The impact of skin tone on perceived facial beauty: A two-culture study

Germano Vera Cruz¹ and Etienne Mullet²

¹Eduardo Mondlane University, Maputo, Mozambique
²Institute of Advanced Studies (EPHE), Paris, France

The impacts of skin tone and shape of face on judgment of beauty regarding female faces were assessed and compared. Sixty adolescents and 60 adults living in Maputo, Mozambique or in Toulouse, France were presented with a set of faces, one at a time, and instructed to rate their attractiveness (in terms of beauty) along a continuous scale. Skin tone contributed to judgments of beauty but its contribution was much weaker than the contribution of the shape of the faces: It explained about 5% of the variance as compared with 85% for shape of faces. The Shape x Tone interaction was not significant; that is, both factors simply added their effect during the judgment process. The Country x Tone interaction was not significant: The impact of skin tone on judgment was similar among Mozambicans and among the French.

Since the 80s, skin bleaching has become a public health concern in many African countries. Skin bleaching involves the application of various cosmetic products (e.g., creams, soap, and lotions) that contain potentially dangerous chemical agents. The aim is to obtain a lightened skin by reducing its melanin content.

Surveys conducted from 1986 to present days show that skin bleaching is practiced, (a) at least occasionally, by two out of three inhabitants of Brazzaville, Congo, (b) by one out of four women living in Bamako, Mali, (c) by one out of two women living in Dakar, Senegal (and one out of four men), and (d) by more than three out of four women and men living in Lagos, Nigeria (e.g., Adebajo, 2002). This phenomenon is

¹ Address correspondence to Dr. Vera Cruz, e-mail: germano.veracruz@gmail.com
not limited to West Africa. A study conducted twenty five years ago showed that seven out of ten patients in two hospitals in Pretoria, South Africa practiced skin bleaching (Hardwick, Van Gelder, Van der Merwe, & Van der Merwe, 1989).

The consequences of skin bleaching can be severe. Skin bleaching has been shown to be associated with (a) premature ageing of skin and with subsequent skin cancers, (b) higher levels of hypertension and diabetes, and (c) higher levels of depression and identity disorders. Pigmentation damages are frequently diagnosed among users. In a sample of West African women using bleaching agents, seven out of ten have been diagnosed with at least one dermatological problem. Among women living in Senegal, the long-term use of skin bleaching agents has been shown to be responsible for a high rate of cutaneous adverse effects. A review of these effects can be found in Kpanake, Munoz Sastre, Sorum, & Mullet (2008). Campaigns of information have been launched in various countries (e.g., Food and Agriculture Organisation, 2006; Ouattara, 2006), but their effect on people’s health behaviors has been rather negligible. People keep being largely unaware of the risks associated with bleaching (Kpanake et al., 2008).

Like many other public health problems (e.g., obesity), skin bleaching is a societal problem having deep psychological roots. Kpanake, Munoz Sastre, and Mullet (2010) examined the motives underlying the practice of regular skin bleaching among Western African (Togolese) adults. Four basic, largely independent motives appeared as clearly dominant. Participants practiced skin bleaching on a regular basis mainly (a) to appear important, (b) to look attractive, (c) because they enjoyed their light colored skin, and (d) because skin bleaching was fashionable. They did not practice skin bleaching as a demonstration of opposition to African culture or to their relatives or as a demonstration of compliance with others’ wishes. They also sometimes practiced skin bleaching as a means of securing a job.

The present study was aimed at examining the extent to which one of the concerns that motivates skin bleaching practices is really grounded. More precisely, the present study was aimed at assessing the impact of skin tone on attractiveness judgment regarding female faces. Early studies have shown that pro-light-skinned/anti-dark-skinned bias is usually found among American preschool children of European and African ancestry (Williams & Morland, 1976). At first, this bias was given a racial interpretation: It was considered as reflecting children’s learning experiences in a multiracial society where (a) prejudice against dark-skinned persons is often
encountered, and (b) dark-skinned persons are generally poorer and less educated than other persons.

