
479 person variation in appearance has also been shown to affect perceived attractiveness [42].

480 A surprising source of variance in both studies was the interaction between identity and  
481 cosmetics. This finding, indicating that cosmetics affected different faces differently, was  
482 analysed further to reveal that the more attractive a face was initially, the less of an increase in  
483 attractiveness cosmetics conferred. While this is an intuitive finding, it has not been  
484 demonstrated before, and was particularly pronounced in the set of supermodels where the effect  
485 size of the interaction was almost as large as that of cosmetics itself. Cosmetics confer attractive  
486 patterns of colouration to faces, enhancing sex typical features in skin reflectance [18,20], as  
487 well as smoothing skin homogeneity and colour distribution [22,24,43]. Female faces that are  
488 considered attractive tend to have lighter skin, darker eyes, and redder lips than the average  
489 female face [44], which are all correlates of attractiveness [20,21], and in a recent study, are  
490 colourations that are conferred to faces by cosmetics [18]. It may be that the more attractive  
491 faces (i.e., of supermodels) already possess the most attractive features that cosmetics can alter,  
492 and so there is little change in attractiveness after an application. That less attractive faces  
493 receive more of an increase from cosmetics also has practical implications. By definition, the  
494 majority of women will lie around average attractiveness, and so a significant number of women  
495 could receive a boost in attractiveness from cosmetics.

496 We also found that the perceived quantity of cosmetics applied to faces played almost no  
497 role in the perceived attractiveness of faces with cosmetics. Recent evidence has shown that  
498 faces with lighter makeup are perceived as more attractive than faces with heavier makeup [45],  
499 which is at odds with our findings here. However, that study used different models for each  
500 cosmetics condition, conflating sources of cosmetics and identity variance, as well as using



501 digitally applied cosmetics. While observers seem to find lighter cosmetics optimally attractive  
502 when given the choice to vary the quantity [16], no study as of yet has systematically shown that  
503 lighter cosmetics are optimally attractive for a given face. Our measurements here, as well as  
504 previous data [15], seem to suggest quantity does not play a large role in perceptions of  
505 attractiveness with cosmetics.

506         Combining image sets from previous research [15] with the findings from Study 1  
507 revealed that, while the models from Study 1 were slightly more attractive than the models from  
508 the previous study, they were rated as significantly more attractive with cosmetics. After  
509 considering the similarity of designs and total variability across all studies (both here and in  
510 [15]), we compared the effect sizes of identity and cosmetics directly. Variability due to  
511 attractiveness between individuals (identity) was smaller among models and supermodels  
512 compared to university students, as predicted, but the effect size of cosmetics was noticeably  
513 larger for professionally-applied cosmetics. However, it is important to note that the sample sizes  
514 of models differed, and larger sample sizes might also result in greater between-person  
515 variability.

516         These findings have relevance for investigating the effects of cosmetics on social  
517 perceptions. There now exist estimates of the effect size of cosmetics when they are self-applied  
518 [15], and when they are applied professionally. In previous work [15], cosmetics explained just  
519 2% of the variation in attractiveness, while the finding from a sample of models showed  
520 cosmetics explained 33% of the variation in attractiveness. This study demonstrated larger effect  
521 sizes of cosmetics when directly compared to previous research [15], though the studies used  
522 different sets of faces, and it is important to note that any effect size estimate calculated is  
523 ultimately based on the context of the research, and should be interpreted within this context

524 [46]. However, the variances in the current and previous research are very similar, and the design  
525 of the studies is identical, meaning direct comparisons are valid and appropriate.

526 The literature examining the effect of cosmetics on social perceptions has, for the most  
527 part, used models with professionally-applied cosmetics in laboratory studies [9,10,17,30,31] as  
528 well as field experiments [12,13,39,40]. With our comparison of the effect size of cosmetics  
529 under both self-applied and professionally-applied conditions, it seems possible that some of the  
530 effects of cosmetics observed in the literature may be inflated. Further, women report higher self-  
531 confidence and engage in more social activities after a professional makeover [41] and this  
532 increase in self-confidence may translate into slight expression or postural differences in images,  
533 which could represent an additional within-person boost in attractiveness due to cosmetics.

534 There are some caveats to the study. Images were obtained from various Internet sources,  
535 and so were not as constrained in lighting or emotional expression as previous research [15].  
536 Study 1 suffered less from this potential issue as images were collected from the same  
537 photographic session. As the images of supermodels with cosmetics were obtained from different  
538 sources, while the images of those women without cosmetics were obtained from the same  
539 source, the magnitude of the interaction between identity and cosmetics should be interpreted  
540 with caution. However, given its presence in Study 1 with more controlled stimuli, we think it  
541 safe to conclude that cosmetics affect more attractive individuals to a lesser extent than others.  
542 Furthermore, that such an effect was obtained in Study 2 with more variable photographs could  
543 be considered strong evidence. Since the images were more variable and cosmetics were  
544 confounded with variations in lighting (both considered noise in the current study), it seems  
545 likely an effect would be obtained under stricter conditions.

