Miscalibrations in judgements of attractiveness with cosmetics

Alex L. Jones\textsuperscript{1}, Robin S. S. Kramer\textsuperscript{2}, Robert Ward\textsuperscript{1}

\textsuperscript{1}School of Psychology, Bangor University, UK.
\textsuperscript{2}School of Psychology, University of Aberdeen, UK.

Acknowledgement

This research was supported by Knowledge Economy Skills Scholarships (KESS). KESS is a pan-Wales higher level skills initiative led by Bangor University on behalf of the HE sector in Wales. It is part funded by the Welsh Government (ESF) convergence program for West Wales and the Valleys. The funding body had no involvement in the present paper.

The authors would like to thank Bernard Tiddeman for his advice, and the two anonymous reviewers for their comments.

Correspondence concerning this article should be addressed to Alex L. Jones, School of Psychology, Bangor University, Bangor LL57 2AS, Gwynedd, United Kingdom.

E-mail: alexjonesphd@gmail.com
Telephone: +44 (0)1248 382791
Abstract

Women use cosmetics to enhance their attractiveness. How successful they are in doing so remains unknown - how do men and women respond to cosmetics use in terms of attractiveness? There are a variety of miscalibrations where attractiveness is concerned - often, what one sex thinks the opposite sex finds attractive is incorrect. Here, we investigated observer perceptions about attractiveness and cosmetics, as well as their understanding of what others would find attractive.

We used computer graphic techniques to allow observers to vary the amount of cosmetics applied to a series of female faces. We asked observers to optimise attractiveness for themselves, for what they thought women in general would prefer, and what they thought men in general would prefer. We found that men and women agree on the amount of cosmetics they find attractive, but overestimate the preferences of women, and when considering the preferences of men, overestimate even more. We also find that models’ self-applied cosmetics are far in excess of individual preferences. These findings suggest attractiveness perceptions with cosmetics are a form of pluralistic ignorance, whereby women tailor their cosmetics use to an inaccurate perception of others’ preferences. These findings also highlight further miscalibrations of attractiveness ideals.

*Keywords:* faces; attractiveness; cosmetics; social cognition; pluralistic ignorance
Self-adornment is an important social behaviour by any standard. Throughout the animal world, the more ornamented sex is typically the one investing less in offspring (Trivers, 1972). Though ornamentation is more balanced in humans, given the more comparable amounts of parental investment, in modern society self-adornment is more prevalent in women. Perhaps the best example of this is the use of facial cosmetics. By some estimates, the value of the global broad cosmetics industry was around €136 billion in 2006 (Rossi, Prlic, & Hoffman, 2007). Here we examine a basic question relating to the use of cosmetics and attractiveness. How do cosmetics affect men’s and women’s perceptions of attractiveness? Research has identified several important and consistent predictors of attractiveness. For example, facial symmetry (Grammer & Thornhill, 1994), averageness (Alley & Cunningham, 1991) and sexual dimorphism (Johnston & Franklin, 1993) greatly influence perceived attractiveness. Skin condition is also important - homogeneously textured skin is a strong signal of health (Samson, Fink & Matts, 2010), while reddened lips may be a cue to healthy circulation (Stephen & McKeegan, 2010). Females also have lighter skin than males (Jablonski & Chaplin, 2000) and males find females with lighter skin more attractive (Russell, 2003). Additionally, this difference in luminance is consistent across ethnicities (Russell, 2009). Furthermore, The contrast between skin and facial features (e.g. eyes and mouth) is greater in female faces, with females possessing darker features and lighter skin, which influences perceptions of femininity (Russell, 2009).

It is no surprise that cosmetics act on these properties. Cosmetics increase facial contrast (Russell, 2009), exaggerating sex typical differences in faces. Cosmetics also homogenise skin texture (Samson et al., 2010), and may alter colour properties linked to healthy skin (Stephen, Coetzee, & Perrett, 2011) such as yellowness and lightness (Stephen, Law-Smith, Stirratt, & Perrett, 2009). Some cosmetic products seem specifically tailored to modifying these perceptions. For example, blushers typically add redness to the face, a colouration tied to circulatory health and a healthy appearance (Stephen, Coetzee, Law-Smith, & Perrett, 2009). Indeed, an application of cosmetics has positive effects on health perceptions (Mulhern et al., 2003). Concordantly, women
with greater facial asymmetries spend longer applying cosmetics (Korichi, Pell de Queral, & Gazano, 2011). It is therefore unsurprising that cosmetic practices have been shown to consistently increase attractiveness (Etcoff et al, 2011; Mulhern et al., 2003). Cosmetics, then, allow women to increase their facial attractiveness by modifying attributes that influence attractiveness in unadorned faces.

