Cognition 191 (2019) 103972

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

Original Articles

When we don't know what we know – Sex and skin color

Mariana L. Carrito[a], Gün R. Semin[b,c,*]

[a] Centre for Psychology at University of Porto, Faculty of Psychology and Education Sciences – University of Porto, Rua Alfredo Allen, 4200–135 Porto, Portugal
[b] William James Center for Research, ISPA – Instituto Universitário, Rua Jardim do Tabaco 41, 1149-041 Lisboa, Portugal
[c] Faculty of Social and Behavioral Sciences, Utrecht University, 3584 CS Utrecht, the Netherlands

ARTICLE INFO

Keywords:
Skin color
Representation of gender
Implicit and explicit association
Memory

ABSTRACT

In 3 experiments, we examine how the abstract category of gender, grounded by the lightness-darkness dimension, derived from the universal sexual dimorphism of skin color, is represented and how such representations lead to visual accentuation processes, i.e. polarization of differences between male and female faces. In the first two experiments, we show that irrespective of whether grayscale male and female faces are presented sequentially or jointly, female faces are judged to be lighter than male faces when participants are asked to indicate the level of lightness of the faces. This pattern was found for the majority of participants who explicitly stated that men and women do not differ in skin color. The third experiment was designed to examine the cognitive consequences of what people implicitly 'know' with a perceptual accentuation study. Participants were provided with male and female faces of equal skin color. Subsequently, in a memory recall task, they were asked to select, from a row of several faces varying in skin color, the original face. They chose, as predicted, lighter versions of faces for females compared to the male faces. This research reveals that the evolutionarily based sexual dimorphism in skin color implicitly grounds gender categories and shapes implicit visual accentuation processes.

1. Introduction

Recent research has shown that gender is marked by brightness, with lighter colors (or white) marking the female gender and darker ones (or black) marking the male gender (Semin & Palma, 2014; Semin, Palma, Acartürk, & Dziuba, 2018). The studies reported so far have examined how gender-related classification processes (e.g., classifying names as male or female) are influenced by shades of a color or black and white, revealing, for instance, that male names in black (or darker colors) and female names in white (or lighter colors) are classified significantly faster than the reverse combinations. Paralleling these findings is the well documented sexual dimorphism in skin color showing that this dimorphism is a distinctive and universal adaptive feature of actual skin color (e.g. Jablonski & Chaplin, 2000).

However, do people represent implicitly female skin color as lighter and male skin color as darker? Thus, is it likely that we actually have incorporated the sexual dimorphism of skin color in our representations by extracting a systematic difference in skin color without being aware of it? The experiments reported here are designed to answer these questions.

In the following, we first provide the background to this research to then furnish an overview of the current studies before we turn to the three experiments.

1.1. Women and men differ in skin color

Recent research on the dimorphism of skin color between males and females has revealed that this is a distinctive and universal adaptive pattern (Jablonski & Chaplin, 2000, 2002; Jablonski, 2004). According to Jablonski and Chaplin: "Throughout the world, human skin color has evolved to be dark enough to prevent sunlight from destroying the nutrient folate but light enough to foster the production of vitamin D" (Jablonski & Chaplin, 2002, p. 74). Consequently, skin pigmentation varies systematically geographically and there is a strong correlation between skin reflectance and latitude (Chaplin, 2004; Jablonski & Chaplin, 2000, 2013, 2018; Walter, 1958). What is more important in the current context is that orthogonal to this systematic geographical variation, there is a systematic difference in the skin color between males and females. Females are consistently lighter than males in all studied populations (Jablonski & Chaplin, 2000; Jablonski, 2004). There have been diverse arguments to explain the observed sexual dimorphism in skin pigmentation related to natural selection (Jablonski & Chaplin, 2000, 2013, 2018) and/or sexual selection (Aoki, 2002; Ihara & Aoki, 1999) mechanisms. The natural selection argument holds that

* Address: William James Center for Research, ISPA – Instituto Universitário, Rua Jardim do Tabaco 41, 1149-041 Lisboa, Portugal.
E-mail addresses: mariana.carrito@gmail.com (M.L. Carrito), gun.r.semin@gmail.com (G.R. Semin).

https://doi.org/10.1016/j.cognition.2019.05.009
Received 9 January 2019; Received in revised form 13 May 2019; Accepted 14 May 2019
Available online 20 June 2019
0010-0277/

females need to maximize cutaneous vitamin D3 production in order to maintain the higher demands they have for calcium requirements during pregnancy and lactation. In contrast, darker pigmentation for males enables the maintenance of folate levels that are required to protect sperm production, which depends on folate for DNA synthesis. On the other hand, sexual selection may operate when individuals base their mating decisions on skin coloration of potential partners (Van den Berghe & Frost, 1986), with skin tone signaling fertility and mate quality and influencing attractiveness perception (Carrito et al., 2016).

1.2. The association between gender and brightness

Semin and his colleagues (Semin & Palma, 2014; Semin et al., 2018) have demonstrated the alignment of female-male with the bipolar dimension of light–dark by using a speeded gender classification task and showing that male names are processed faster when they are presented in a black typeface (Semin & Palma, 2014, Exp. 1) or a dark color (Semin & Palma, 2014, Exp. 2) than when they are presented in white or a light color, with the opposite pattern for female names. In a further experiment (Semin et al., 2018, Exp. 2), they utilized a gender classification task involving the disambiguation of very briefly appearing non-descript stimuli in the form of black and white "blobs". The black blobs were classified predominantly as male and the white ones as female names. Finally, processes driving light and dark object choices for male and female targets were examined by tracking the number of fixations and their duration in an eye-tracking experiment (Semin et al., 2018, Exp. 3). As predicted, when choosing for a male target, participants look longer and made more fixations on dark objects, while the reverse was found in choices for light objects when the target for whom a choice was made was a female. Furthermore, in all these studies no effects for participant gender were found.