546           There now exists convincing evidence that alterations to within-person facial appearance  
547 via cosmetics, whether self-applied or professionally-applied, do not overcome between-person  
548 variability in attractiveness due to simple identity. Facial attractiveness is, to an extent, more  
549 about what you have, rather than what you do with it. However, we have uncovered here  
550 interesting caveats to this overarching and consistent finding. An increased skill level in applying  
551 cosmetics seems to offer a larger increase in attractiveness than self-applied cosmetics does –  
552 larger effects were clear when a professional makeup artist applied cosmetics. Furthermore, we  
553 have shown cosmetics affect faces of varying levels of attractiveness differently, particularly  
554 within a sample of faces with lower variation in attractiveness between individuals. More  
555 attractive individuals simply have less to gain from using cosmetics. These findings have  
556 theoretical implications for attractiveness research. Cosmetics is perhaps the most common form  
557 of modification of facial appearance, and we have shown that the currently reported literature,  
558 with its reliance on professionally-applied cosmetics, highlights an effect that does not seem  
559 achievable through everyday use.

560           How cosmetics affect attractiveness is a growing literature, and many studies use  
561 professionally-applied cosmetics as a means to examine this change. We have shown that  
562 professionally-applied cosmetics seem to explain a larger proportion of variation in  
563 attractiveness judgements than self-applied cosmetics, a category which the vast majority of  
564 cosmetics users fall under. This could suggest an inflation of the effect of cosmetics in the  
565 current literature, with cosmetics increasing attractiveness beyond what is achievable through  
566 everyday means. Additionally, we have illustrated that cosmetics affect women differently –  
567 more attractive women, particularly supermodels, gain less of a boost in attractiveness from  
568 cosmetics than do less attractive women. Importantly, the effect size of identity, or between-

569 person variance in attractiveness, was larger than the effect of cosmetics in both studies. We  
570 conclude that, when it comes to cosmetics, individual differences in facial appearance are  
571 ultimately more important than even a professional application of cosmetics.

572

### 573 **Acknowledgements**

574 The authors would like to thank Mila Mileva for help in collecting and preparing stimuli,  
575 testing participants, and providing advice on cosmetics looks. Thanks also go to Sarah Kramer  
576 for her help in participant testing and recruitment.

577

### 578 **Ethical approval**

579 All procedures performed in the studies involving human participants were in accordance  
580 with the ethical standards of the institutional and/or national research committee and with the  
581 1964 Helsinki declaration and its later amendments or comparable ethical standards.

582

### 583 **Author Contributions**

584 Conceived and designed the studies: ALJ RSSK. Materials collection: ALJ RSSK.  
585 Performed the studies: ALJ RSSK. Analysed data: ALJ RSSK. Wrote the paper: ALJ RSSK.

586

587

## References

- 588 1. Brain R. *The decorated body*. London: Hutchinson & Co; 1979.
- 589 2. Jablonski NG. *Skin: A natural history*. Berkeley: University of California Press; 2006.
- 590 3. Russell R. Why Cosmetics Work. In: Adams RB, Ambady A, Nakayama K, Shimojo S,  
591 editors. *The Science of Social Vision*. Oxford: Oxford University Press; 2010. pp. 186–204.
- 592 4. Etcoff N. *Survival of the prettiest: The science of beauty*. New York: Anchor Books; 1999.
- 593 5. Rossi E, Prlic A, Hoffman R. A study of the European cosmetics industry [Internet]. 2007.  
594 Available: [ec.europa.eu/DocsRoom/documents/13125/attachments/1/.../pdf](http://ec.europa.eu/DocsRoom/documents/13125/attachments/1/.../pdf)
- 595 6. Korichi R, Pelle-de-Queral D, Gazano G, Aubert A. Why women use makeup: implication  
596 of psychological traits in makeup functions. *J Cosmet Sci*. 2008;59: 127–137.
- 597 7. Miller LC, Cox CL. For appearances' sake: Public self-consciousness and makeup use. *Pers*  
598 *Soc Psychol Bull*. 1982;8: 748–751.
- 599 8. Robertson J, Fieldman G, Hussey T. “Who wears cosmetics?” Individual differences and  
600 their relationship with cosmetic usage. *Individ Differ Res*. 2008;6: 38–56.
- 601 9. Nash R, Fieldman G, Hussey T, Lévêque J-L, Pineau P. Cosmetics: They influence more  
602 than Caucasian female facial attractiveness. *J Appl Soc Psychol*. 2006;36: 493–504.  
603 doi:10.1111/j.0021-9029.2006.00016.x
- 604 10. Etcoff N, Stock S, Haley LE, Vickery SA, House DM. Cosmetics as a feature of the  
605 extended human phenotype: modulation of the perception of biologically important facial  
606 signals. Suzuki S, editor. *PLoS ONE*. 2011;6: e25656. doi:10.1371/journal.pone.0025656
- 607 11. Mileva VR, Jones AL, Russell R, Little AC. Sex differences in the perceived dominance  
608 and prestige of women with and without cosmetics. *Perception*. 2016; 0301006616652053.  
609 doi:10.1177/0301006616652053
- 610 12. Guéguen N, Jacob C. Enhanced female attractiveness with use of cosmetics and male  
611 tipping behavior in restaurants. *J Cosmet Sci*. 2011;62: 283–290.
- 612 13. Guéguen N. The effects of women's cosmetics on men's approach: An evaluation in a bar.  
613 *North Am J Psychol*. 2008;10: 221–228.
- 614 14. Cash TF, Dawson K, Davis P, Bowen M, Galumbeck C. Effects of cosmetics use on the  
615 physical attractiveness and body image of American college women. *J Soc Psychol*.  
616 1989;129: 349–355. doi:10.1080/00224545.1989.9712051
- 617 15. Jones AL, Kramer RSS. Facial cosmetics have little effect on attractiveness judgments  
618 compared with identity. *Perception*. 2015;44: 79–86. doi:10.1068/p7904

