CITATION

Comfort, W. E., de Andrade, B. N., Wingenbach, T. S. H., Causeur, D., & Boggio, P. S. (2021). Implicit responses in the judgment of attractiveness in faces with differing levels of makeup. *Psychology of Aesthetics, Creativity, and the Arts.* Advance online publication. https://doi.org/10.1037/aca0000408

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

Abstract

Makeup is a form of body art which has been used for over 7000 years and is present in the great majority of human cultures, often used to enhance facial attractiveness and to accentuate features that represent femininity. This study examined how cumulative levels of facial makeup influenced approach and avoidance tendencies and on facial muscle responses associated with emotional response obtained through facial electromyography (EMG) in a passive viewing task. Experiment 1 employed the joystick variant of the approach-avoidance task, where 30 subjects categorised female faces by visual orientation (portrait/landscape) in 7 cumulatively-added makeup levels. In Experiment 2, facial EMG was recorded from 40 subjects in the passive viewing of the same images. The present study shows that makeup application modulates implicit responses and reveals two distinct implicit preferences, behavioural and affective, with a male behavioural preference for heavy eye cosmetics, a female behavioural preference for light makeup, and an overall affective preference in both men and women for makeup accentuating visual contrast in the eye and mouth regions. These results are consistent with the conception that perceptual cues underlying cosmetic enhancement are key determinants in aesthetic facial preferences.

Keywords: cosmetics, facial attractiveness, facial electromyography, approach-avoidance task.

Abstract word count: 190

Text word count: 7977

Introduction

Judgements of facial attractiveness have been shown to be remarkably consistent between individuals and cultures (Langlois et al., 2000), in marked contradiction to the commonly-held belief that "beauty is in the eye of the beholder". Recent studies have shed further light on the perceptual bases of facial attractiveness, by using techniques traditionally employed in studies of the recognition of facial expressions to further elucidate common perceptual cues for the evaluation of facial attractiveness, moving beyond the concepts laid out in human ethology and evolutionary psychology of facial symmetry, averageness and skin texture as the principal determinants of facial attractiveness in humans (Fink & Neave, 2005; Rhodes, 2006; Little et al., 2011).

One key area of interest is the extent to which the application of cosmetics enhances female facial attractiveness in the face of several conflicting claims from the health and beauty industry. Female faces are judged to be significantly more attractive following the application of differing layers of eye, lip and full-face makeup (Mulhern et al., 2003), and this effect is enhanced for female observers compared to males. Mulhern et al. (2003) investigated the effect of makeup on female facial attractiveness evaluated by male and female participants using five cosmetic conditions: no makeup, foundation only, eye makeup only, lip makeup only, and full facial makeup. Their results showed that in explicit ratings of attractiveness, faces with full makeup were rated as more attractive than those with no makeup or less makeup (e.g. lipstick only). While their design allowed for the evaluation of cosmetic enhancement by region, in naturalistic settings women more commonly combine cosmetic products applied to the face as a whole. As such, both the number and combination of makeup products used should be addressed when investigating the cosmetic enhancement of facial attractiveness.

In addition, women judged to be more attractive when wearing makeup are also perceived as healthier, more confident and more professionally successful by male and female participants (Nash et al., 2006), as part of a generalised "attractiveness halo" effect for attractive faces (for a review, see Zebrowitz & Montepare, 2008). However, differences in judgements of facial attractiveness due to makeup are statistically negligible in comparison to differences due to identity (Jones & Kramer, 2016) and thus the enhancement effect of makeup on overall attractiveness appears to be slight and may potentially be due to an interaction with individual physiognomy.

Facial attractiveness is highly correlated with femininity or sexual dimorphism (Koehler et al., 2004; Little et al., 2011). Perceptual analyses of female and male faces have highlighted differences in luminance contrast in facial regions such as the eyes and mouth as one of the prime determinants of sexual dimorphism (judging whether a face is more feminine or masculine) regardless of face gender (Russell, 2003; 2009). Accordingly, makeup appears to lead to an enhancement of this contrast effect with products such as eyeshadow that accentuate the femininity of the face. At the same time, products such as foundation are used to mask imperfections and smooth skin texture, leading to an increase in overall facial symmetry and averageness which are seen as common evolutionary-defined cues for assessing attractiveness (Rhodes, 2006; Russell, 2010). This is consistent with attractiveness research, which has highlighted specific facial characteristics as determinants of female facial attractiveness beyond facial symmetry, such as high cheekbones, large eyes and lips, thin eyebrows, and small noses and chins (Cunningham et al., 1995; Baudouin & Tiberghien, 2004).

Previous studies have shown a remarkable consistency in male preferences for female facial attractiveness across different races and cultures (Cunningham et al., 1995; Thornhill & Grammer, 1999; Fink & Penton-Voak, 2002). One of the main divergences in attractiveness preferences across race is related to the degree of sexual dimorphism that female faces display (Penton-Voak et al., 2004). White British and Japanese men show significant preference for more feminine female faces of the same race, while Jamaican men show less preference for sexual dimorphism, evaluating female faces with greater masculinity as significantly more attractive than feminised morphs (Penton-Voak et al., 2004). Overall however, increased sexual dimorphism is preferred in female but not male faces (Morrison et al., 2010), primarily as an indicator of health and increased fertility (Law-Smith et al., 2006).

The majority of studies conducted to date on the effect of makeup on judgements of facial attractiveness have employed an overt rating scale, with participants assigning explicit scores to faces on the basis of their conscious perception of the individual's attractiveness, similar to the design by Osborn (1996). There have been relatively few studies conducted to date measuring implicit responses to faces with differing levels of makeup. One of the few studies to do so (Richetin et al., 2004), used the Implicit Association Test to test differences in reaction time in response to pairings of female faces with and without makeup and positive and negative stimuli such as personality traits, pleasant/unpleasant words and professions of high-/low-social status. They found that faces with makeup were associated with positive personality traits and high-status professions more than faces with no makeup, similarly to the results from Nash et al. (2006). As makeup had no effect on reaction time in response to pleasant and unpleasant words, the implicit processing of makeup may be dependent on social context, and not merely affected by the emotional valence of the stimuli.

Another method for measuring implicit response, traditionally in the context of emotional valence, is facial electromyography (EMG). Facial EMG is capable of delivering great sensitivity and accuracy in the detection of the movement of facial muscles associated with emotional expressions such as the *M. corrugator supercilii* (associated with frowning and negative affect) and the *M. zygomaticus major* (associated with smiling and positive affect). Facial EMG can detect face muscle activations that are so subtle that they are not visible in the face due to the overlaying fatty tissue and skin (Rinn, 1984). In addition, recordings from facial EMG can capture responses to low-intensity emotional stimuli and even in situations where the participant has no conscious awareness of producing an emotional response (Cacioppo et al., 1986; Dimberg et al., 2000). An example of face muscle activity that can be measured using EMG despite a lack of participants' awareness of the muscle activations is the phenomenon termed 'facial mimicry' (see a review by Hess & Fischer, 2014). That is, presenting participants with stimuli portraying facial emotional expressions on a computer screen will produce only a subliminal perception of muscle activation in the participant in accordance to the observed facial expression. Facial mimicry has been demonstrated in the corrugator when participants observe emotional expressions of negative valence and in the zygomaticus when observing positively valenced emotional expressions (Achaibou et al., 2008; Dimberg, 1982; Dimberg & Thunberg, 1998; Lundqvist, 1995; Lundqvist & Dimberg, 1995), and increased levator activity when observing facial expressions of disgust and greater frontalis activation when observing facial expression of fear and surprise (Lundqvist, 1995; Lundqvist & Dimberg, 1995). That facial muscle activity associated with emotional facial expressions can be measured without participants being aware of these activations makes facial EMG an ideal implicit measure. Facial EMG holds great utility as an implicit measure of individual affect, since face muscle activations (as measured via EMG) can reflect underlying emotional states. For example, when watching pictures related to positive and negative emotions, participants' EMG activity will increase in the zygomatic and corrugator regions respectively, and subjective affective valence ratings are in line with these increased muscle activations (e.g. Larsen et al., 2003). Facial EMG thus allows to measure participants' affective responses independent of participants awareness of their emotional states or accompanying facial muscle activations.