Finally, we should draw attention to the oft reported relationship between valence and the brightness dimension and the sexual dimorphism in skin color. There is a substantial literature that has rested on the idea that darkness is associated with negativity, and lightness with positivity. Indeed, numerous studies have shown that the dimension of brightness-darkness grounds valence (e.g. Alter, Stern, Granot, & Balcetis, 2016; Lakens, Semin, & Foroni, 2012; Meier, Fetterman, & Robinson, 2015; Meier, Robinson, & Clore, 2004; Specker et al., 2018). However, the research to date shows that the brightness dimension can function orthogonally when grounding different categories, much like mental metaphors from physical space that structure time (e.g. Boroditsky, 2001; Majid, Gaby, & Boroditsky, 2013), valence (Meier & Robinson, 2004), number (Dehaene, Bossini, & Giraux, 1993) and power (Schubert, 2005). As we have reported in our earlier studies that have controlled for valence, brightness has been shown to be orthogonal to the positive-negative dimension when marking gender (Semin & Palma, 2014; Semin et al., 2018).

1.3. The current Studies: an overview

The research to date has not examined directly whether or not the difference in skin color of males and females is explicitly available to participants. That is the question that the three experiments reported here examine. We investigate this question by using two converging paradigms and their relationship. The first two experiments ask participants to match the skin color of a male and/or female target person's monochromatically presented face by using a visual analogue scale (VAS) which varies from a light to a dark skin color. Moreover, in the same experiments, participants were explicitly questioned, through multiple choice questions, about their beliefs on skin color sex differences/similarities. One rather inviting implication of only implicit access to skin color differences concerns the cognitive consequences of the implicit representation of sexual dimorphism of skin color. It is well known that placing a set of stimuli in the same category versus splitting them into different categories can affect similarity ratings (Tajfel & Wilkes, 1963). Thus, providing *explicit* category labels is sufficient to increase perceived differences between categories and reduce perceived differences within the categories.

Lastly, using a delayed reproduction task (experiment 3) provided the rather intriguing and novel possibility to use the perceptual accentuation phenomenon as an additional test of the explicit-implicit access to the sexual dimorphism of skin color. This involved asking participants to identify which of 7 face images (varying in skin shading and including the one with the target skin color) had been presented previously. In this, 'implicit' task, where participants would not be provided with a label, recall was expected to be biased in favor of the sexual dimorphism in skin color if participants implicitly activate the gender category. Female skins would then be recalled as lighter than male skins. If the sexual dimorphism of skin color activates perceptual accentuation, then this would be the first time that this process would be demonstrated without supplying categorical labels.

2. Experiment 1

In experiment 1, participants were presented with images of male and female faces that were converted to grayscale. Their task was to indicate the skin color of each sequentially presented image on a VAS ranging from a light to a dark shade of skin color. This task was followed by an explicit question regarding the relative skin color of males and females, namely if males (females) have lighter or darker skin than females (males) or if there was no difference between the skin color of males and females. While the first experimental task constitutes the 'implicit' mode of assessing the relationship between sex and skin color, the second task constituted the 'explicit' assessment. The first task is 'implicit' insofar as it requires the representation of sexual dimorphism in skin color to shape judgments of skin color shading systematically without the participant knowing the comparative hypothesis driving the task. The second task is explicit, as it requires participants to report their accessible knowledge of sexual dimorphism in skin color.

2.1. Method

2.1.1. Participants

Participants from this and subsequent experiments were screened online using the Prolific platform (https://prolific.ac/) and were paid for their participation. In line with the host institution's ethical research guidelines, in all 3 experiments, the participants were asked to read and agree with the informed consent. Participants were assured that all data collected would be treated anonymously and would only be published in scientific outlets and that they were free to withdraw at any time without providing a reason. Across all three experiments, the studies were performed in accordance with the Declaration of Helsinki for experimentation with human subjects. Sample size was estimated using G*Power 3.1.9.2 software, considering a medium effect size ($F = 0.25$), an alpha of 0.05 and a power of 0.9. A total sample of 46 participants (23 per group) was obtained. Considering the possibility of one or two participants not performing the task properly, a per-group size of 24 was considered in this and subsequent experiments. The inclusion criteria were ethnicity (Caucasian), age (between 18 and 35 years old), first language (English speakers) and visual health (not reporting color blindness or other vision problems that were not corrected using glasses or contact lenses).

2.1.2. Stimuli

Seventy-two photographs (36 male and 36 female) were selected from the St Andrews database of Caucasian adult faces. The photographs were taken under controlled illumination conditions and were posteriorly color calibrated (see Whitehead, Re, Xiao, Ozakinci, & Perrett, 2012 for detailed procedures). The faces were delineated with 192 points (with X and Y coordinates) using *Psychomorph* software (Tiddeman, Burt, & Perrett, 2001). The delineation process intends to outline different face areas, which is required for averaging processes and skin color transformations. Groups of three different facial photographs, from the same sex, were averaged together to create 12 composite male faces and 12 composite female faces. The use of composite faces instead of the original photographed ones allowed us to obtain lower levels of inter-individual differences. Also, this procedure ensures that the stimuli have average levels of skin color, allowing subsequent standardization of color manipulations.

Skin areas of these 24 faces (including lips and eyebrows but excluding eyes - sclera, iris, and pupil) were color-transformed considering natural variations of luminance following Lefevre and Perrett (2015). To do so, we created two face-shaped uniform color masks in *Matlab*, representing high luminance (+2.7 units of L*) and low luminance (−2.7 units of L*). These masks were used to create two new versions of each face, a lighter one that was color transformed towards the high luminance mask (−100% transform) and a darker one that was color transformed towards the low luminance mask (+100% transform). Face shape and texture were kept constant. An oval grey mask was later applied to occlude the hair, neck, ears, and background of the three versions of each face: the original, the lightened and the darkened one. All images were converted to grayscale and resized to 480 × 480 pixels, with a resolution of 300 × 300 dpi (see Fig. 1). Grayscale images were used to remove color information which could influence participants' responses.

The same color masks were also used to create two color patches, one representing high luminance and the other representing low luminance. These patches were designed to anchor the extremes in a later VAS. These images were cropped as 150 × 150 pixels squares, with uniform coloring in all area of the patches.

2.1.3. Procedure

The Qualtrics platform (https://www.qualtrics.com) was used to implement the procedures in this and subsequent experiments. Participants started by answering a short demographic information questionnaire. Subsequently, they were instructed to look at the grayscale images of faces and imagine the skin color of each presented person. In each trial, one face was presented at a time along with a VAS (−50 – +50) as illustrated in Fig. 2, with two color patches anchoring the extremes of the scale (see above in the stimuli section). Participants were told that they could place the marker in any point on the scale ranging from a value of −50 as representing the lighter patch (+2.7 L*) to the value of +50 as representing the darker patch (−2.7 L*). The following instructions were given: "Please indicate whether the skin color of the presented face is more like the color shading on the right or the left patch. To do so, place the pointer to the position you think best corresponds to what you think the person's skin color is likely to be." Participants were presented with 24 faces, 12 of each sex, in a random order. Within each sex group of faces, a third (4 of the faces) was presented in its grayscale version of their original coloration, a third was presented in a darkened version, and a third was presented in a lighter version. The reason why we chose to present participants with original, darkened and lightened faces of both sexes was to avoid giving away information about skin color that could directly influence participants' response.