- 619 16. Jones AL, Kramer RSS, Ward R. Miscalibrations in judgements of attractiveness with  
620 cosmetics. *Q J Exp Psychol.* 2014;67: 2060–2068. doi:10.1080/17470218.2014.908932
- 621 17. Mulhern R, Fieldman G, Hussey T, Lévêque J-L, Pineau P. Do cosmetics enhance female  
622 Caucasian facial attractiveness? *Int J Cosmet Sci.* 2003;25: 199–205. doi:10.1046/j.1467-  
623 2494.2003.00188.x
- 624 18. Jones AL, Russell R, Ward R. Cosmetics alter biologically-based factors of beauty:  
625 Evidence from facial contrast. *Evol Psychol.* 2015;13: 210–229.  
626 doi:10.1177/147470491501300113
- 627 19. Russell R. Sex, beauty, and the relative luminance of facial features. *Perception.* 2003;32:  
628 1093–1107. doi:10.1068/p5101
- 629 20. Russell R. A sex difference in facial contrast and its exaggeration by cosmetics. *Perception.*  
630 2009;38: 1211–1219. doi:10.1068/p6331
- 631 21. Stephen ID, McKeegan AM. Lip colour affects perceived sex typicality and attractiveness  
632 of human faces. *Perception.* 2010;39: 1104–1110. doi:10.1068/p6730
- 633 22. Fink B, Grammer K, Matts P. Visible skin color distribution plays a role in the perception  
634 of age, attractiveness, and health in female faces. *Evol Hum Behav.* 2006;27: 433–442.  
635 doi:10.1016/j.evolhumbehav.2006.08.007
- 636 23. Samson N, Fink B, Matts PJ. Visible skin condition and perception of human facial  
637 appearance. *Int J Cosmet Sci.* 2010;32: 167–184. doi:10.1111/j.1468-2494.2009.00535.x
- 638 24. Jones AL, Porcheron A, Sweda JR, Morizot F, Russell R. Coloration in different areas of  
639 facial skin is a cue to health: The role of cheek redness and periorbital luminance in health  
640 perception. *Body Image.* 2016;17: 57–66. doi:10.1016/j.bodyim.2016.02.001
- 641 25. Porcheron A, Mauger E, Russell R. Aspects of facial contrast decrease with age and are  
642 cues for age perception. de Beeck HPO, editor. *PLoS ONE.* 2013;8: e57985.  
643 doi:10.1371/journal.pone.0057985
- 644 26. Kościński K. Hand attractiveness--its determinants and associations with facial  
645 attractiveness. *Behav Ecol.* 2012;23: 334–342. doi:10.1093/beheco/arr190
- 646 27. Rhodes G. The evolutionary psychology of facial beauty. *Annu Rev Psychol.* 2006;57:  
647 199–226. doi:10.1146/annurev.psych.57.102904.190208
- 648 28. Penton-Voak I. In retreat from nature? Successes and concerns in Darwinian approaches to  
649 facial attractiveness. *J Evol Psychol.* 2011;9: 173–193. doi:10.1556/JEP.9.2011.2.5
- 650 29. Morrison ER, Morris PH, Bard KA. The stability of facial attractiveness: Is it what you've  
651 got or what you do with it? *J Nonverbal Behav.* 2013;37: 59–67. doi:10.1007/s10919-013-  
652 0145-1