Employing facial EMG in response to faces of varying facial attractiveness results in a modulation of activity mainly in the zygomatic and corrugator regions (Hazlett & Hoehn-Saric, 2000; Gerger et al., 2011) and over the *levator labii superioris* muscle (Principe & Langlois, 2011), associated with emotional reactions of disgust. Hazlett and Hoehn-Saric (2000) found an interesting sex difference in facial EMG response to facial attractiveness, with female subjects revealing increased corrugator response when presented with highly-attractive female faces and greater zygomatic response when viewing highly-attractive male faces. Overall however, there appears to be a linear negative correlation between facial attractiveness and corrugator and levator response (Principe & Langlois, 2011), and, to a lesser extent, a positive correlation between attractiveness and zygomatic response (Gerger et al., 2011), regardless of the gender of the subject. Facial EMG modulation has been shown to be consistent with explicit ratings of facial attractiveness (Gerger et al., 2011) and sexual arousal (Hazlett & Hoehn-Saric, 2000), but has not yet been compared to other measurements of implicit response, such as approach/avoidance behaviour. Though, facial EMG might reveal affective responses beyond what can be assessed with explicit measures. Another more recent study (Tagai et al., 2017) investigated the effect of different levels of makeup on amplitude differences of ERP components associated with face processing such as the N170 and VPP. The authors observed that the processing of faces with light makeup was accompanied by a decrease in N170 and VPP amplitude as compared to faces with heavy makeup. This result was consistent with the explicit classification of facial attractiveness, with slightly softer faces being evaluated as more attractive than faces with heavy makeup, possibly due to the greater fluency and ease of visual processing of faces with lighter makeup.

Several studies have investigated trustworthiness and facial emotion through an approach/avoidance paradigm, whether through measurements of amygdalar activation (Todorov et al., 2008) or manipulation of a virtual manikin or physical joystick (Heuer et al., 2007; Krieglmeyer & Deutsch, 2010). As facial attractiveness and approachability/trustworthiness are highly correlated (Todorov, 2008; Sofer et al., 2015), additional feedback from an approach/avoidance measure such as joystick position may provide useful data in response to emotional modulation by both attractiveness and makeup levels. Concomitant effects of facial attractiveness on emotional processing have long been established (Nakamura et al., 1998); several areas associated with reward and positive-valenced emotions, such as the orbitofrontal cortex and amygdala, are also activated when viewing and categorising faces by level of attractiveness (Winston et al., 2007). In addition, this activation frequently occurs even when the task is not specific to categorisation of facial attractiveness (Chatterjee et al., 2009), indicating that the reward-inducing properties of attractive faces are at least partly automatised. More recently, approach-related behaviour has been directly linked with the reward value of faces explicitly categorised as more attractive in both male and female participants (Kramer et al., 2020), with greater physical “lean” and approach response towards attractive than unattractive female faces even in the absence of active task demands. As emotionally-expressive facial cues have been shown to lead to the modulation of task-selective motor response, the authors argue for a similar modulatory effect from facial cues signalling attractiveness (Kramer et al., 2020).

This close correlation between beauty and emotional and behavioural response opens up several possibilities for testing the perception of facial attractiveness using implicit measures more commonly used for the analysis of affective valence or intensity (Chatterjee & Vartanian, 2016). Interestingly, an interaction between positive emotional feedback and visual fluency may lie behind one of the main determinants of facial attractiveness, prototypicality or averageness. Winkielman et al. (2006) found that more prototypical random-dot patterns were categorised more quickly and consistently rated as more attractive than less prototypical displays, together with increased zygomatic EMG response, revealing the close association between the increased perceptual fluency of prototypical stimuli and higher measures of attractiveness and positive affect. Principe and Langlois (2012) investigated the effect of face prototypicality on emotional response when categorising faces by attractiveness, and found that previous familiarisation with human-chimpanzee morphed faces led to a shift in preferences; with human-chimpanzee morphs categorised as more attractive and with a correspondent increase in zygomatic activity in those participants than for participants who received no previous familiarisation. This reveals that our internal prototypes for facial attractiveness are both malleable and subject to previous cultural experience.

Taken together, these results reveal a diverse set of social and cognitive mechanisms underlying the perception of makeup, ranging from a generalised “halo effect”, with more heavily-applied cosmetics associated with greater professional success, competence, and even physical health and wellbeing (Nash et al., 2006; Richetin et al., 2004), to visual cues signalling femininity, youth and attractiveness, particularly a skin-smoothing effect of foundation (Russell, 2010) and an increase in luminance contrast provided by eye makeup (Russell, 2003, 2009) associated with sexual dimorphism. Some researchers have argued for the integration of these components as part of an 'extended phenotype' of cosmetic use, as a cultural tool to increase one's social and sexual success (Etcoff et al., 2011; Mileva et al., 2016).

Present study: The implicit techniques of facial EMG and AAT response were employed as dependent measures in the present study to investigate the effect of cumulative levels of applied makeup on participants' implicit emotional and behavioural responses to facial attractiveness and ethnicity while conducting a perceptual categorisation task. We sought to investigate the implicit perception of makeup by testing participants' responses to varying conditions of makeup, from a basic layer of foundation to the "heavier" application of eyeshadow. As makeup is commonly applied in different stages in response to social context, with foundation and lipstick used in more "everyday" contexts than other products, we designed a set of facial stimuli containing the cumulative addition of cosmetic products from a base of foundation and lipstick to the greater visual contrast of pencilling, mascara, eyeliner and eyeshadow respectively. The inclusion of gradually-applied makeup levels in the stimuli also sought to differentiate participants' implicit response to the qualitative changes to facial features and configuration caused by different makeup products, within a naturalistic setting. Previous studies investigating implicit or physiological responses to makeup differences have employed either a no-makeup/makeup design (Richetin et al., 2004) or a no/light/heavy makeup design (Tagai et al., 2016; 2017), which may not have shown sufficient sensitivity towards differences in intermediate levels of makeup application.

As EMG markers associated with negative emotional response such as corrugator and levator activation have previously shown to be negatively correlated with explicit face attractiveness ratings (Principe & Langlois, 2011), and conversely zygomatic activation, associated with positive affect, has been shown to be positively correlated with explicit face attractiveness ratings (Gerger et al., 2011), we opted to include these specific muscle sites in the design of the current study. Additionally, we included the frontalis muscle (*M. Frontalis, pars lateralis*) as a site of EMG response, due to the previously reported association of this muscle with the inducement of stress (Kukde & Neufeld, 1994), and negative affect (Cacioppo et al., 1986) similarly to the corrugator. We hypothesised that the increased cosmetic enhancement of facial attractiveness would lead to a decrease in activation at the three muscle sites associated with negative affect and emotional response (corrugator, levator and frontalis), and a concurrent increase in zygomatic activation, associated with positive affect. In addition, in line with studies reporting a perceptual preference for increased visual fluency in faces with light compared to heavy makeup (Tagai et al., 2016; 2017), we expected to observe a drop-off in zygomatic activation following intermediate levels of makeup application, and an increase in corrugator activation in response to heavy levels of makeup application, consistent with previously-reported visual fluency effects in aesthetic preference (Gerger & Leder, 2015).