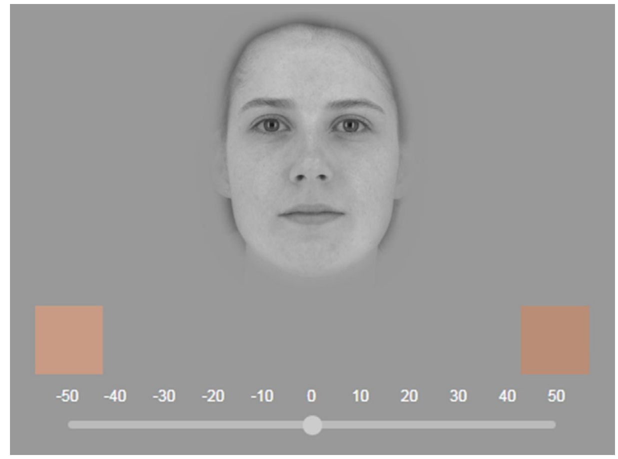

Fig. 2. Representation of the face task in experiment 1.

Finally, participants were asked to answer a question regarding what they considered to be the typical skin color of women and men. Using one of the three options they had to indicate if they considered the skin color of males to be darker, equal or lighter compared than the skin color of females. The order of the words "females", "males", "lighter" and "darker" across the three sentences was counterbalanced across participants.

Fig. 1. Colour transformation applied to faces in Experiment 1 and 2. Left versions represent decreased darkening (−100%), the ones in the middle represent the original images, and the right versions represent increased darkening (+100%). Note that only one of the versions of each face was presented to the participants. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

2.2. Results

2.2.1. Perceived typical skin color of women and men

On the question regarding the typical skin color of men and women, a chi-square goodness of fit test was performed to determine whether the verbal options would be equally chosen. Responses to the verbal questions were not equally selected, X^2 (2, N = 48) = 9.13, p = .01. Fifty percent of the participants chose the "The skin of males/females is as light/dark as the skin of females/males" ([F = M]) option; 35.42% of participants chose [MD − FL] – "The skin of males is darker than the skin of females" (or "The skin of females is lighter than the skin of males"); 14.58% of participants chose [FD − ML] – "The skin of females is darker than the skin of males" (or "The skin of males is lighter than the skin of females").

2.2.2. Level of darkness chosen for male and female faces

For each participant, the mean degree of darkness selected for the male and female faces was calculated. These values were normally distributed (Kolmogorov–Smirnov, $p > .129$), except in the case of male color choices of male participants which showed an acceptable level of skewness of 0.392 ($SE = 0.47$) and kurtosis of -0.950 ($SE = 0.92$). One sample t-tests revealed that darkness choices were greater than chance (i.e. 0%, which would mean a choice not different from the original face) for male faces, t (47) = -6.55, $p < .001$, $d_s = -1.89$, and female faces, t (47) = -10.64, $p < .001$, $d_s = -3.07$. A 2-way mixed ANOVA (between-subjects factor: sex of participant; within-subjects factor: sex of the face) revealed a main effect of sex of the face, F (1, 46) = 21.21, $p < .001$, $\eta_p^2 = 0.316$, indicating that participants selected a lighter skin color for females (M = -21.38, SE = 2.02) than males (M = -13.20, SE = 2.00). The effect of sex of participant did not reach statistical significance (p = .22). Similarly, the interaction between the sex of face and the sex of participants was not significant (p = .53).

2.2.3. Level of darkness chosen for both sexes faces, grouping participants based on their explicit responses.

To examine the relationship between explicit answers and [VAS], we split the [VAS] data by [verbal] categories. As the number of participants for each explicit category varied considerably, we performed Wilcoxon signed-rank tests for each group of [verbal] ([MD − FL], [F = M] or [FD − ML]) separately. Participants that selected [MD − FL] ($Z = -3.20$, p = .001, $r = -0.77$) as well as participants that selected [F = M] ($Z = -2.29$, p = .022, $r = -0.47$) chose a lighter coloration for the female compared to the male faces. The level of darkening chosen for male and female faces did not differ for participants that considered [FD − ML] ($Z = -0.85$, p = .398). Table 1 shows descriptive statistics for the choices of all groups.

3. Discussion and experiment 2

In experiment 1, we found a clear tendency for participants to attribute a lighter coloration to female when compared to male faces. We also found a sliding difference on the [VAS] responses from those who on the explicit response choices said that males are darker than females (50%) to those who say that there is no difference in skin color between the genders (35.42%) to those who said that females are darker than males (14.58%). Notably, those who said that there is no difference in skin color between the gender nevertheless show a systematic and significant difference on their [VAS] responses judging females as having a lighter skin color than males. Indeed, even those who endorse the opposite of the sexual dimorphism in skin color do not show any systematic relationship on their [VAS] responses.

It is possible to argue that when participants depict the skin color shading of sequentially presented males and females they are unable to judge their relative relationship. However, when both a male and a female face are presented and judged jointly then there may not only be a consistency between the two types of tasks but also a more uniform outcome across participants. To this end, in our second study participants were presented with a grayscale male and a female face simultaneously and responded with the VAS to each of them and could see their responses to both faces.

The other consideration relates to the presentation order of the face task and the explicit task. It is possible to argue that the order of explicit and implicit tasks sensitizes participants to the skin color relationship and influences the possible outcomes. To this end, the order of task presentation (Verbal question and face task) was counterbalanced.

Table 1
Descriptive statistics of the level of darkness chosen in the face task of experiment 1, aggregated by the response given in the verbal task. Negative values represent lighter choices.