Both popular media and scholarly sources (Buss, 1988; Singh, 2004) suggest that women use cosmetics as a mate attraction tactic, in a similar way to clothing (Durante, Li, & Haselton, 2008). One possibility is that women apply cosmetics specifically to appeal to men. If so, then we might expect women to apply cosmetics so as to be maximally attractive to men, and for men to respond more favourably than women to cosmetics use. An alternative is that women use cosmetics not to attract men, but to compete with women. For example, women are more sensitive to the opinions of other women than of men when it comes to their own attractiveness (Graziano et al., 1993), and the primary avenue of competition between women is appearance (Campbell, 2004). Women perceive attractive other women as socially dominant (Dijkstra & Buunk, 2001), and are jealous when other women surpass them in attractiveness (Buss et al., 2000). It is therefore plausible that women are motivated to appear attractive to both men and women, and use cosmetics to achieve this goal.

However, there are a number of perceptions individuals hold about the beliefs, attitudes and preferences of the opposite sex that are highly inaccurate. A specific example relates to attractive body shapes. Women believe men prefer much thinner body shapes than men actually do (Fallon & Rozin, 1985), while men believe women have a preference for much more muscular body shapes than women do (Pope et al., 2000). These attractiveness miscalibrations seem to be a form of pluralistic ignorance. The ideals of attractiveness held by one sex are very different from the personal preferences of the other sex, which are surely the preferences that should be appealed to. These false perceptions are upheld through the assumption that the opposite sex really does prefer the assumed trait (Prentice & Miller, 1993), assumptions that may also be fuelled by media sources
Cosmetics are an easy way of modifying appearance compared with diets and exercise, and act upon the most salient signal of our attractiveness - the face (Currie & Little, 2009). Are there similar miscalibrations of perceived attractiveness with cosmetics?

We seek to address two issues relating to cosmetics use and attractiveness. We first examine how men and women respond to cosmetics, and in doing so attempt to uncover who, if anyone, women are tailoring their cosmetics to. Secondly, we examine whether there are errors and miscalibrations in this tailoring. To address these issues, we photographed a sample of young women before and after they applied facial cosmetics for “a night out”. By generating a smooth sequence of images between these photographs, we allowed participants to select the level of cosmetics they found maximally attractive for each face. We asked participants to judge not only their personal preferences of attractiveness, but also what they thought others might prefer. Though there are several factors influencing attractiveness, there is still considerable inter-individual variance in attractiveness judgements (Kościński, 2008). This may be compounded by cosmetics use, which exhibits cultural variation (Russell, 2010). We therefore wanted to contextualise the question by asking observers to assess attractiveness from different perspectives: what they personally preferred, and what they thought women in general and men in general preferred. By providing a perspective for the attractiveness judgement, we first intended to reduce the variance of responses that might arise if people answered the question according to different criteria. Most importantly, we can also examine differences between individual preferences of attractiveness and cosmetics, and what they believe others prefer - does the use of cosmetics demonstrate pluralistic ignorance?

Method

Models and stimulus generation
Forty-four women (age in years, $M = 21.18$, $SD = 1.94$) participated as models. All self-reported as being of White ethnicity, and gave full consent for their photographs to be used. All were paid £6 for their participation.

**Stimulus sequence generation.** We asked models to remove all traces of facial jewellery as well as thoroughly clean their face of all cosmetic products. Models also tied their hair back from their face. We photographed models using a Nikon D3000 SLR camera mounted on a tripod, at a distance of approximately one metre, in a room with no other sources of lighting. Models were photographed against a white background, with a Nikon SS-400 flash angled 45º towards the ceiling. We photographed each model three times, and used the clearest exposure as our final stimulus. After the initial photograph, models were provided with a range of best-selling foundations, lipsticks, mascaras and blushers, and were instructed to apply their cosmetics as though they were going on a “night out”. They were then photographed again to capture their appearance with cosmetics. Between shots, camera settings were kept constant, including lens aperture (F5.3), exposure time (1/60 seconds) and ISO speed rating (200). All photographs were subsequently rotated so the pupils lay along the same transverse plane, and were cropped to just above the hairline, to below the chin, and to just outside the widest point of the face.