The previously-reported effect of the cosmetic enhancement of facial attractiveness in an implicit experimental design (IAT; Richetin et al., 2004) and a recent study highlighting the utility of the approach-avoidance task (AAT) as an implicit measure of facial attractiveness (Kramer et al., 2020), led us to include AAT response as a dependent variable in our study. That is, beyond an affective preference for certain makeup products or combinations of such products, we tested a behavioural preference for cosmetic products in terms of participants' reaction time towards engaging more closely or more distally with cosmetically-enhanced facial stimuli. Similarly to the EMG response, we expected to observe a faster approach and slower avoidance time to intermediate levels of makeup application, and a slower approach and faster avoidance time to both no makeup and heavy levels of makeup application, driven primarily by visual fluency effects.

Experiment 1 - Methods

Subjects

The sample of the approach/avoidance task in Experiment 1 was composed of 15 women and 15 men, heterosexual, between 19 and 27 years (M: 21.77, SD: 2.743) and of Caucasian ethnicity. Sample size and composition were calculated based on the effect sizes reported in previous related research studies (Tagai et al., 2016; 2017; N = 38-45, $\eta^2 \approx 0.3$), using G*Power 3.1.9.2 (Faul et al., 2007). All participants were informed about the procedure but not informed about the specific objective in the study (approach-avoidance response to different cosmetics), and signed an informed consent form indicating their willingness to participate in the experiment. Participant recruitment took place through digital media such as social media and scientific research recruitment sites. Participants enrolled as students received course credit for their participation in the experiment. The study adhered to the Declaration of Helsinki guidelines and was approved by the institutional ethics committee and national ethics committee.

Procedure

Stimuli

The experimental stimuli were composed of 126 emotionally-neutral images of 18 female faces with 7 different levels of makeup (no makeup, added foundation, added lipstick, added mascara, added pencilling, added eyeliner, added eyeshadow) from a previously-constructed face database. The database used in the current study is composed of facial photographs taken of 60 women aged 19 – 32, in a frontal pose, and of three distinct ethnicities as identified by their self-classification on the electoral roll as of Asian, Caucasian or African descent (20 faces for each ethnicity). All face models were recruited through digital media such as social media and scientific research recruitment sites, and models enrolled as students received course credit for their participation in the study. The models were photographed in a frontal pose following the application of makeup by a professional makeup artist. Makeup was applied in a standardised manner for an "everyday" setting, with the first 5 levels (no makeup – pencilling) corresponding to daytime use and the last 2 levels (eyeliner – eyeshadow) corresponding to nighttime use, to closely mirror makeup use in naturalistic settings (see Figure 1 for an example). All face images in the database were feature-aligned and digitally standardised for luminance, visual contrast and visual spatial frequency. One hundred and twenty-six images of 18 individuals were selected from the database following image processing and standardisation. In addition, the images selected for the study showed a linear increase in perceived lightness (as measured through HSV and CIELAB) following each successive stage of application, with an additional increase in Global Contrast Factor (GCF; Matkovic et al., 2005) on the last four levels, consistent with the makeup "looks" employed in Etcoff et al. (2011).

FIGURE 1 GOES HERE

The images selected for the current study were tested in an online validation task whereby each face was rated according to emotional valence (1: negative valence – 7: positive valence) and facial attractiveness (1: very unattractive – 7: very attractive). In addition, participants were required to indicate whether they had previously met or knew the person shown in the task. Twelve participants (6 male/female, mean age: 23.14, SD: 2.47) completed the validation task on an online research platform (Google Forms), with no participants indicating they were familiar with the identities presented. A repeated-measures ANOVA revealed no significant differences in median facial attractiveness or emotional valence scores between the three ethnicity groups (Asian, Caucasian, African descent), with mean emotional valence scores ranging between 3 and 5 on the rating scale.

Approach / Avoidance Task (AAT)

The approach-avoidance task used in the present study was based on the Approach-Avoidance Task (AAT) used in the study by Wiers et al. (2009). The task version was designed and executed using the Inquisit psychological research software (Millisecond, Inc.).

Images of 18 facial identities and 7 cumulative makeup levels were presented in both vertical (portrait) and horizontal (landscape) orientations, totalling 252 images. Initial portrait resolution was 1500 x 2000 pixels while initial landscape resolution was 2000 x 1500 pixels, measuring approx. 168° of visual angle. All facial stimuli were unframed and presented against a grey background. Participants were instructed to maintain their attention in the centre of the screen and to move the joystick forwards or backwards according to the image orientation, with the movement assigned to either orientation counterbalanced between participants. The image size increased or decreased according to the extension of the joystick (Thrustmaster® PC USB) in a backwards and forwards direction respectively, up to a maximum increase or decrease of 70% percent of the original image size (see Rinck & Becker, 2007, for a more detailed technical description). That is, pulling the joystick towards the participant resulted in a continuous increase in image size, with a maximum increase of 70% of the original image size, while pushing the joystick away from the participant resulted in a continuous decrease in image size, with a maximum decrease of 70% of the original image size. Each image was presented four times over the course of four blocks, with a total of 1008 trials (144 trials per makeup condition), and an equal number of portrait/landscape presentations. Each image stayed on screen until the joystick was fully extended in either direction, and the next trial was initiated. The reaction time on each trial was calculated as the difference between onset of stimulus presentation and the terminus of joystick extension, to ensure a standardised response for all participants, as image contraction/inflation is also an exteroceptive cue of approach/avoidance (Wiers et al., 2009). The order of image presentation was randomised with no replacement.

Explicit Rating Task

Immediately following the completion of the AAT, all participants were instructed to rate the images shown in the AAT rated according to emotional valence (1: negative valence – 7: positive valence) and facial attractiveness (1: very unattractive – 7: very attractive). In addition, participants were required to indicate whether they had previously met or knew the person shown in the task, with no participants indicating they were familiar with the identities presented.

AAT Pre-Processing

The initial phase of data analysis consisted in excluding trials containing incorrect responses. Only 3.12% of trials contained an incorrect response and means comparisons revealed no significant differences in error rate between the experimental conditions of makeup level and participant gender. Next, AAT difference scores were calculated from the subtraction of the median approach value (pulling the joystick) from the median avoidance value (pushing the joystick) for each image (see Table 1 in the Supplement). Thus, positive values correspond to a faster approach time and slower avoidance time, values close to zero correspond to equal speeds of approach and avoidance, and a negative index corresponds to faster avoidance times and slower approach times. The mean AAT score for each face ethnicity was then computed for every participant. Two experimental factors were examined: makeup level and gender of the participant.

To compare the indices between the different levels of makeup, a mixed ANOVA analysis was performed, with participant gender as a between-subjects factor, and makeup level as a within-subjects factor.

Results – Experiment 1

Explicit Rating Task

A repeated-measures ANOVA revealed a significant main effect of makeup on median attractiveness scores ($F\ (6, 23) = 6.476$, $p < .001$, partial $\eta^2 = .188$), with Bonferroni post-hoc testing revealing significantly lower scores in response to M1 than all other makeup levels ($p < .05$), with no other significant differences between makeup levels (*p*

= n.s.). No significant main effect of gender on median attractiveness (F (1, 28) = 1.679, *p* = n.s.) or emotional valence (F (1, 28) = 1.082, *p* = n.s.) was observed, and no significant main effect of makeup level on emotional valence (F (6, 23) = 1.585, *p* = n.s.) was observed (see Figure 2 for details).