	[F = M] N = 24 *Mean (SE)*	[MD − FL] N = 17 *Mean (SE)*	[FD − ML] N = 7 *Mean (SE)*
Male faces	−15.30 (2.95)	−10.89 (3.40)	−11.61 (4.89)
Female faces	−20.57 (2.88)	−24.34 (3.19)	−17.00 (5.92)

3.1. Method

3.1.1. Participants

Forty-eight online participants, 24 men, and 24 women, aged between 18 and 35 (M = 27.29 years, SD = 4.78), were selected based on the same inclusion criteria as in experiment 1. All participants reported Caucasian ethnicity and absence of color vision problems. Again, all participants were informed that the material gathered during the research would be treated as confidential and that they could withdraw from the survey at any time.

3.1.2. Stimuli

The exact same three versions (original coloration, darken version and lightened version) of each twenty-four composite faces such as the two-color patches used in experiment 2 were utilized in this experiment. Again, for the face images, hair, neck, ears, and background were occluded by an oval grey mask and images were converted to grayscale and resized to 480 × 480 pixels, with a resolution of 300 × 300 dpi.

3.1.3. Procedure

Participants started by supplying demographic information. As previously they were presented with grayscale images of faces and were instructed to look at them imagining the skin color of each person presented. This time, each trial included a presentation of pairs of faces, with a male face presented beside a female one. Below each pair of faces, two VASs (-50 – $+50$) were presented with two color patches anchoring the extremes of the scale, with one VAS to be used for the male face and the other for the female face. Instructions were given so that participants could place the marker in any point of the scales between the lighter patch and the darker patch as in experiment 1. All twenty-four faces were split randomly into two sets, wherein each set of twelve faces had the same number of male and female faces. Each set was allocated to be presented on one of the sides of the screen, with this allocation being randomly assigned across participants. Thus, each participant saw twelve pairs of opposite-sex faces, in which half of the trials presented the male face on the left side of the screen, while on the other half the male face was presented in the right side of the screen. Also, as in experiment 1, only one of the darkening versions of each face was presented, with darkened male versions being paired with lightened female ones (4 pairs), lightened male versions being paired with darkened female ones (4 pairs), and original male and female versions being paired together (4 pairs). Again, the purpose of this procedure was to avoid giving away information to the participants about possible natural sexually dimorphic color cues during the experimental task. The trial order was randomized across participants. Likewise, we randomized the allocation of the color patches in each extreme of the VAS (the light one on the right extreme with the dark one on the left extreme and vice-versa) and the order of the male and female VAS (the male one appearing above the female one and vice-versa).

As in experiment 1, participants were asked to answer a multiple-choice question regarding the typical skin color of women and men: [F = M]; [MD − FL] and [FD − ML]. Again, the order of the words ("females", "males", "lighter" and "darker") across the three sentences was counterbalanced across participants. Now, however, the explicit question appeared after the experimental face task for half of the participants and before the face task for the other half.

3.2. Results

3.2.1. Judgements of the skin color of women and men

The large majority of the 24 participants that started the experiment by responding to the explicit task selected [F = M] (87.5%). The [MD − FL] option was selected by 8.3%, with the remaining 4.2% choosing [FD − ML]. Chi-square goodness of fit test confirmed that, for these participants, verbal responses were not equally selected, X^2 (2, $N = 24$) = 31.75, $p < .001$. Regarding the participants that responded to the explicit question at the end of the experiment, the choices were also not equally distributed, X^2 (2, $N = 24$) = 7.75, $p = .021$, and 54.2% of participants selected [F = M]. The second most selected option was [MD − FL] selected by 37.5%, with [FD − ML] being chosen by 8.3% of the participants.

3.2.2. Level of darkness chosen for male and female faces

Once more, mean degree of darkness chosen was calculated for male and female faces. Values were normally distributed (Kolmogorov–Smirnov, $p > .20$). One sample t-tests revealed that darkness choices were greater than chance for male faces, t (47) = −3.97, $p < .001$, $d_s = -1.15$, and female faces, t (47) = −6.00, $p < .001$, $d_s = -1.73$. A 3-way mixed ANOVA (between-subjects factors: sex of participant and order of tasks; within-subjects factor: sex of the face) revealed a main effect of sex of the face, $F(1, 44) = 8.89$, $p = .005$, $\eta_p{}^2 = 0.168$, reflecting a greater tendency of participants to choose a lighter coloration for female faces ($M = -11.36$; $SE = 1.89$) when compared to male faces ($M = -6.28$; $SE = 1.59$). The effect of sex of participant was not statistically significant ($p = .45$), neither was the order of tasks ($p = .18$) or the remaining interaction effects ($p_s > 0.071$).

3.2.3. Level of darkness chosen for both sexes faces, grouping participants based on explicit responses

Wilcoxon signed-rank tests were performed for each group of participants (grouping was performed based on [verbal]). Participants that, in the explicit task, chose [MD − FL] selected a lighter tone for female than male faces ($Z = -2.50$, $p = .013$, $r = -0.79$) as did those who chose [F = M] ($Z = -2.21$, $p = .027$, $r = -0.38$). Participants that selected [FD − ML] did not choose differently in the face task when considering male and female faces ($Z = -0.37$, $p = .715$). Table 2 shows descriptive statistics for the choices of all groups.

Table 2
Descriptive statistics of the level of darkness chosen in the face task of experiment 2, aggregated by the response given in the verbal task. Negative values represent lighter choices.

	[F = M] N = 24 *Mean (SE)*	[MD − FL] N = 17 *Mean (SE)*	[FD − ML] N = 7 *Mean (SE)*
Male faces	−4.87 (1.67)	−9.57 (3.94)	−10.02 (8.62)
Female faces	−9.69 (1.94)	−17.93 (4.91)	−9.17 (9.89)

3.3. Discussion of experiment 2

Experiment 2 was designed to examine whether the simultaneous presentation of male and female faces result in differences in the implicit knowledge of the sexually dimorphic nature of skin color. This was not the case since participants generally selected lighter colorations for female compared to male faces. Once again, these findings were observed not only for participants that explicitly stated that males were darker than females but also for those who said that males and females have the same skin color. Irrespective of task order, the majority of participants in both groups chose to explicitly state that females and males have the same skin color, although the percentage of such response was higher for participants who responded to the verbal question first. The percentage of participants considering females and males to have the same skin color was lower for the group that performed the verbal task after the face task, probably showing that for some few participants, doing the face task resulted in a conscious recollection of a knowledge that was previously implicit.