The racial interpretation has subsequently been challenged. Preschool children in countries with few dark-skinned persons demonstrated the same bias, which, as a result had to be attributed to other sources, possibly to a general pro-light/anti-dark bias that is related to preferences for colors in general (and not preferences for skin features). A strong, positive relationship between the strength of this bias and the strength of preference for light colors was found among children in many countries (Neto & Williams, 1997). Also, it was shown that children of solely black parentage, or of solely white parentage or of mixed black and white parentage demonstrated the same bias, practically to the same extent (Neto & Paiva, 1998). Hill (2002), using data from the National Survey of Black Americans (NSBA), examined the relationship between Black interviewers’ perception of Black respondents’ physical attractiveness, and Black interviewers’ assessment of respondent’s skin tone (in terms of darkness). Physical attractiveness and skin tone were shown to be positively correlated, and the strength of the association was stronger regarding female respondents than male respondents. It can be concluded that, people in general, whether African or European, child, adolescent, or adult, prefer White over Black, whether it is displayed on cardboard (Neto & Williams, 1997), dolls (see, Powell-Hopson & Hopson, 1988), cartoons (Jordan & Fernandez-Reif, 2009), or human skin (Hill, 2002).

Findings from these studies appear to support the practice of skin bleaching. If light colors are preferred to dark colors, then lighting one’s face may certainly have positive consequences in terms of personal attractiveness. There is, however, a methodological problem that is associated with some of these studies, which may, at least partly, invalidate this inference. Participants were usually presented pairs of stimuli that differed in tone (faces, dolls or cardboards), and instructed to choose the one they liked most. Such a methodology has a basic virtue, which is simplicity. It does not allow, however, precisely assessing the impact of tone on preference. An example can illustrate this point. Suppose a group of people is instructed to choose between $999 and $1,001. Everybody will choose the second option but this observation will tell nothing about the size of the difference between offers.

In the present study, participants were not presented with pairs of faces and instructed to choose one of them. They were presented with a set of faces, one at a time, and instructed to rate their attractiveness (in terms of beauty) along a continuous scale. In other words, to a dichotomous
methodology was substituted a continuous methodology: Information Integration Theory (Anderson, 2008). By varying skin tone and shape of faces, it was thus possible to precisely assess the impact of each factor (in terms of part of explained variance), and, by the same token, to precisely assess whether the two factors independently contributed to attractiveness or whether they interacted. In addition, the present study was cross-cultural in character. Two very different samples were considered, a sample of participants living in Southern Africa (Mozambique), and a sample of participants living in Western Europe (France).

The first hypothesis was that skin tone would significantly contribute to attractiveness of faces. The second hypothesis was that the extent to which skin tone contributes to attractiveness of faces among African and among Europeans would be similar. The third hypothesis, based on the work by Neto and colleagues (e.g., Neto & Williams, 1997) was that a majority of participants would prefer light faces to dark faces. The fourth hypothesis, also based on the same sources, was that the extent to which light faces are preferred to dark faces among African and among Europeans would be similar.

Two research questions were also considered. The first one was about the relative impact of skin tone and shape of faces on attractiveness judgments: Which of the two factors contribute most? The second one was about the way in which skin tone and shape of faces combine their effect in participant’s judgments: Do they interact, and, if so, in what way?

**METHOD**

**Participants.** The participants were 60 adolescents (aged 15-18, $M=16.28$, $SD=1.71$) and 60 adults (aged 19-60, $M=27.89$, $SD=11.62$) living in Maputo, Mozambique or in Toulouse, France. The adolescents were recruited in secondary schools. The adults were recruited in the streets of the two cities. Sixty-nine participants were females and 51 were males, and the female/male ratio was similar in both countries. There were 30 adolescents and 30 adults from each country.

**Material.** The material was composed of nine pictures showing a female face. These nine pictures were obtained by crossing three levels of physical beauty (very attractive, mildly attractive, and not very attractive) and three levels of skin tone (Black, Métis, and White). The three faces that expressed the three levels of physical beauty were selected from a pilot study in which 25 female faces (all Métis from White and Black ancestors)
were presented to 40 Mozambican adults who were instructed to rate each face on a 0-20 beauty scale. These females had formally consented that their picture is used in a scientific study under the condition of strict anonymity. In all pictures facial expression was kept neutral, and females did not wear any make-up of facial adornment.