INSERT FIGURE 2 HERE

AAT Difference Scores

A repeated-measures ANOVA was conducted to test the effect of two factors (makeup level and participant gender) on the AAT difference scores computed from the subtraction of the avoidance response by the approach response. The analysis revealed a significant main effect of gender (F (1, 41) = 8.233, *p* = .005, partial η^2 = .086), with a significantly higher AAT value for the female group (M = 15.82, SD = 4.649) as compared to the male group (M = -3.046, SD = 4.649), indicating a faster approach time and a slower avoidance time in response to all images, and a significant interaction between gender and makeup level (F (6, 41) = 2.299, *p* = .034, partial η^2 = .025). Simple effects testing Bonferroni-corrected for multiple comparisons revealed a significantly higher female response at level M2 (M = 32.98) compared to M5 (M = -0.667; *p* = .017) and M7 (M = 8.4; *p* = .027), and significantly higher male responses at level M1 (M = 5.389) compared to M2 (M = -14.62; *p* = .038) and M4 (M = -18.26; *p* = .032), at level M6 (M = 8.744) compared to M2 (*p* = .027) and M4 (*p* = .009), and at level M7 (M = 13.23) compared to M2 (*p* = .008), M3 (M = -8.033; *p* = .029), M4 (*p* = .003) and M5 (M = -7.778; *p* = .030), indicating a female preference for light compared to heavy makeup and a male behavioural preference for no or heavy makeup as opposed to medium makeup. In addition, there was a near-significant main effect of makeup (F (6, 41) = 2.108, p = .051). See Figure 3 below for more details.

INSERT FIGURE 3 HERE

Experiment 2 - Methods

Subjects

In Experiment 2 facial EMG recordings were collected for 40 participants, 20 men and 20 women, heterosexual, aged 20 to 26 (M: 22.72, SD: 3.879) and of Caucasian ethnicity. Sample size and composition were calculated based on the effect sizes reported in previous related research studies (Tagai et al., 2016; 2017; N = 38-45, $\eta^2 \approx 0.3$), using G*Power 3.1.9.2 (Faul et al., 2007). All participants were informed about the procedure but not informed about the specific objective in the study. Furthermore, deception was employed, by instructing participants the EMG recording device measured skin conductance response (SCR), not muscle activity, to prevent participants from modulating their facial expressions. Participants signed an informed consent form indicating their willingness to participate in the experiment. Participant recruitment took place through digital media such as social media and scientific research recruitment sites. Participants enrolled as students received course credit for their participation in the experiment. The study adhered to the Declaration of Helsinki guidelines and was approved by the institutional ethics committee and national ethics committee.

Procedure

Stimuli

The experimental stimuli were composed of 108 emotionally-neutral images of 18 female faces with 6 different levels of makeup (no makeup, added foundation, added lipstick, added eyebrow pencil and mascara, added eyeliner, and added eyeshadow) from a previously-constructed face database. In this experiment only 6 separate levels were used: 1: no makeup; 2: added foundation; 3: added lipstick; 4: added eyebrow pencil and mascara; 5: added eyeliner; 6: added eyeshadow, to reduce the total number of trials, as no significant differences were reported between the M4 (eyebrow pencil) and M5 (mascara) levels used in Experiment 1, and both products serve as similar perceptual cues (cues signaling higher visual contrast in eye region).

Facial Electromyography (EMG) Recording

Psychophysiological data was collected through surface EMG recording with four shielded electrode pairs to measure voltage changes linked to muscle activity while participants passively viewed images of female faces with different cumulative levels of makeup, composed of the same images used in the first experiment. The passive viewing task was designed and executed using the E-Prime 2.0 psychological presentation software (Psychology Software Tools, Inc., Pittsburgh, PA). After placing all electrodes on the left side of the face, participants were instructed to maintain their attention in the centre of the screen and passively view images of female faces while recording the EMG signal, with each stimulus level corresponding to a specific marker in the EMG signal. In addition, instructions were given to maintain a relaxed and still posture so as to minimise interference with the recording device, which participants were deceptively informed was for monitoring their skin conductance response. The experiment consisted of 3 blocks, with a total of 324 trials (54 trials per makeup condition). Participants were instructed to take a 5-minute break in between each block, with a total experiment time of approximately 50 minutes. Each image was presented for 2000ms, with an inter-stimulus interval (ISI) of 2000ms, and a 500ms pre-stimulus baseline containing a fixation cross. Prior to the participant debriefing, participants were asked what they perceived to be the objectives of the experiment, with no participant correctly identifying the objective.

Data was recorded using the BIOPAC MP150 system with Acqknowledge software (Version 4, Biopac Systems, Inc., Goleta, CA) and a separate EMG110C unit for each of the four facial muscles sires recorded with the current study. The electrodes were positioned over the following muscle sites, according to the guidelines of Fridlund and Cacioppo (1986): corrugator (Corrugador supercilii); zygomatic (Zygomaticus major); levator (Levator labii superiors); and frontalis (Frontalis, pars lateralis). Silver-silver chloride (Ag-AgCl) shielded surface electrode pairs (EL254S) filled with conductive gel (Signa Gel with saline solution) with a contact area of 4mm diameter were used. EMG amplifiers were set to a gain of 2000 and real-time data filtering was conducted through a bandwidth with lower and upper thresholds of 10 Hz and 500 Hz, respectively. Grounding was performed through an additional electrode placed in the middle of the forehead. The sampling rate was held constant at 1000 Hz throughout the experiment. Prior to electrode placement, the surface area of participant's face was wiped with cotton wool and an ethanol solution to remove excess oils and dead skin and thus secure electrode attachment with double-stick adhesive rings. During the task, the experimenter observed the participant through a webcam (recording offline) placed above the monitor, and documented any instances of movement such as coughing or sneezing, for later removal of experimental artefacts in the EMG data.

Facial EMG Pre-Processing

EMG data preparation was conducted with a custom-made MATLAB script (Mathworks, Natick, MA). First, artifacts were removed according to the documentation during data collection by excluding artefactitious data segments per participant in the respective channel. The EMG data was then filtered with a 28 Hz high-pass filter, rectified and smoothed with a moving average of 50 ms. A total of 2.179 %, 4.362 %, 4.341 % and 2.685 % of trials were excluded from the data recorded at the corrugator, zygomatic, levator and frontalis respectively. Each trial was segmented in 100 ms bins resulting in a 500 ms initial baseline period, a 2000 ms period corresponding to stimulus presentation, and a 2000 ms interstimulus interval (ISI). Further data preparation was conducted in Excel (Microsoft Office, Microsoft, Inc.). A spike filter was applied to the EMG data defined by a deviation of +/- 3 SD of the total mean from one bin to the next. All trials that exceeded this definition were winsorized, such that extreme values were set to the next-highest value, as described by Field (2013). To compare the mixed factors of makeup condition (within-subjects) and participant gender (between-subjects), the bins from each trial were z-standardised according to the participant mean, as a secondary dataset. The bins from the two datasets (EMG values and z-scores) were then baseline-corrected, subtracting the mean value from the baseline period. Statistical analysis of the within-subjects factor of makeup condition was conducted using the means across participants for each makeup condition (M1 – M7). An R script (R-Project) was used to conduct functional ANOVA (FANOVA) analyses on the EMG response observed during stimulus presentation, separately for each muscle site (corrugator, zygomatic, levator and frontalis).

Functional ANOVA applies the assumptions of analysis of variance to functional observations that, while discrete to specific timepoints in the data, are sampled frequently over a defined period (Ramsay & Silverman, 2005). FANOVA was employed in this case for its utility in analysing the time course of facial EMG response to a complex visual stimulus such as a cosmetically-enhanced face, with a Type-II sum of squares for testing main effects and interactions (Langsrud, 2003). In addition, a functional generalized F-test designed for electrophysiological data analysis was employed whereby exact F statistics and p-values are estimated using Monte Carlo simulation (Causeur et al., 2019b). Data from all time periods (baseline, stimulus presentation, ISI) was included in the analysis but only data from the period of stimulus presentation was included for the purpose of significance testing.