One might argue that social desirability is the driving motive for the high percentage of participants making the choice of same skin color for males and females, in other words, participants did not wish to appear biased. However, if the desire not to appear biased was expressed in a preference for equal shading of skin color for males and females, then the pre- and post-measurement percentages of skin color relation for males and females should have been reversed, that is the percentage of participants choosing equal skin coloration should have been higher in the post-test. This is because participants should have been more sensitized to potential bias in their indication of skin color differences after performing the experimental task and thus have made a higher percentage of equal skin color choices. That this is not the case suggests that social desirability did not play a role in the choices made by participants.

4. Experiment 3

In experiment 3, we examined if the implicit knowledge of sex differences in skin color not only impacts face perception but also gives rise to memory distortions. Implicit memory effects have been extensively investigated in psychological studies (e.g. Roediger, 1990; Schacter, 1987). In line with the findings in the two preceding experiments that demonstrated the shared implicit knowledge of sexual dimorphism in skin color, we expected that in the recall, the skin color difference between females and males would be visually accentuated. Moreover, no category labels were provided.

This argument is derived from the classic paper by Tajfel and Wilkes (1963) who demonstrated that providing explicit category labels is sufficient to increase perceived differences between categories and reduce perceived differences within the categories. Levin and Banaji (2006), more recently, illustrated this by showing that to categorize a face as Black or White is sufficient to affect the perceived lightness of the skin tone of a target. For instance, when faces were morphed to create a racially ambiguous Black–White image, *explicit* labelling of the face as either White or Black shifted the perceived skin tone of the target. When the category Black was used and when the participant was asked to match the skin tone of the target then they perceived the skin tone of the targets to be darker compared to when the target was labelled White.

In experiment 3 participants were not given any labels. They were presented with images of male and female faces sequentially. They observed these faces, one at a time, for a brief period, and were then asked to identify the presented target among 7 versions of the original face that varied in the level of luminance presented in a random order within a row, including the original.

4.1. Method

4.1.1. Participants

Ninety-seven online participants, 45 men, and 52 women, aged between 18 and 35 ($M = 27.78$ years, $SD = 4.31$), were selected based on the same criteria used in experiments 1 and 2. Fifty of those volunteers (group 1) participated in the experimental condition where the color manipulation of faces shown in the recall phase ranged between −100% and 100% darkness. The remaining 47 participants (group 2) were in the second experimental condition in which the color manipulation of faces shown in the recall phase ranged between −75% and 75% darkness. The rationale for the color manipulation range is explained in detail in the stimuli section below. All participants reported Caucasian ethnicity and absence of color vision problems.

4.1.2. Stimuli

The same twenty-four composite faces from the previous experiments, 12 of each sex, were used for further color transformation using the same color masks, representing high luminance (+2.7 units of L*) and low luminance (−2.7 units of L*). The luminance of the composite faces was manipulated in order to obtain a set of seven images for each face, ranging from −100% (decreased darkening) to 100% (increased darkening) for group 1 of participants, and from −75% (decreased darkening) to 75% (increased darkening) for group 2, with the middle image being the original composite face for both groups of participants. To our knowledge, this investigation is the first to explore the impact of sexually dimorphic skin color differences on the memory of faces. Therefore, different ranges of luminance were applied to faces of the recall phase, as there are no previous findings addressing how the range of variation might impact participants' responses in such a task. Also, the decision of using differentiated ranges was justified acknowledging the influence of question format on response style behavior across psychological research (Wetzel, Böhnke, & Brown, 2016). Systematic variance may arise depending on measurement method (Kieruj & Moors, 2010) and we expected participants to show an increased preference for midpoint/ average faces when presented with a wider range of manipulation, although such an effect was expected not to hinder accentuation processes. Again, the hair, neck, ears, and background were occluded from view. All images were resized to 480 × 480 pixels, with a resolution of 300 × 300 dpi.

4.1.3. Procedure

First, a short questionnaire was administered to obtain demographic information. The experiment started with a fixation cross for one second, followed by an image for two seconds, and a blurred pixelated image for another two seconds (see Fig. 3). The blurred pixelated image was shown to prevent participants from basing their responses on the iconic memory of the coloration that had just been presented. In the next step of the experiment, participants were presented with seven versions of the face they had seen. These 7 faces varied in terms of their levels of darkness and were presented in a random order on a row. Participants' task was to choose the picture that was presented previously in isolation.

4.2. Results

Participants could choose between 7 faces, one identical to the one that had been previously shown, and other 6 varying in shading. For the ease of interpretation, lighter versions have been coded between −3 and −1, 0 as the original version and darker versions as between 1 and 3. For each participant, the mean degree of darkness chosen was calculated for male and female faces. Values were normally distributed (Kolmogorov–Smirnov, $p_s > 0.59$), except for responses given by participants of group 2 towards female faces which showed acceptable skewness of 0.262 ($SE = 0.35$) and kurtosis of −0.627 ($SE = 0.68$). One sample t-tests revealed that the level of darkness selected in recall was greater than chance for male faces, t (96) = −5.51, $p < .001$, $d_s = -1.12$, and for female faces, t (96) = −6.79, $p < .001$, $d_s = -1.38$.

A 3-way mixed ANOVA (between-subjects factors: sex of participant and luminance range in the recall phase; within-subjects factor: sex of the face) revealed a main effect of sex of the face, F (1, 93) = 4.46, $p = .037$, $\eta_p^2 = 0.046$, with lighter versions of faces selected for female faces ($M = -0.56$, $SE = 0.08$) compared to male faces (M = -0.42, $SE = 0.07$) (see Fig. 4).

The effect of luminance range in the recall phase was also statistically significant, F (1, 93) = 11.85, $p = .001$, $\eta_p^2 = 0.113$, with lighter versions of faces selected for group 2 who were presented with a narrower range of lightness-darkness in the row of possible choices ($M = -0.72$, $SE = 0.10$) compared to group 1 who were presented with a wider range of lightness-darkness as possible choices (M = -0.26, $SE = 0.09$). The effect of sex of participant was not significant ($p = .079$) and the same was true for all interaction effects (all other $p_s > 0.466$).

4.3. Discussion of experiment 3

Experiment 3 explored whether implicit representations of the categorical association between gender and skin color would be visually accentuated as memory distortions. As predicted, participants recalled lighter versions of female faces when compared to the versions recalled for male faces. In fact, this accentuation effect in memory occurred even in the absence of explicit categories, as the procedure did not include any verbal information about gender.