The most attractive face was rated 14.00 (SD=3.08), the less attractive was rated 3.67 (SD=2.67), and the face that was selected for representing the intermediate level was rated 9.74 (SD=3.20). The three faces were then either darkened until obtaining the skin tone that is common among Black people living in Mozambique, or lighted until obtaining the skin tone that is common among White people living in Western Europe or let untouched.

The material was also composed of 22 additional faces, the ones that had not been selected from the set of 25 used in the pilot study. Seven of these faces were darkened, seven were lighted and the other ones were left untouched.

**Procedure.** Three different sets of faces were created. The first set was composed of three faces taken from the set of nine (e.g., “Black” combined with “Very attractive”, “Métis” combined with “Not very attractive”, and “White” combined with “Mildly attractive”) to which were added the 22 faces forming the background. The second set was composed of three other faces taken from the set of nine, and of the 22 additional faces. The third set was composed of three remaining faces from the set of nine to which were added the 22 faces. In each set, the order of presentation of the faces was made random except that the first experimental face was presented in fifth position.

Participants were instructed to rate facial beauty in three sessions. On Day 1, they were presented one of the three sets of faces, taken randomly. On Day 15, they were presented with another set, also taken randomly. On Day 30, they were presented with the remaining set. No participant realized that three of the faces with which they were presented in one or another of the three sessions were either Black, or Métis or White, depending of the session, while the color of the other faces did not change.

As the issue of skin tone can rise ethical problems for some participants and influence their responses, they were only told that this study was about beauty judgment. That means that participants were not aware of the skin tone dimension of the study.
RESULTS

Prior reordering of factor levels

The pilot session allowed the experimenters to select three levels of attractiveness which mean ratings were clearly different (from about 4 to 14). This does not mean, however, that, for all participants, these three levels were always ordered in the same way. Individual differences, as reflected in the high standard deviations observed in the pilot study, were considerable. Also, preference for skin tone was not always the same from one participant to the other: 28 participants (10 Mozambicans and 18 French) rated the Black faces higher than the White or Métis faces, 32 participants (12 Mozambicans and 20 French) rated the Métis faces higher than the Black or White faces, and 60 participants (38 Mozambicans and 22 French) rated the White faces higher than the Black or Métis faces, \( \chi^2(2) = 8.55, p < .02 \). The relationships between preference for tone and age or gender were not significant.

One basic precaution that had to be taken before considering group data; that is, before averaging data over participants, was to make sure that such an aggregation process makes sense. When considerable inter-individual differences exist, as in the present case, the reordering of levels is needed. This will be illustrated via an example. Suppose that data observed for Participant A have been displayed in the left panel of Figure 2. In this panel, (a) the mean beauty judgments are on the y-axis, (b) the three levels of skin tone are on the x-axis, and (c) the three curves correspond to the three levels of face attractiveness. Participant A likes very much Black faces (and Face X), likes Métis faces (and Face Y) and strongly dislikes White faces (and Face Z). Suppose now that data observed for Participant B have been displayed in the center panel of Figure 2. Participant B likes very much White faces (and Face X), likes Métis faces (and Face Y) and strongly dislikes Black faces (and Face Z).

It is safe to conclude that Participants A and B differ in their preferences for skin tone, but are similar (a) in their preference for shapes, and (b) in the way the two factors produce liking: The more a skin tone is liked, and the more a shape is liked, the more the corresponding face is considered as beautiful. In the framework of Information Integration Theory, there is no problem with such an apparent contradiction. Valuation processes and integration processes are considered as independent processes. Participants can differ in the way they value (like) skin tone, and even in the way they value (like) shapes of faces but, nevertheless, use the same rule at the integration stage. Through proper reordering of levels of factors, it is possible to separate the valuation processes and the integration
process. Reordering (and, more generally re-scaling) are current practice in Information Integration studies (e.g., the program FM#1, created by Shanteau, 1984; the study by Makris & Mullet, 2003, on musical preferences, and the study by Karpowicz-Lazreg & Mullet, 2001, on the form-color Bauhaus’ problem). Coombs and Avrunin (1977) have provided a detailed rationale for rescaling procedures.