Following bandpass filtering and baseline correction, the mean facial EMG response at the four muscle sites: corrugator, zygomatic, levator and frontalis, was calculated for each 100 ms bin including the baseline, stimulus presentation time and ISI. The within-subjects analysis of makeup level was plotted against mean EMG response (μV) and time (ms), while the mixed between- and within-subjects analysis of participant gender and makeup level was plotted against mean z-score and time (ms). Detection of extreme curves was conducted for all participants defined as curves showing large variation with respect to the mean curve under the same conditions of muscle site and gender. A FANOVA using Type-II sum of squares was then conducted to test the effect of makeup condition on EMG response, and makeup condition and gender on participant z-scores (Causeur et al., 2019a; 2019b).

Experiment 2 - Results

Facial EMG Response

The Type-II functional ANOVA of the corrugator EMG response revealed a significant main effect of makeup condition (F = 41.25, $p = .003$), with Bonferroni-corrected pairwise comparisons revealing significant differences between M1 and M2 ($p < .001$), M1 and M3 ($p < .001$), M2 and M3 ($p < .001$), M3 and M5 ($p = .042$), and significant differences between M6 and all other makeup levels (M1: $p < .001$; M2: $p < .001$; M3: $p < .001$; M4: $p = .011$; M5: $p < .001$) (see Figure 4 for details). No other significant effects were observed at other muscle sites for EMG response (Zygomatic: F = 28.75, $p = 0.325$; Levator: F = 22.87, $p = 0.796$; Frontalis: F = 20.19, $p = 0.859$).

FIGURE 4 GOES HERE

The Type-II functional ANOVA conducted on the z-scores again revealed a significant main effect of makeup condition for the corrugator muscle (F = 45.61, $p < .001$), with Bonferroni-corrected pairwise comparisons revealing a significant difference between M1 and M3 ($p = .048$) and between M2 and M3 ($p = .006$), with no significant main effects of makeup condition observed for the other muscle sites (Zygomatic: F = 23.61, $p = 0.785$; Levator: F = 23.71, $p = 0.768$; Frontalis: F = 19.86, $p = 0.918$) (see Figure 5 for details). In addition, significant main effects of participant gender were observed for the corrugator (F = 6.873, $p < .001$) and levator (F = 2.845, $p < .001$), revealing significantly higher z-scores in both instances for female participants compared to males (see Figure 6 for details). No significant main effects of gender were observed for the zygomatic (F = 1.509, $p = 0.197$) or frontalis (F = 0.974, $p = 0.529$) sites. No significant interactions were observed between gender and makeup condition during the time period of stimulus presentation (Corrugator: F = 24.37, $p = 0.586$; Zygomatic: F = 40.43, $p = 0.138$; Levator: F = 27.22, $p = 0.488$; Frontalis: F = 25.42, $p = 0.503$).

FIGURE 5 GOES HERE

FIGURE 6 GOES HERE

General Discussion

Overall, the results from this study indicate a greater behavioural tendency for greater approach and lower avoidance to no and light makeup in female participants, and a specifically male behavioural tendency for greater approach and lower avoidance for heavy over light makeup, and reveal differences in corrugator response (indicating negative affect) towards varying levels of makeup in both men and women, as well as a higher corrugator and levator response in women than men towards all makeup levels. Given the linear increase in reported attractiveness observed over all makeup levels in the explicit rating task (although insignificant from M2 – M7), the results from our EMG analysis are in line with past research showing a negative linear relationship between facial attractiveness and corrugator response (Hazlett & Hoehn-Saric, 2000; Principe & Langlois, 2011). Furthermore, the present study extends past research in both social psychology (Mulhern et al., 2003; Nash et al., 2006), visual perception (Koehler et al., 2004; Russell, 2009), and aesthetic neuroscience (Chatterjee et al., 2009; Tagai et al., 2017), by revealing the interplay between gender, aesthetic preference and visual fluency through the use of implicit responses, as well as the contribution of visual cues linked to female facial attractiveness to implicit emotional response. For the first time, the present study shows changes in electromyographical activity linked to facial affect in response to different levels of makeup in face stimuli.

Overall, the AAT task revealed a main effect of gender on behavioural response. However, this effect appears to be in part due to the faster reaction times of the male group as compared to the female group across all images, as confirmed by separate analyses of the median approach and avoidance RTs (see Tables 2 and 3 in Supplement), revealing that male participants were faster in their response to both approach and avoidance of the images. Thus, differences in the behavioural response to the distinct makeup levels should be considered with respect to a separate baseline for each gender. Interestingly, simple effect analyses conducted on the significant interaction between gender and makeup level showed that female participants responded more positively to faces with light makeup (foundation) than heavy makeup applied to the eye regions, with AAT scores for M2 higher than M5 (pencil) and M7 (eyeshadow), consistent with a previously-reported "visual fluency" effect of light makeup in female participants (Tagai et al., 2016; 2017). In contrast, males showed an "all-or-nothing" effect of behavioural preference towards makeup, with simple effects analysis showing significantly higher AAT scores in response to no makeup (M1) and heavy eye makeup (M6 and M7) as opposed to light and intermediate (M2 – M5) levels of makeup. These two distinct patterns of results appear to correspond to separate mechanisms of visual expertise and sexual preference, as described below.

With regards to a previously-reported light makeup advantage for visual fluency, the behavioural AAT response indicated a partial preference of female participants for faces with little makeup as compared to medium and high makeup faces, primarily due to a slower avoidance response to these faces (see Table 1 in the Supplement). As a previous study by Tagai et al. (2016) found a recognition bias for light makeup faces in female participants, this effect is in line with past research, and indicates an additional bias in terms of approach-avoidance behaviour for light makeup. However, given the present study does not systematically vary the information content within each makeup level (i.e. visual spatial frequency, skin tone), these results do not offer widespread support for or against a visual fluency account of the cosmetic enhancement of facial attractiveness. Instead, the higher response to foundation (M2) in female participants may simply be due to increased sensitivity to the visual cues provided by foundation, as all female participants reported regularly using facial cosmetics. By contrast, the specifically-male preference for heavy makeup over medium makeup appears to indicate the presence of a secondary effect of sexual propensity towards heavy makeup primarily accentuating greater visual contrast in the eye regions, consistent with the corrugator response recorded in Experiment 2, discussed later.

The surprising result of higher AAT response for no makeup compared to light makeup in male participants, appears to have no previous correspondence in the literature, as "no makeup" conditions have been consistently rated as the least attractive faces according to past studies utilising explicit ratings of attractiveness in response to cosmetic enhancement (Mileva et al., 2016; Mulhern et al., 2003; Osborn, 1996), a result confirmed by the explicit rating of the attractiveness of the faces in the present study, revealing that "no makeup" was judged as significantly less attractive than all other makeup levels. However, at an implicit behavioural level, positive male responses to faces with no makeup may be due to the evaluation of such faces as neutral, non-sexualised stimuli, as opposed to cosmetically-enhanced female faces viewed as potential mates and rejected at the initial phase of makeup application. Similarly, female participants may show more positive approach-avoidance behaviour towards faces with no makeup due to being viewed as neutral non-threatening competitors (Stockley & Campbell, 2013).