The effect of luminance range turned out significant and, as expected, participants from group 2 (−75 – +75% range) chose faces with higher levels of lightness when compared to participants from group 1 (−100 – +100% range). Such an effect may indicate some concordance in the level of lightness recalled by participants of both groups, considering that faces presented in the recall row were color manipulated across different luminance ranges. As maximum lightness increase corresponded to +2.7 units of L* in group 1 and +2.025 units of L* in group 2, participants of group 2 would have to select more extreme versions to obtain an identical level of luminance to the ones chosen by group 1. Hence, if participants had wanted to choose a −50% luminance face, they would have to select a more extreme value in group 2 but a more moderate/central value in group 1.

Lastly, it is also evident that, once again, the categorization effect evidenced in this recall task, is independent of the sex of participants, showing that both women and men distort their memory in recognizing the target face along the same representations of sexually dimorphic skin color.

5. General discussion and conclusions

Earlier research (Semin & Palma, 2014; Semin et al., 2018) revealed, across a number of different experimental paradigms and across three different cultures, a systematic interface between the brightness dimension and gender. None of these studies established a direct link between the sexual dimorphism of skin color and the types of tasks that were used in the diverse experimental paradigms. The nature of the tasks used in these studies revealed that the processes driving the categorization processes that are involved are implicit but not that these implicit processes derive directly from a socially shared notion of sexual dimorphism of skin color or from an implicit representation of the association between sex and skin color. The three studies reported here were aimed at unravelling this issue – namely, that the sexual dimorphism in skin color drives the categorization patterns that were observed. Moreover, the studies reported here were designed to illuminate whether the sexual dimorphism of skin color is a consciously accessed and a socially shared representation or not.

In experiments 1 and 2, we found a significant effect of 'sex of face'

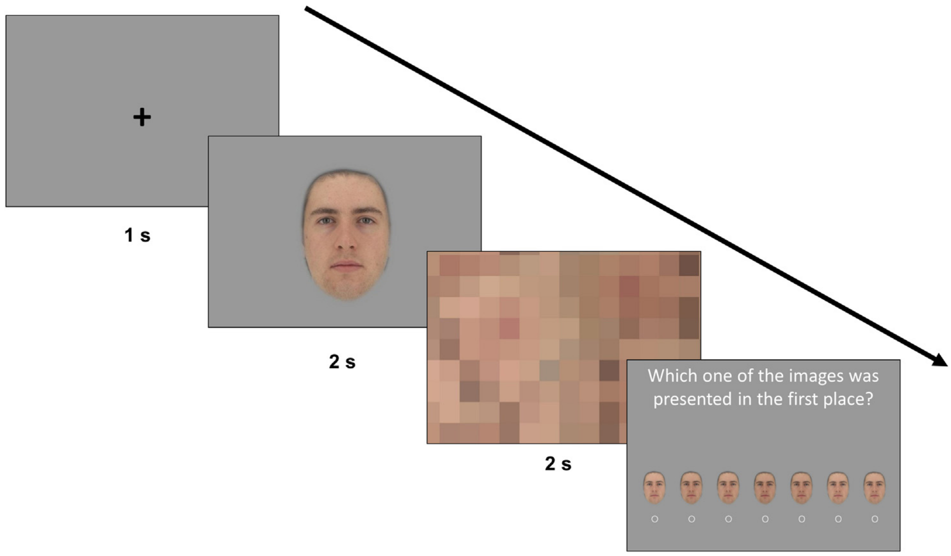

Fig. 3. Representation of the sequence of events in the procedure of experiment 3.

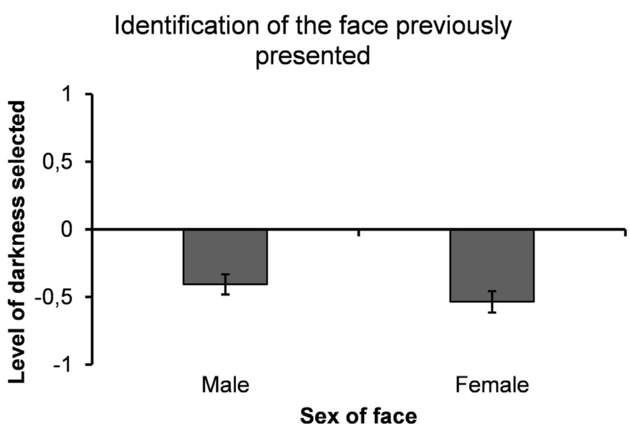

Fig. 4. Mean level of darkness of the faces selected in experiment 3. Negative values represent choices of lighter versions. Error bars show standard errors of the mean.

when asking participants to indicate the level of darkness/lightness of faces, namely that the majority of participants chose a lighter coloration for female compared to male faces, while they explicitly expressed that men and women have equal skin tones. These contradictory responses on the two variables reveal that participants' implicit knowledge as measured by [VAS] is driven by a representation of sexually dimorphic skin luminance that does not inform their explicit responses. These findings, consistent across both experiments, suggest that a substantial proportion of participants have no access to why they respond the way they do on [VAS], irrespective of whether the explicit task is before or after the face task (Exp. 2) or irrespective of whether the task involves a comparison where both faces are presented sequentially (Exp. 1) or simultaneously (Exp. 2). What is noteworthy is that when participants in experiment 2 fill out the verbal question first, a remarkable 88% choose the option that males and females have the same skin color. This drops to, a still respectable, 54% if they respond to the explicit item after filling the VAS scales. This suggests that they are possibly influenced by their VAS responses preceding their explicit judgments and may reflect a visual but not reflective memory effect. In other words, their shading responses may remind a percentage of participants by their visual responses that males are darker than females, that then translates to an overall lower percentage in [verbal]. Moreover, because for participants choosing asymmetric responses, the [F = M] verbal choice (females are darker than males) was so rare, it may be possible that explicit judgments are also influenced by implicit knowledge in a way that prevents participants to choose the option that is the complete opposite to what they encountered in their daily lives. Nevertheless, [VAS] results for all participants show the predicted direction: females have a lighter skin color than males.