![Figure 1. Hypothetical patterns of results corresponding to Person A (left hand panel), Person B (center panel) and mean judgments (right hand panel). In each panel, (a) the judgments of beauty are on the y-axis, (b) the three levels of skin tone are on the x-axis, and (c) the three curves correspond to the three levels of shape of face.](image)

In the right panel of Figure 2, the data shown have been obtained by averaging A’s data and B’s data without reordering. The resulting pattern is totally different from the one shown in the two previously considered panels. The skin tone factor has no longer any effect on liking, by contrast with the shape factor that had kept its effect. This was clearly not what individual data had shown. In the first panel of Figure 3, the levels have been reordered for each participant before averaging, from the least preferred level to the most preferred level. The integration pattern observed tells exactly the same story as the ones shown for A and B: The more a skin tone is liked and the more the shape of a face is liked, the more the corresponding face is liked. In the present study, this is precisely what we
are interested in: Identifying the factors that impact on face preferences, and assessing their respective contribution, independently of the direction of their effects (which can, anyway, be independently assessed).

Figure 2. Hypothetical patterns of results corresponding to mean judgments of Persons A and B computed once the levels of both factors have been properly reordered. In each panel, (a) the judgments of beauty are on the y-axis, (b) the three levels of skin tone are on the x-axis, and (c) the three curves correspond to the three levels of shape of face. In the right panel the information integration rule is of an additive type. In the center panel it is of a conjunctive type. In the right panel it is of a disjunctive type.

The center and right panels of Figure 3 also show two possible integration patterns that correspond to two alternative possible ways in which both factors are combined during the judgment process (see the second research question). In the center panel, the more a tone is the preferred one, the more the shape of the face is important (the conjunctive case). For a face to be judged very attractive, it must be of the preferred tone and of the preferred shape. In the right panel, the more a tone is the preferred one, the less the shape of the face is important (the disjunctive case). In this second case, a face can be judged attractive if it is of the preferred tone or of the preferred shape.
Impact of skin tone

Figure 3. Observed patterns of results for both samples. In each panel, (a) the judgments of beauty are on the y-axis, (b) the three levels of skin tone are on the x-axis, and (c) the three curves correspond to the three levels of shape of face.

Analysis of Variance

For each participant, data were transformed through the reordering of the levels of the two factors. An analysis of variance was conducted on the reordered data, with a Gender x Age x Country x Shape x Tone, 2 x 2 x 2 x 3 x 3 design. As the Gender and Age factors were neither significant nor involved in any significant interaction, a second, simpler ANOVA was conducted with a Country x Shape x Tone, 2 x 3 x 3 design. The results of this ANOVA are shown in Table 1.

Figure 3 shows the main findings. Curves are slightly ascending; that is, the effect of skin tone was not very strong. The difference between the highest and lowest observed mean ratings (11.49 – 9.50 = 1.99) was about 2 points – about one-tenth of the length of the rating scale. In contrast, curves are clearly separated; that is, the effect of face shape was apparently strong. The difference between the highest and lowest observed mean ratings (14.64 – 6.25 = 8.39) was about 8 points; that is, about two-fifth of the length of the rating scale.

The Country x Shape interaction was significant. As illustrated in the clearly unequal separation of the curves from one panel to the other, the effect of shape was stronger among Mozambican participants (15.83 – 4.82
= 11.01) than among French participants (13.44 – 7.69 = 5.75). The Country x Tone interaction was not significant. The effect of tone was similar among Mozambican participants (11.65 – 9.44 = 2.21) and among French participants (11.33 – 9.56 = 1.77).