With regards to the facial EMG response in Experiment 2, activity at the corrugator muscle site over the course of stimulus presentation was significantly higher in response to no makeup (M1) than light makeup (M2 and M3) and eyeshadow (M6), revealing a more relaxed corrugator pose in response to addition of these levels of cosmetic application, indicative of decreased negative affect (Principe & Langlois, 2011). In addition, lipstick (M3) was found to play a key role in the attenuation of corrugator response, with the lowest corrugator response recorded and significantly lower than all levels with the exception of mascara and pencilling (M4). Finally, eyeshadow (M6) displayed the next lowest corrugator response, with significantly lower values to all other makeup levels with the exception of lipstick. Overall, these results indicate a significant effect of makeup application on the attenuation of corrugator response. While corrugator response has been associated with increased cognitive load (Lishner et al., 2008) which may have contributed to the smoothing effect of foundation (M2) on corrugator activity, these results primarily indicate lower negative affect in response to an increase in the stages of makeup application. Interestingly, this effect was observed in both genders, indicating a similar affective response to the visual cues of makeup. Notably however, the addition of eye makeup such as mascara, pencilling and eyeliner (associated with higher visual contrast) did not produce a significant decrease in corrugator response as compared to the no makeup condition. Thus, we found no evidence for the effect of these products on facial attractiveness as gauged by corrugator response. Instead, significant attenuation of corrugator response was observed only in response to the addition of eyeshadow. Visual contrast accentuating the eye and lip regions has been proposed as one of the major determinants of facial attractiveness and femininity in female faces (Russell, 2009; 2010). Our results suggest that the visual contrast in the eye regions must be sufficiently intense to produce a change in affective response contributing to perceived attractiveness. This is consistent with the observation of a linear increase in global contrast factor (Matkovic et al., 2005) on levels M4 – M7 of the images employed in the AAT task as well as a gradual, but non-significant, increase in explicitly-rated attractiveness of these makeup levels. The significant decrease in corrugator activity observed for the addition of lipstick highlights the importance of this region in providing a visual cue likely associated with luminance contrast (Russell, 2003; 2009), to determine an appropriate affective response for guiding the evaluation of facial attractiveness. The marked reduction in corrugator response towards the presence of lipstick and eyeshadow does not support the role of visual fluency in reducing cognitive load as the sole determinant of perceived attractiveness, at least on an affective level, instead indicating that the evaluation of facial attractiveness as enhanced by makeup relies on a wide set of visual cues eliciting distinct behavioural and affective reactions.

An analysis of the participants' z-scores over the course of stimulus presentation revealed higher corrugator and levator activity for female subjects than male subjects. Interestingly, this gender difference occurred at an early peak of EMG response, likely corresponding to an orienting response towards novel facial stimuli (Achaibou et al., 2008; Dimberg, 1982). While corrugator activity has specifically been correlated with early visual processing (approximately 200 ms after stimulus onset; Achaibou et al., 2008), we cannot discount the possibility that this difference may be due to enhanced attention towards facial stimuli containing makeup in female participants. Alternatively, this result may reflect an initially adverse negative emotional response in women towards female faces prior to subsequent modulation, indicating increased female intrasexual competition (Stockley & Campbell, 2013).

Similarly, the increase in corrugator and levator response in women than men in response to all images is unlikely to be due to increased cognitive load due to the greater familiarisation the women had with the makeup products applied (all female participants reported regularly using makeup at least once per week). While a recent large-scale study examining facial muscles according to the Facial Action Coding

System (FACS; Ekman et al, 2002) has shown greater expressiveness in female facial actions associated with positive valence, a corresponding difference in negative facial affect between male and female faces was not reported (McDuff et al., 2017). Given these factors and the well-established link between corrugator response and negative affect (Larsen et al., 2003; Neta et al., 2009), the corrugator response observed in the present study can reasonably be attributed to an affective index of aesthetic preference for facial cosmetics, revealing a more negative affective response to female faces with and without makeup in women than men.

An important caveat must be made with respect to the limitations of utilising facial EMG in the measurement of differences in aesthetic judgement. We were unable to compile a complete affective 'profile' of the valence and intensity of participants' emotional response due to the variability and lack of significant voltage changes to the different stimulus types used in Experiment 2, at all muscle sites with the exception of the corrugator supercilii and levator (Figure 3.). While EMG measurement was more sensitive than AAT response to intermediate differences in makeup application, for example in the distinct perceptual cues associated with lipstick and eye makeup, as a whole facial EMG may be insufficiently sensitive towards the effect of relatively subtle physiological cues on facial attractiveness, and future studies investigating the affective responses underlying aesthetic experience should consider pairing the technique with an explicit attractiveness rating task, for example. The inclusion of eye-tracking measures to monitor which precise face regions the participant attends to while rating attractiveness, may also provide a useful 'attentional' index of aesthetic preference. Previous studies have tested the role of eye gaze in evaluating female facial attractiveness, indicating both that attractive faces receive longer gaze durations and a greater number of directed saccades than unattractive faces (Leder et al., 2016) and that participants attend longer to the nose than other facial regions during the evaluation of facial attractiveness (Zhang et al., 2017). Future research may be directed at the role of cosmetic enhancement in guiding attention during attractiveness judgments.

A further caveat is the limited support this study found for the smoothing, texturing and colour distribution effects of foundation on the evaluation of facial attractiveness, commonly associated with signals of youth and individual health (Fink & Matts, 2008; Jones et al., 2015; Porcheron et al., 2013). While there was a slight (but non-significant) increase in mean AAT score from M1 (no makeup) to M2 (foundation) for female participants, this effect was inverted in male subjects, showing significantly greater behavioural preference for no makeup than foundation. However, EMG recording of the corrugator site revealed a significantly lower response to faces with foundation than faces with no makeup in both genders, suggesting that this cue of facial attractiveness is more dependent on one's affective response than the enhancement of visual contrast in the eye regions for example, which was reflected in both EMG and AAT response.

Overall, the present study found two clear indices of the implicit evaluation of facial attractiveness as modulated by changes in facial cosmetics; a behavioural index, characterised in female participants by a preference for faces with light makeup, and in males by an all-or-nothing preference for faces with no makeup or heavy eye cosmetics. The second index corresponds to the individual's negative affective response, reflected primarily in terms of reduced electromyographical response at the corrugator muscle site to facial cosmetics accentuating visual contrast in the mouth and eye regions. Given the counterintuitive results reported, particularly with regards to a male implicit behavioural preference for no makeup over light makeup, the evaluation of female facial attractiveness appears to rely on a complex set of perceptual and behavioural cues, highlighting the importance of implicit measures in further investigations.

References

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

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

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

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

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

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

Chatterjee, A., & Vartanian, O. (2016). Neuroscience of aesthetics. *Annals of the New York Academy of Sciences*, *1369*(1), 172-194. https://doi.org/10.1111/nyas.13035

Cunningham, M. R., Roberts, A. R., Barbee, A. P., Druen, P. B., & Wu, C. H. (1995). "Their ideas of beauty are, on the whole, the same as ours": Consistency and variability in the cross-cultural perception of female physical attractiveness. *Journal of personality and social psychology*, *68*(2), 261. https://doi.org/10.1037/0022-3514.68.2.261

Dimberg, U. (1982). Facial Reactions to Facial Expressions. *Psychophysiology*, *19*(6), 643–647. https://doi.org/cmpnb7

Dimberg, U., & Thunberg, M. (1998). Rapid facial reactions to emotional facial expressions. *Scandinavian Journal of Psychology*, *39*(1), 39–45. https://doi.org/10.1111/1467-9450.00054

Dimberg, U., Thunberg, M., & Elmehed, K. (2000). Unconscious facial reactions to emotional facial expressions. *Psychological Science*, *11*(1), 86-89. https://doi.org/10.1111/1467-9280.00221

Ekman, P., Friesen, W. V., & Hager, J. C. (2002). Facial action coding system: The manual on CD ROM. *A Human Face, Salt Lake City*, 77-254.