Experiment 3 provided evidence that the association between the brightness dimension and gender gives rise to memory distortions. Participants tended to recall lighter versions of faces when presented with female faces than when presented with male ones. This is also consistent with the classic accentuation effect evidenced by Tajfel and Wilkes (1963) (see also Corneille, Klein, Lambert, & Judd, 2002). Subsequent work focusing on the impact of categorization on memory distortions has been documented on faces (e.g. Corneille, Huart, Becquart, & Brédart, 2004; Huart, Corneille, & Becquart, 2005), and other objects (e.g. Hansen, Olkkonen, Walter, & Gegenfurtner, 2006). Alternatively, and considering the literature on sex differences in facial contrast (Russell, 2009; Russell, Kramer, & Jones, 2017) advocating female faces to have greater luminance contrast between internal facial features (eyes and lips) and the surrounding skin, it is possible that our participants relied their responses also on knowledge about sexually dimorphic luminance contrast. Future work should address awareness and accentuation of facial contrast as past work has shown it to enhance identification (Lee & Perrett, 2000).

The difference between the current study and Tajfel and Wilkes (1963) seminal work, as well as its follow-ups, is, however, that the current study is not conducted with explicitly provided categories but rather implicitly accessed representations of the categorical association between gender and skin color. This suggests that the implicitly prominent representation of the natural difference in skin color for the sexes, as seen in the first two experiments, accentuates a difference in recall, as demonstrated in Experiment 3. This implicit accentuation phenomenon is in itself a novel contribution and opens a new road into examining accentuation phenomena driven by implicit categorization processes.

Finally, we would have expected an accentuation in opposite categorical properties, male skin color being recalled as darker and female skin color being recalled as lighter than the stimulus luminance. Surprisingly, this was not the case since participants recalled overall lighter than original versions of the face of both sexes. It is possible that such finding is due to the ethnical backgrounds of participants, who were all Caucasian, first language English speakers, and hence may encounter frequently both sex people with lighter skin.

One of the central contributions of this research is for how categories are grounded. In the case of gender, we find that it is marked by the brightness dimension that is directly associated with gender and that this marking is predominantly an implicit one. The second contribution of the current research is to categorization processes, namely accentuation (Tajfel & Wilkes, 1963). The final study in this paper has demonstrated that accentuation processes can be driven by 'implicit' categories, namely gender in this case, rather than one that is driven by explicitly accessible labels. The origin of these accentuation effects is believed to result from a universal adaptive pattern (Jablonski & Chaplin, 2000; Jablonski, 2004) and may arise early in life when individuals encounter sexually dimorphic traits in the faces of parents or siblings, being latter further enhanced by socio-cultural factors. In fact, recent findings report that female children in a non-Western civilization, by the age of 6, attribute light objects to women and dark ones to men (Sebastián-Enesco & Semin, submitted for publication). However, similar future studies should continue exploring whether the reported effects are found across different ethnic groups and cultural backgrounds. Beyond adding to the literature documenting the impact of visual accentuation processes, practical implications of our findings include the understanding and prediction of cultural groundings of color, namely in ceremonies (e.g. weddings where the bride wears white and the groom wears black) and moral concepts (e.g. association between white and purity), while also adding evidence to perception and memory bias that influence legal actions (e.g. eyewitness testimonies).

Overall, these studies along with the previous ones (Semin & Palma, 2014; Semin et al., 2018) spell out the steps of how a physically given difference, namely sexual dimorphism grounds the categorical representation of gender and shapes cognitive processes implicitly. These processes range from speeded gender classification tasks to preferences of commercial objects differing in color or brightness (e.g. Semin & Palma, 2014; Semin et al., 2018), gaze duration and number of fixations in relation to object choices for males and females (Semin et al., 2018), and now perceptual accentuation. Thus, the studies in combination illuminate the grounding of an abstract category and its cognitive implications.

Funding

The second author gratefully acknowledges the financial support provided by the Portuguese Foundation for Science and Technology [grant ref: IF/00085/2013/CP1186/CT0001].

Declaration of Competing Interest

None.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2019.05.009.

References

Alter, A. L., Stern, C., Granot, Y., & Balcetis, E. (2016). The "bad is black" effect: Why people believe evildoers have darker skin than do-gooders. *Personality and Social Psychology Bulletin, 42*(12), 1653–1665. https://doi.org/10.1177/0146167216669123.

Aoki, K. (2002). Sexual selection as a cause of human skin colour variation: Darwin's hypothesis revisited. *Annals of Human Biology, 29*(6), 589–608. https://doi.org/10.1080/03014460210000019144.

Boroditsky, L. (2001). Does language shape thought?: Mandarin and English speakers' conceptions of time. *Cognitive Psychology, 43*, 1–22. https://doi.org/10.1006/cogp.2001.0748.

Carrito, M. L., Santos, I. M. B., Lefevre, C. E., Whitehead, R. D., Silva, C. F., & Perrett, D. I. (2016). The role of sexually dimorphic skin colour and shape in attractiveness of male faces. *Evolution and Human Behavior, 37*(2), 125–133. https://doi.org/10.1016/j.evolhumbehav.2015.09.006.

Chaplin, G. (2004). Geographic distribution of environmental factors influencing human skin coloration. *American Journal of Physical Anthropology, 125*(3), 292–302. https://doi.org/10.1002/ajpa.10263.

Corneille, O., Huart, J., Becquart, E., & Brédart, S. (2004). When memory shifts toward more typical category exemplars: Accentuation effects in the recollection of ethnically ambiguous faces. *Journal of Personality and Social Psychology, 86*(2), 236–250. https://doi.org/10.1037/0022-3514.86.2.236.

Corneille, O., Klein, O., Lambert, S., & Judd, C. M. (2002). On the role of familiarity with units of measurement in categorical accentuation: Tajfel and Wilkes (1963) revisited and replicated. *Psychological Science, 13*(4), 380–383. https://doi.org/10.1111/1467-9280.00468.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology: General, 122*(3), 371–396. https://doi.org/10.1037/0096-3445.122.3.371.

Hansen, T., Olkkonen, M., Walter, S., & Gegenfurtner, K. R. (2006). Memory modulates color appearance. *Nature Neuroscience, 9*, 1367. https://doi.org/10.1038/nn1794.

Huart, J., Corneille, O., & Becquart, E. (2005). Face-based categorization, context-based categorization, and distortions in the recollection of gender ambiguous faces. *Journal of Experimental Social Psychology, 41*(6), 598–608. https://doi.org/10.1016/j.jesp.2004.10.007.

Ihara, Y., & Aoki, K. (1999). Sexual selection by male choice in monogamous and polygynous human populations. *Theoretical Population Biology, 55*, 777–793. https://doi.org/10.1006/tpbi.1998.1388.