A series of 160 landmarks were added to each model using JPsychomorph (Tiddeman, Burt, & Perrett, 2001). For each model, we generated a sequence of images that moved from their natural appearance (defined as 0%) to their night out look with cosmetics (defined as 100%). To avoid any floor or ceiling effects, sequences started at -50%, which exaggerated their appearance without cosmetics, and extended to 150%, which exaggerated their appearance with cosmetics. The transform sequence can be simply thought of as taking the difference between the 100% image and the 0% image, and multiplying this difference by the desired transform level (e.g., 50%). Finally, these values are added to the original image to complete the transform.

JPsychomorph uses a wavelet Markov Random Field (MRF) method for interpolating realistic, fine grain textures (Tiddeman, Stirratt, & Perrett, 2005). Using this method, high
resolution information such as colour and texture in a transformed image is calculated by assuming the new pixel distribution is dependent on the values in the local neighbourhood of landmark points. With this method, variations in intensity change linearly across any given sequence, with the result being that blurring at intermediate steps is greatly reduced. This ensures highly realistic images within the sequences for each model. Additionally, a Procrustes fit is performed before sequence generation, which aligns the images as precisely as possible, resulting in a cleaner transform and further reducing transformation artefacts.

For each model, the final sequence contained 21 images, ranging from -50% to 150% in increments of 10%. Models appeared to increase the amount of cosmetics worn as participants moved through the sequence. An example of the difference between the two cosmetics conditions with average faces appears in Figure 1, and an example set of photographs in Figure 2 demonstrates a 50% shifted image and the two anchor points, illustrating the realism of the sequences.

Participants

Forty-four observers (age $M = 20.06$, $SD = 1.97$, 22 males) from Bangor University participated for course credit.

Procedure

The presentation order of models was randomised, and each trial began with a random image from their sequence. Underneath the face was a white circle, with a red bar over a random point of the circumference. The bar could be moved around the circle using left and right arrow keys. Each position of the bar corresponded to an image in the model’s image sequence, and movement of the bar altered the image. One full cycle caused the image to move smoothly through the sequence completely and back again (-50% to 150% to -50%). For each trial, the starting position of the bar and the image corresponding to its position were randomised.

Participants assessed all the models in three separate blocks, adopting a different perspective in each block: Participants were asked to optimise the attractiveness of the faces for themselves
(Self Perspective); for what they thought women in general would prefer (To Women Perspective); and for what they thought men in general would prefer (To Men Perspective). These definitions were placed at the top of the screen throughout the respective blocks. On each trial, participants moved through the sequence of each model with the arrow keys, and pressed the spacebar once they felt they had reached optimal attractiveness for the current perspective. Responses were untimed and unspeeded. The percentage of cosmetics use corresponding to the selected image was then recorded as the participant’s response, although this value was not seen by participants. The block order was counterbalanced across participants.

Results

For each model, we calculated the average selection for optimal attractiveness for all three perspectives, for both male and female observers. These results are shown in Figure 3. A 2 (Observer Sex: Female, Male) x 3 (Perspective: Self, To Women, To Men) ANOVA revealed a significant main effect of Observer Sex, $F(1, 43) = 7.54, p = .009, \eta^2_p = .15$, such that women selected slightly higher amounts of cosmetics than did men. While significant, this difference of approximately 3% was small compared to the significant main effect of Perspective, $F(2, 86) = 149.41, p < .0001, \eta^2_p = .77$. Post-hoc comparisons between perspectives highlight ways in which observers of both sexes reported inaccurate views about how others respond to cosmetics. As shown by the comparison of the ‘To Men’ and ‘To Women’ perspective in Figure 3, observers judged that men would prefer significantly more cosmetics than would women, $t(43) = 8.55, p < .0001, r^2 = .63$. However, this judgement was incorrect. Looking at the results for the ‘Self’ perspective in Figure 3, we find that if anything, women trended towards preferring slightly more cosmetics than men did, $t(43) = 1.89, p = .06, r^2 = .07$.