**Table 1. Results of the ANOVA.**

<table>
<thead>
<tr>
<th>Factor</th>
<th>df</th>
<th>Effect</th>
<th>df</th>
<th>Error</th>
<th>F</th>
<th>p</th>
<th>η²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Country</td>
<td>1</td>
<td>2.13</td>
<td>118</td>
<td>31.66</td>
<td>0.07</td>
<td>ns</td>
<td>.00</td>
</tr>
<tr>
<td>Shape</td>
<td>2</td>
<td>6327.73</td>
<td>236</td>
<td>9.54</td>
<td>663.30</td>
<td>.001</td>
<td>.85</td>
</tr>
<tr>
<td>Tone</td>
<td>2</td>
<td>356.03</td>
<td>236</td>
<td>0.83</td>
<td>430.47</td>
<td>.001</td>
<td>.78</td>
</tr>
<tr>
<td>Country x Shape</td>
<td>2</td>
<td>651.47</td>
<td>236</td>
<td>9.54</td>
<td>68.29</td>
<td>.001</td>
<td>.37</td>
</tr>
<tr>
<td>Country x Tone</td>
<td>2</td>
<td>4.26</td>
<td>236</td>
<td>0.83</td>
<td>5.15</td>
<td>ns</td>
<td>.04</td>
</tr>
<tr>
<td>Shape x Tone</td>
<td>4</td>
<td>1.27</td>
<td>472</td>
<td>0.98</td>
<td>1.30</td>
<td>ns</td>
<td>.01</td>
</tr>
<tr>
<td>Country x Shape x Tone</td>
<td>4</td>
<td>0.39</td>
<td>472</td>
<td>0.98</td>
<td>0.40</td>
<td>ns</td>
<td>.00</td>
</tr>
</tbody>
</table>

**DISCUSSION**

Using Information Integration Theory (Anderson, 2008) as a methodological framework, the present study assessed the impact of skin tone on attractiveness judgment regarding female faces among African and among European participants. The first hypothesis was that skin tone would significantly contribute to attractiveness of faces. Results fully supported the hypothesis. This finding was consistent with previous findings on color preferences (e.g., Hill, 2002; Neto & Williams, 1997). The absence of effect of age and gender was consistent with previous findings suggesting that people's views of facial beauty are highly similar regardless of respondents’ demographic characteristics (Perrett, May, & Yoshikawa, 1994).

The second hypothesis was that the extent to which skin tone contributes to attractiveness of faces among African and among Europeans would be similar. Results support the hypothesis. The Country x Tone
interaction was not significant. Close examination of the means shows that the impact of skin tone on judgment was slightly stronger among Mozambicans than among the French, a difference that is consistent with the findings by Harvey, LaBeach, Pridgen, & Gocial (2005) showing that the importance of skin tone can be slightly moderated by racial context. The third hypothesis was that a majority of participants would prefer light faces to dark faces. Results fully supported the hypothesis. Half of the participants preferred the White faces, and the remaining participants preferred either the Black face or the Métis face. This result was consistent with previous findings on color preferences (e.g., Neto & Paiva, 1998). The fourth hypothesis was that the extent to which light faces are preferred to dark faces among African and among Europeans would be similar. Results did not support the hypothesis. Only about one-third of the French preferred the White faces versus about two-thirds of the Mozambicans.

The first research question was about the relative impact of skin tone and shape of faces on attractiveness judgments. Results were clear-cut. Even if skin tone undoubtedly contributes to attractiveness judgments, its contribution is much weaker than the contribution of the shape of the faces. Skin tone explained about five per cent of the variance as compared with 85% that was explained by the shape of the faces. This result was consistent with findings by (a) Fink, Grammer, and Thornhill (2001) who, in their study on the effect of skin texture on facial attractiveness of females did not find, among male White people, any significant correlation between darkness and attractiveness, (b) Wade and Bielitz (2005) who, in their study on the effect of skin tone on perceived attractiveness of African Americans, did not find, among White people, any significant effect of color on attractiveness judgments or any interaction involving color, and (c) Jefferson and Stake (2009) who, in their study on African American women’s dissatisfaction with body features, showed that only about 16% of them reported some degree of dissatisfaction with their skin tone.

The second one was about the way in which skin tone and shape of faces combine their effect in participant’s judgments. Results were also clear-cut. No Shape x Tone interaction was found; that is, both factors simply added their effect during the judgment process. The effect of skin tone on attractiveness judgment was the same, irrespective of the shape of face. This finding was consistent with results from the study by Fink, Grammer, and Thornhill (2001) who suggested that, in judgments of facial beauty, skin features are evaluated in addition to the facial shape.