Jablonski, N. G. (2004). The evolution of human skin and skin color. *Annual Review of Anthropology, 33*(1), 585–623. https://doi.org/10.1146/annurev.anthro.33.070203.143955.

Jablonski, N. G., & Chaplin, G. (2000). The evolution of human skin coloration. *Journal of Human Evolution, 39*, 57–106. https://doi.org/10.1006/jhev.2000.0403.

Jablonski, N. G., & Chaplin, G. (2002). Skin deep. *Scientific American, 287*(4), 74–81.

Jablonski, N. G., & Chaplin, G. (2013). Epidermal pigmentation in the human lineage is an adaptation to ultraviolet radiation. *Journal of Human Evolution, 65*, 671–675. https://doi.org/10.1016/j.jhevol.2013.06.004.

Jablonski, N. G., & Chaplin, G. (2018). The roles of vitamin D and cutaneous vitamin D production in human evolution and health. *International Journal of Paleopathology, 23*, 54–59. https://doi.org/10.1016/J.IJPP.2018.01.005.

Kieruj, N. D., & Moors, G. (2010). Variations in response style behavior by response scale format in attitude research. *International Journal of Public Opinion Research, 22*(3), 320–342. https://doi.org/10.1093/ijpor/edq001.

Lakens, D., Semin, G. R., & Foroni, F. (2012). But for the bad, there would not be good: Grounding valence in brightness through shared relational structures. *Journal of Experimental Psychology: General, 141*(3), 584–594. https://doi.org/10.1037/a0026468.

Lee, K. J., & Perrett, D. I. (2000). Manipulation of colour and shape information and its consequence upon recognition and best-likeness judgments. *Perception, 29*, 1291–1312. https://doi.org/10.1068/p2792.

Lefevre, C. E., & Perrett, D. I. (2015). Fruit over sunbed: Carotenoid skin colouration is found more attractive than melanin colouration. *The Quarterly Journal of Experimental Psychology, 68*(2), 284–293. https://doi.org/10.1080/17470218.2014.944194.

Levin, D. T., & Banaji, M. R. (2006). Distortions in the perceived lightness of faces: The role of race categories. *Journal of Experimental Psychology: General, 135*(4), 501–512. https://doi.org/10.1037/0096-3445.135.4.501.

Majid, A., Gaby, A., & Boroditsky, L. (2013). Time in terms of space. *Frontiers in Psychology, 4*, 554. https://doi.org/10.3389/fpsyg.2013.00554.

Meier, B. P., Fetterman, A. K., & Robinson, M. D. (2015). Black and white as valence cues: A large-scale replication effort of Meier, Robinson, and Clore (2004). *Social Psychology, 46*(3), 174–178. https://doi.org/10.1027/1864-9335/a000236.

Meier, B. P., & Robinson, M. D. (2004). Why the sunny side is up: Associations between affect and vertical position. *Psychological Science, 15*(1), 243–247.

Meier, B. P., Robinson, M. D., & Clore, G. L. (2004). Why good guys wear white. *Psychological Science, 15*(2), 82–87. https://doi.org/10.1111/j.0963-7214.2004.01502002.x.

Roediger, H. L. (1990). Implicit memory: Retention without remembering. *American Psychologist, 45*(9), 1043–1056. https://doi.org/10.1037/0003-066X.45.9.1043.

Russell, R. (2009). A sex difference in facial contrast and its exaggeration by cosmetics. *Perception, 38*(8), 1211–1219. https://doi.org/10.1068/p6331.

Russell, R., Kramer, S. S., & Jones, A. L. (2017). Facial contrast declines with age but remains sexually dimorphic throughout adulthood. *Adaptive Human Behavior and Physiology, 1–11*. https://doi.org/10.1007/s40750-017-0068-x.

Schacter, D. L. (1987). Implicit memory: History and current status. *Journal of Experimental Psychology Learning Memory and Cognition, 13*(3), 501–518. https://doi.org/10.1037/0278-7393.13.3.501.

Schubert, T. W. (2005). Your highness: Vertical positions as perceptual symbols of power. *Journal of Personality and Social Psychology, 89*(1), 1–21. https://doi.org/10.1037/0022-3514.89.1.1.

Sebastián-Enesco, C., & Semin, G. R. (2019). The brightness dimension as a marker of gender across cultures and age (submitted for Publication).

Semin, G. R., & Palma, T. A. (2014). Why the bride wears white: Grounding gender with brightness ☆. *Journal of Consumer Psychology, 24*(2), 217–225. https://doi.org/10.1016/j.jcps.2013.09.003.

Semin, G. R., Palma, T. A., Acartürk, C., & Dziuba, A. (2018). Gender is not simply a matter of black and white, or is it? *Phil. Trans. R. Soc. B, 373*, 20170126. https://doi.org/10.1098/rstb.2017.0126.

Specker, E., Leder, H., Rosenberg, R., Hegelmaier, L. M., Brinkmann, H., Mikuni, J., & Kawabata, H. (2018). The universal and automatic association between brightness and positivity. *Acta Psychologica, 186*, 47–53. https://doi.org/10.1016/J.ACTPSY.2018.04.007.

Tajfel, H., & Wilkes, A. L. (1963). Classification and quantitative judgement. *British*

Journal of Psychology, 54(2), 101–114. https://doi.org/10.1111/j.2044-8295.1963.tb00865.x.

Tiddeman, B., Burt, M., & Perrett, D. (2001). Prototyping and transforming facial textures for perception research. *Computer Graphics and Applications, IEEE, 21*, 42–50. https://doi.org/10.1109/38.946630.

Van den Berghe, P. L., & Frost, P. (1986). Skin color preference, sexual dimorphism and sexual selection: A case of gene culture co-evolution?*. *Ethnic and Racial Studies, 9*, 87–113.

Walter, H. (1958). Der zusammenhang von hautfarbenverteilung und intensitat der ultravioletten strahlung. *Homo, 9*, 1–13.

Wetzel, E., Böhnke, J. R., & Brown, A. (2016). Response biases. *The ITC international handbook of testing and assessment* (pp. 349–363). New York, NY, US: Oxford University Press.

Whitehead, R. D., Re, D., Xiao, D., Ozakinci, G., & Perrett, D. I. (2012). You are what you eat: Within-subject increases in fruit and vegetable consumption confer beneficial skin-color changes. *PloS One, 7*, e32988. https://doi.org/10.1371/journal.pone.0032988.