The large gap between the ‘Self’ and other perspectives highlights a second, related error of judgement. Observers of both sex judged that other people would prefer higher levels of cosmetics use than they themselves did, reflected by the difference between the ‘Self’ and ‘To Women’
perspectives, \( t(43) = 10.33, p < .0001, r^2 = .71 \), and the ‘Self’ and ‘To Men’ perspectives, \( t(43) = 15.22, p < .0001, r^2 = .84 \).

Finally, we see a third intriguing error of judgement, this time by the models rather than the observers. Models used more makeup than observers found attractive, evident from the fact that the personal preferences of observers (Self Perspective) was significantly different from 100% (i.e. how much the models actually applied) for both sexes, \( t > 445.49, p < .0001 \).

We checked whether model attractiveness and the application of cosmetics might have been entangled in some way. For example, observers might have preferred the way in which more attractive models used cosmetics, or perhaps more attractive individuals benefit differently from using cosmetics than less attractive individuals. To this end, we repeated our main analysis as a 2 (Observer Sex) x 3 (Perspective) ANCOVA, with the mean attractiveness of the models without cosmetics, rated by 27 participants (age \( M = 21, SD = 3.35, 14 \) females) on a seven-point scale, as a covariate. The pattern of results was unaffected, although effect sizes were smaller, with significant main effects of Observer Sex, \( F(1, 42) = 5.73, p = .02, \eta^2_p = .12 \); Perspective, \( F(2, 86) = 4.42, p = .01, \eta^2_p = .09 \); and no interaction, \( F(1, 43) = 0.23, p = .79, \eta^2_p = .00 \). We also checked whether there were any correlations between rated attractiveness and the amount of cosmetics applied in the different blocks, all \( r\)s between -.14 and .17, all \( p\)'s > .27, indicating little influence of attractiveness on personal or perceived preferences for cosmetics.

We considered whether the consistent preference for faces wearing less than 100% cosmetics was due to artefacts from the sequence generation process. Our intermediate faces reflect an average of two faces (e.g., the 50% image is a composite of the 0% image and the 100% image). With traditional averaging methods, composites from multiple faces produce smoother skin texture and a more attractive appearance (Little & Hancock, 2002). Although the use of wavelet MRF texture transform is intended to reduce or remove this possibility (Tiddeman, Stirratt, & Perrett, 2005), we examined the issue empirically. We cropped a 114 x 114 patch of skin from the right cheek of each model in her 0% no cosmetics image, her 50% sequence image, and her self-applied
100% cosmetics image. Each of these sets was rated for attractiveness (on a scale of 1-7) by a
different set of 12 participants. If the preference for intermediate faces were due simply to more
attractive skin textures due to averaging, then we would expect the skin of the 50% transform to be
more attractive than both the 0% and 100% natural images. However, although both the 0% and
50% patches were rated as more attractive than the 100% patch ($M = 3.51, SE = .12$), the 50% patch
was actually non-significantly less attractive ($M = 3.89, SE = .12$) than the 0% patch ($M = 4.06, SE
= .11$), $p = .09$. It therefore seems unlikely that the consistent preference for intermediate
composites reflects a skin-smoothing artefact of the averaging process.

5.4 Discussion

We found a variety of miscalibrations and errors in people’s judgements about cosmetics
use. Observers of both sexes judged, incorrectly, that men would prefer significantly more
cosmetics than would women. Instead we found that men and women had similar preferences for
cosmetics use. Observers of both sexes also significantly overestimated how much cosmetics other
people would prefer. These misjudgements of other people’s reactions and preferences to cosmetics
also seem to have been carried by the models themselves: The amount of cosmetics the models
actually wore was significantly higher than observers’ personal preferences, and even higher than
observers’ overinflated estimates of what others would prefer. Taken together, these results suggest
that women are likely wearing cosmetics to appeal to the mistaken preferences of others. These
mistaken preferences seem more tied to the perceived expectancies of men, and, to a lesser degree,
of women.