Etcoff, N. L., Stock, S., Haley, L. E., Vickery, S. A., & House, D. M. (2011). Cosmetics as a feature of the extended human phenotype: Modulation of the perception of biologically important facial signals. *PloS one*, *6*(10), e25656. https://doi.org/10.1371/journal.pone.0025656

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175-191. https://doi.org/10.3758/BF03193146

Fridlund, A. J., & Cacioppo, J. T. (1986). Guidelines for human electromyographic research. *Psychophysiology*, *23*(5), 567-589. https://doi.org/btrbt5

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

Fink, B., & Matts, P. J. (2008). The effects of skin colour distribution and topography cues on the perception of female facial age and health. *Journal of the European Academy of Dermatology and Venereology*, *22*(4), 493-498. https://doi.org/10.1111/j.1468-3083.2007.02512.x

Fink, B., & Neave, N. (2005). The biology of facial beauty. *International Journal of Cosmetic Science*, *27*(6), 317-325. https://doi.org/c978s9

Fink, B., & Penton-Voak, I. (2002). Evolutionary psychology of facial attractiveness. *Current Directions in Psychological Science*, *11*(5), 154-158. https://doi.org/10.1111/1467-8721.00190

Gerger, G., & Leder, H. (2015). Titles change the esthetic appreciations of paintings. *Frontiers in Human Neuroscience*, *9*, 464. https://doi.org/10.3389/fnhum.2015.00464

Gerger, G., Leder, H., Tinio, P. P., & Schacht, A. (2011). Faces versus patterns: Exploring aesthetic reactions using facial EMG. *Psychology of Aesthetics, Creativity, and the Arts, 5*(3), 241. https://doi.org/10.1037/a0024154

Langsrud, Ø. (2003). ANOVA for unbalanced data: Use Type II instead of Type III sums of squares. *Statistics and Computing*, *13*(2), 163-167.

Hazlett, R. L., & Hoehn-Saric, R. (2000). Effects of perceived physical attractiveness on females' facial displays and affect. *Evolution and Human Behavior*, *21*(1), 49-57. https://doi.org/10.1016/S1090-5138(99)00036-7

Heuer, K., Rinck, M., & Becker, E. S. (2007). Avoidance of emotional facial expressions in social anxiety: The approach–avoidance task. *Behaviour Research and Therapy*, *45*(12), 2990-3001. https://doi.org/10.1016/j.brat.2007.08.010

Hess, U., & Fischer, A. (2014). Emotional mimicry: why and when we mimic emotions. *Social and Personality Psychology Compass*, *8*(2), 45–57. https://doi.org/10.1111/spc3.12083

Jones, A. L., & Kramer, R. S. (2016). Facial Cosmetics and Attractiveness: Comparing the Effect Sizes of Professionally-Applied Cosmetics and Identity. *PloS One*, *11*(10), e0164218. https://doi.org/10.1371/journal.pone.0164218

Jones, B. C., Hahn, A. C., Fisher, C. I., Wincenciak, J., Kandrik, M., Roberts, S. C., ... & DeBruine, L. M. (2015). Facial coloration tracks changes in women's estradiol. *Psychoneuroendocrinology*, *56*, 29-34. https://doi.org/10.1016/j.psyneuen.2015.02.021

Koehler, N., Simmons, L. W., Rhodes, G., & Peters, M. (2004). The relationship between sexual dimorphism in human faces and fluctuating asymmetry. *Proceedings of the Royal Society of London B: Biological Sciences*, *271*(Suppl 4), S233-S236. https://doi.org/10.1098/rsbl.2003.0146

Kramer, R. S., Mulgrew, J., Anderson, N. C., Vasilyev, D., Kingstone, A., Reynolds, M. G., & Ward, R. (2020). Physically attractive faces attract us physically. *Cognition*, *198*, 104193. https://doi.org/10.1016/j.cognition.2020.104193

Krieglmeyer, R., & Deutsch, R. (2010). Comparing measures of approach–avoidance behaviour: The manikin task vs. two versions of the joystick task. *Cognition and Emotion*, *24*(5), 810-828. https://doi.org/10.1080/02699930903047298

Kukde, M. P., & Neufeld, R. W. (1994). Facial electromyographic measures distinguish covert coping from stress response to stimulus threat. *Personality and Individual Differences*, *16*(2), 211-228. https://doi.org/10.1016/0191-8869(94)90160-0

Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, *126*(3), 390. https://doi.org/cfb9cd

Larsen, J. T., Norris, C. J., & Cacioppo, J. T. (2003). Effects of positive and negative affect on electromyographic activity over zygomaticus major and corrugator supercilii. *Psychophysiology*, *40*(5), 776-785. https://doi.org/dkc6vh

Law-Smith, M., Perrett, D. I., Jones, B. C., Cornwell, R. E., Moore, F. R., Feinberg, D. R., Boothroyd, L. G., Durrani, S. J., Stirrat, M. R., Whiten, S., Pitman, R. M., & Hillier, S. G. (2006). Facial appearance is a cue to oestrogen levels in women. *Proceedings of the Royal Society of London B: Biological Sciences*, *273*(1583), 135-140. https://doi.org/10.1098/rspb.2005.3296

Leder, H., Mitrovic, A., & Goller, J. (2016). How beauty determines gaze! Facial attractiveness and gaze duration in images of real world scenes. *i-Perception*, *7*(4), 2041669516664355. https://doi.org/10.1177/2041669516664355

Lishner, D. A., Cooter, A. B., & Zald, D. H. (2008). Rapid emotional contagion and expressive congruence under strong test conditions. *Journal of Nonverbal Behavior*, *32*(4), 225-239. https://doi.org/10.1007/s10919-008-0053-y

Little, A. C., Jones, B. C., & DeBruine, L. M. (2011). Facial attractiveness: evolutionary based research. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, *366*(1571), 1638-1659. https://doi.org/10.1098/rstb.2010.0404

Lundqvist, L.-O. (1995). Facial EMG reactions to facial expressions: A case of facial emotional contagion? *Scandinavian Journal of Psychology*, *36*(2), 130–141. https://doi.org/bs85wt

Lundqvist, L.-O., & Dimberg, U. (1995). Facial expressions are contagious. *Journal of Psychophysiology*, *9*(3), 203–211.

Matkovic, K., Neumann, L., Neumann, A., Psik, T., & Purgathofer, W. (2005). Global contrast factor-a new approach to image contrast. *Computational Aesthetics in Graphics, Visualization and Imaging*, *2005*(159-168), 1. https://doi.org/fhxp

McDuff, D., Kodra, E., el Kaliouby, R., & LaFrance, M. (2017). A large-scale analysis of sex differences in facial expressions. *PloS one*, *12*(4), e0173942. https://doi.org/10.1371/journal.pone.0173942

Mileva, V. R., Jones, A. L., Russell, R., & Little, A. C. (2016). Sex differences in the perceived dominance and prestige of women with and without cosmetics. *Perception*, *45*(10), 1166-1183. https://doi.org/f86f6v

Morrison, E. R., Clark, A. P., Tiddeman, B. P., & Penton-Voak, I. S. (2010). Manipulating shape cues in dynamic human faces: Sexual dimorphism is preferred in female but not male faces. *Ethology*, *116*(12), 1234-1243. https://doi.org/bgg6dd

Mulhern, R., Fieldman, G., Hussey, T., Lévêque, J. L., & Pineau, P. (2003). Do cosmetics enhance female Caucasian facial attractiveness?. *International Journal of Cosmetic Science*, *25*(4), 199-205. https://doi.org/frmj9p

Nakamura, K., Kawashima, R., Nagumo, S., Ito, K., Sugiura, M., Kato, T., ... & Kojima, S. (1998). Neuroanatomical correlates of the assessment of facial attractiveness. *Neuroreport*, *9*(4), 753-757. https://doi.org/c7hrsm