- 653 30. Cox CL, Glick WH. Resume evaluations and cosmetics use: When more is not better. *Sex*  
654 *Roles*. 1986;14: 51–58.
- 655 31. Huguet P, Croizet J-C, Richetin J. Is “what has been cared for” necessarily good? Further  
656 evidence for the negative impact of cosmetics use on impression formation. *J Appl Soc*  
657 *Psychol*. 2004;34: 1752–1771. doi:10.1111/j.1559-1816.2004.tb02796.x
- 658 32. Tiddeman B, Burt DM, Perrett D. Prototyping and transforming facial textures for  
659 perception research. *IEEE Comput Graph Appl*. 2001;21: 42–50. doi:10.1109/38.946630
- 660 33. Peirce JW. PsychoPy - psychophysics software in Python. *J Neurosci Methods*. 2007;162:  
661 8–13. doi:10.1016/j.jneumeth.2006.11.017
- 662 34. Coetzee V, Greeff JM, Stephen ID, Perrett DI. Cross-cultural agreement in facial  
663 attractiveness preferences: The role of ethnicity and gender. McCormick C, editor. *PLoS*  
664 *ONE*. 2014;9: e99629. doi:10.1371/journal.pone.0099629
- 665 35. Howell DC. *Statistical methods for psychology*. 4th ed. Belmont, CA: Wadsworth; 1997.
- 666 36. Cross JF, Cross J. Age, sex, race, and the perception of facial beauty. *Dev Psychol*. 1971;5:  
667 433.
- 668 37. Wickham LH, Morris PE. Attractiveness, distinctiveness, and recognition of faces:  
669 Attractive faces can be typical or distinctive but are not better recognized. *Am J Psychol*.  
670 2003;116: 455. doi:10.2307/1423503
- 671 38. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical  
672 primer for t-tests and ANOVAs. *Front Psychol*. 2013;4: 863. doi:10.3389/fpsyg.2013.00863
- 673 39. Guéguen N, Jacob C. Lipstick and tipping behavior: When red lipstick enhance waitresses  
674 tips. *Int J Hosp Manag*. 2012;31: 1333–1335. doi:10.1016/j.ijhm.2012.03.012
- 675 40. Jacob C, Guéguen N, Boulbry G, Ardiccioni R. Waitresses’ facial cosmetics and tipping: A  
676 field experiment. *Int J Hosp Manag*. 2010;29: 188–190. doi:10.1016/j.ijhm.2009.04.003
- 677 41. Holme SA, Beattie PE, Fleming CJ. Cosmetic camouflage advice improves quality of life.  
678 *Br J Dermatol*. 2002;147: 946–949.
- 679 42. Jenkins R, White D, Van Montfort X, Burton AM. Variability in photos of the same face.  
680 *Cognition*. 2011;121: 313–323.
- 681 43. Samson N, Fink B, Matts PJ. Visible skin condition and perception of human facial  
682 appearance. *Int J Cosmet Sci*. 2010;32: 167–184. doi:10.1111/j.1468-2494.2009.00535.x
- 683 44. Said CP, Todorov A. A statistical model of facial attractiveness. *Psychol Sci*. 2011;22:  
684 1183–1190. doi:10.1177/0956797611419169

685 45. Tagai K, Ohtaka H, Nittono H. Faces with light makeup are better recognized than faces  
686 with with heavy makeup. *Front Psychol.* 2016;7. doi:10.3389/fpsyg.2016.00226

687 46. Fritz CO, Morris PE, Richler JJ. Effect size estimates: current use, calculations, and  
688 interpretation. *J Exp Psychol Gen.* 2012;141: 2–18. doi:10.1037/a0024338

689

## 690 **Supporting Information**

691 **S1 Dataset. Data from Study 1.** Each participant rated all 33 YouTube models, but each model  
692 appeared in a randomly selected cosmetics condition. All conditions are stated in the data. We  
693 averaged across participants for each image, building a score for each identity under both  
694 cosmetics conditions.

695

696 **S2 Dataset. Data from Study 2.** Each participant rated all 45 supermodels, but each model  
697 appeared in a randomly selected cosmetics condition. All conditions are stated in the data. We  
698 averaged across participants for each image, building a score for each identity under both  
699 cosmetics conditions.

700

701 **S3 Dataset. Data from the quantity raters in both studies.** Sheet 1 contains the quantity from  
702 Study 1, and Sheet 2 contains the quantity data for Study 2. Participants compared each model  
703 without and with cosmetics, indicating how much cosmetics the faces were wearing.