Finally, an unexpected finding is still in need of explanation. The impact of shape of face on attractiveness judgments was stronger among
African participants than among European participants. This difference may be attributed to the way the three faces were selected. They were selected on the basis of a Mozambican criterion of beauty. If they had been retained on the basis of attractiveness judgments emitted by a French sample, it is possible that other faces would have been selected, and the impact of the face shape factor would have been stronger among French participants, and weaker among Mozambican participants.

**Limitations**

The present study had limitations. The first one resides in the way the samples were constituted. Participants were volunteers, and although special efforts were made to contact people from different geographic areas in both cities and from different educational levels, we are unsure about the representativeness of our samples. As a result, findings must be generalized with care to other geographic areas in both countries, and, of course to other countries.

The second limitation is that we used pictures, not real persons. Even though examining pictures is potentially different from seeing people, pictures are commonly used in studying people’s judgments (e.g., Jordan & Fernandez-Reif, 2009). In the present study, we used typical faces, and used credible skin tones. The participants had little trouble in making judgments; this is an important sign of the ecological validity of the material used. Moreover, in the present study, we needed to use pictures for the following reason. We examined how cues were weighted, how they were combined, and how different groups of respondents differed in weighting and combining. One condition for examining the processes of weighting and combining, independently of other processes, is that each participant has the same information presented in the same way.

The third limitation is that the effect of only one feature -- skin tone -- has been examined in detail. This feature is one of the many that facial characteristics that have been shown to impact on perceived attractiveness (Harvey et al., 2005; Thornhill & Gangestad, 1999). Among other ones are symmetry of traits, absence of senescence features, and sex hormone markers (e.g., homogeneous, smooth skin). In the present study all these other features have been subsumed under the “shape of the faces” factor, and their effect was not analytically assessed. In the same vein, the two factors considered in the design were attributed only three levels rather than, in daily reality, multiple values. This simplification was necessitated by the practical requirement to limit the number of pictures. Further studies
are needed to more precisely evaluate the shape of the relationship between each cue, presented with multiple levels, and the facial beauty ratings.

**Implications**

Face attractiveness undoubtedly depends on skin tone but the extent to which it depends from this factor is at best weak. In addition, the direction of the effect is certainly not homogeneous. If, in the present study, a majority of participants from Southern Africa preferred White faces, a majority of Western European participants preferred Black faces or Métis faces. Overall, and taking into account the health risks associated with skin bleaching, the net benefits associated with the lightening of one’s face are probably very weak. (Conversely, the benefits, in terms of attractiveness that are associated with skin tanning – a common practice among Europeans, are probably equally weak.) In addition, and from an evolutionary viewpoint, assessing facial attractiveness amounts, to some extent, to, consciously or unconsciously, assess health and reproductive capacity. Altering one’s skin tone or having excessive recourse to make-up makes facial features become unreliable. This may have unpredictable consequences at the time of mating.

Given past, and persistent racism in human and societal relationships, mainly based on skin color (Hill, 2002), one may wonder whether the results from this study, and also the ones by Fink, Grammer, and Thornhill (2001), Jefferson and Stake (2009), and Wade and Bielitz (2005) are really credible. Close examination of past laws that were expected to protect the White race from “contamination” lead, however, to conclusions that are consistent with the present findings. In South Africa, during the apartheid era and until 1985, sexual relationships between Black and White people were illegal, and mixed marriages were “unthinkable”: Transgressors were severely punished (Prohibition of Mixed Marriages Act, Act No 55 of 1949). If people’s attractiveness would have largely depended on skin tone, then these laws would have been superfluous. Black people would have been “naturally” protected against the idea that White people can be attractive, and reciprocally. The very existence of these laws and their systematic implementation constitute a material demonstration that, regarding face beauty, and attractiveness, skin tone is not very important: About 11% of U.S. Black people who married in 2007 married White people (Pew Report, 2010).
REFERENCES


(Manuscript received: 30 June 2013; accepted: 7 October 2013)