Although we saw little difference in personal preferences, our observers reported that men
would prefer relatively higher levels of cosmetics than would women. This is consistent with the
belief that men prefer an exaggerated sex typical appearance - in this case, a ‘supernormal’ face. In
animal studies, supernormal attributes such as experimentally lengthened tails can increase mating
success (Winquist & Lemon, 1994). Cosmetics likely function in humans in a similar fashion
(Etcoff et al., 2011), and this may explain the discrepancy between men’s actual preferences and
their estimated preferences. Non-human animals have been shown to exaggerate colour cues or other visual features to increase their mating success, though these behaviours can cause the opposite sex to decipher what is a real or deceptive signal of mate value (Trivers, 1985). While the general consensus is that more cosmetics are more attractive, at the individual level men may realise these alterations are essentially deceptive, since they modify and mask important cues to mate value (Samson et al., 2010; Stephen & McKeegan, 2010).

Our sample of models self applied their cosmetics, resulting in an ecologically valid use. A surprising proportion of studies looking at cosmetics employ the use of a professional makeup artist to apply cosmetics (Etcoff et al., 2011). In reality, the application of cosmetics in these studies reflects the preferences and ideals of one individual and tells us little about how individuals alter their own appearance. Our participants were instructed to apply cosmetics for a “night out”, and so may have incorporated ideas of low lighting, a particular dress style, etc. However, as demonstrated in Figure 1, there was nothing especially unusual in the way the sample of models applied cosmetics. The style appears to conform closely to what Russell (2010) calls the ‘received style’, with cosmetics being applied to exaggerate areas of contrast around the eyes and mouth. Additionally, though there is the potential for the over-application of cosmetics to change with another style, it does not explain the results where observers over-estimate what others prefer, nor that observers believe that men prefer more cosmetics than women. However, it is possible that an application of cosmetics made models feel more confident, which could have led to minute but perceptible changes in expression or posture. Other studies have demonstrated that individuals appear more confident after an application of cosmetics (Mulhern et al., 2003), and that individuals report greater satisfaction with their social interactions whilst wearing makeup (Miller & Cox, 1982).

A possible limitation concerns the fact that the images in our sequences are derived from a pair of images for each model, and as such, are not exactly like a real face. Were this the case, we could not be sure the actual amount of cosmetics worn by the models is beyond optimal levels, but
is instead due to preferences for an ‘averaged’ skin texture which may be present throughout the sequence. A simple comparison of faces in the sequence and natural images would be confounded by the presence of cosmetics. However, examination of the attractiveness of skin patches from a face in the sequence and the two natural photographs demonstrated the 100% cosmetics skin was less attractive than both the 50% transform and the natural appearance. These results indicate that the amount of cosmetics worn by models is indeed beyond optimal levels of attractiveness, and is particularly valid as small skin samples are predictive of overall face attractiveness (Jones, Little, Burt, & Perrett, 2004).

It seems that the models in our sample were using cosmetics to appeal to mistaken beliefs about what men, and to a lesser degree other women, would prefer. Such mistaken beliefs are also common when women consider what body shape men find attractive (Bergstrom, Neighbors & Lewis, 2004). Media portrayals of idealised, attractive women tend to possess supernormal properties (Barrett, 2010) such as flawless skin, low waist to hip ratios (Singh, 2004) and youthful features. Many of these portrayals cause great dissatisfaction in women regarding their own levels of facial and bodily attractiveness (Li et al., 2010), and can be influential from an early age (Hargreaves & Tiggermann, 2004). Further associations between beauty and positive life outcomes are propagated by media aimed at adolescent (Clay et al., 2005), and older individuals (Becker et al., 2005). These errors can be viewed under the broad umbrella of pluralistic ignorance, of which our results are another example. Our observers believed others, especially men, preferred higher amounts of cosmetics than the observers personally did. Additionally, our models seemed to be perpetuating the mistaken belief that more cosmetics are more attractive, by wearing excessive amounts. Our female observers indicated the amount of cosmetics for optimal attractiveness was lower than the amount worn by the models, who were a sample of similarly aged peers. This suggests that while there is a sense of what is optimally attractive, it may be overlooked in order to conform to stereotyped ideals, and mistaken notions of what others will find attractive.
References


MISCALIBRATIONS AND COSMETICS