Nash, R., Fieldman, G., Hussey, T., Lévêque, J. L., & Pineau, P. (2006). Cosmetics: They influence more than Caucasian female facial attractiveness. *Journal of Applied Social Psychology*, *36*(2), 493-504. https://doi.org/c34j72

Neta, M., Norris, C. J., & Whalen, P. J. (2009). Corrugator muscle responses are associated with individual differences in positivity-negativity bias. *Emotion*, *9*(5), 640-648. https://doi.org/10.1037/a0016819

Osborn, D. R. (1996). Beauty Is as Beauty Does?: Makeup and Posture Effects on Physical Attractiveness Judgments. *Journal of Applied Social Psychology*, *26*(1), 31-51. https://doi.org/10.1111/j.1559-1816.1996.tb01837.x

Penton-Voak, I. S., Jacobson, A., & Trivers, R. (2004). Populational differences in attractiveness judgements of male and female faces: Comparing British and Jamaican samples. *Evolution and Human Behavior*, *25*(6), 355-370. https://doi.org/10.1016/j.evolhumbehav.2004.06.002

Porcheron, A., Mauger, E., & Russell, R. (2013). Aspects of facial contrast decrease with age and are cues for age perception. *PloS one*, *8*(3), e57985. https://doi.org/10.1371/journal.pone.0057985

Principe, C. P., & Langlois, J. H. (2011). Faces differing in attractiveness elicit corresponding affective responses. *Cognition and Emotion*, *25*(1), 140-148. https://doi.org/10.1080/02699931003612098

Principe, C. P., & Langlois, J. H. (2012). Shifting the prototype: Experience with faces influences affective and attractiveness preferences. *Social Cognition*, *30*(1), 109-120. https://doi.org/10.1521/soco.2012.30.1.109

Ramsay, J. O., & Silverman, B. W. (2005). *Functional Data Analysis (2nd Edition)*. Springer Series in Statistics. Springer. https://doi.org/10.1007/b98888

Rhodes, G. (2006). The evolutionary psychology of facial beauty. *Annual Review of Psychology*, *57*, 199-226. https://doi.org/ddg2hh

Richetin, J., Croizet, J. C., & Huguet, P. (2004). Facial make-up elicits positive attitudes at the implicit level: Evidence from the implicit association test. *Current Research in Social Psychology*, *9*(11), 145-164.

Rinck, M., & Becker, E. S. (2007). Approach and avoidance in fear of spiders. *Journal of Behavior Therapy and Experimental Psychiatry, 38*(2), 105-120. https://doi.org/10.1016/j.jbtep.2006.10.001

Rinn, W. E. (1984). The neuropsychology of facial expression: a review of the neurological and psychological mechanisms for producing facial expressions. *Psychological Bulletin*, *95*(1), 52–77. https://doi.org/10.1037/0033-2909.95.1.52

Russell, R. (2003). Sex, beauty, and the relative luminance of facial features. *Perception*, *32*(9), 1093-1107. https://doi.org/10.1068/p5101

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. (2010). Why cosmetics work. In R.B. Adams, N. Ambady, K. Nakayama, and S. Shimojo, (Eds.) *The Science of Social Vision* (pp. 186-204). Oxford Series in Social Cognition vol. 7. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195333176.003.0011

Sofer, C., Dotsch, R., Wigboldus, D. H., & Todorov, A. (2015). What is typical is good: The influence of face typicality on perceived trustworthiness. *Psychological Science*, *26*(1), 39-47. https://doi.org/10.1177/0956797614554955

Stockley, P., & Campbell, A. (2013). Female competition and aggression: interdisciplinary perspectives. https://doi.org/10.1098/rstb.2013.0073

Tagai, K., Ohtaka, H., & Nittono, H. (2016). Faces with light makeup are better recognized than faces with heavy makeup. *Frontiers in Psychology*, *7*, 226. https://doi.org/10.3389/fpsyg.2016.00226

Tagai, K., Shimakura, H., Isobe, H., & Nittono, H. (2017). The light-makeup advantage in facial processing: Evidence from event-related potentials. *PloS One*, *12*(2), e0172489. https://doi.org/10.1371/journal.pone.0172489

Thornhill, R., & Grammer, K. (1999). The body and face of woman: One ornament that signals quality? *Evolution and Human Behavior*, *20*(2), 105-120. https://doi.org/10.1016/S1090-5138(98)00044-0

Todorov, A. (2008). Evaluating faces on trustworthiness: an extension of systems for recognition of emotions signaling approach/avoidance behaviors. *Annals of the New York Academy of Sciences*, *1124*(1), 208-224. https://doi.org/fk35vn

Todorov, A., Baron, S. G., & Oosterhof, N. N. (2008). Evaluating face trustworthiness: a model based approach. *Social Cognitive and Affective Neuroscience*, *3*(2), 119-127. https://doi.org/10.1093/scan/nsn009

Wiers, R. W., Rinck, M., Dictus, M., & Van den Wildenberg, E. (2009). Relatively strong automatic appetitive action-tendencies in male carriers of the OPRM1 G-allele. *Genes, Brain and Behavior*, *8*(1), 101-106. https://doi.org/cbvrc7

Winkielman, P., Halberstadt, J., Fazendeiro, T., & Catty, S. (2006). Prototypes are attractive because they are easy on the mind. *Psychological Science*, *17*(9), 799-806. https://doi.org/dt4c3m

Winston, J. S., O'Doherty, J., Kilner, J. M., Perrett, D. I., & Dolan, R. J. (2007). Brain systems for assessing facial attractiveness. *Neuropsychologia*, *45*(1), 195-206. https://doi.org/d38755

Zebrowitz, L. A., & Montepare, J. M. (2008). Social psychological face perception: Why appearance matters. *Social and Personality Psychology Compass*, *2*(3), 1497-1517. https://doi.org/ff3s4w

Zhang, Y., Wang, X., Wang, J., Zhang, L., & Xiang, Y. (2017). Patterns of eye movements when observers judge female facial attractiveness. *Frontiers in Psychology*, *8*, 1909. https://doi.org/10.3389/fpsyg.2017.01909

Figure Legends

Figure 1. Examples of the different makeup levels used in Experiment 1 (a) and Experiment 2 (b). For illustrative purposes, this figure shows the same individual at different stages of makeup application. A total of 18 individuals were shown in all makeup levels. The individual shown gave explicit written consent for the publication of her face images.

Figure 2. Graphs of median response scores of a) attractiveness ("How attractive is this face from 1 to 7?") and b) emotion ("How emotional is this face from 1 to 7?), as measured on the explicit rating task in Experiment 1. Scores from 1 to 7 reflect faces judged as a) 1: not at all attractive to 7: very attractive; and b) 1: very emotionally negative, to 4: emotionally neutral, to 7: very emotionally positive. Error bars show ± 1 standard error of the mean.

Figure 3. Graph of AAT difference scores for each makeup level (no makeup, foundation, lipstick, mascara, pencil, eyeliner, eyeshadow), with separate lines for gender. AAT scores refer to median avoidance RT – median approach RT, with higher scores reflecting faster approach and slower avoidance of the image. Error bars show ± 1 standard error of the mean.

Figure 4. Graphs of EMG values (μv) recorded at the corrugator, zygomatic, levator, and frontalis sites in response to viewing of 6 makeup levels (M1: no makeup; M2: foundation; M3: lipstick; M4: mascara + pencil; M5: eyeliner; M6: eyeshadow). Curves show average EMG voltage change across participants for each 100 ms bin over the periods of pre-stimulus baseline (-500 to 0 ms), stimulus presentation (0 to 2000 ms) and interstimulus interval (ISI; 2000 to 4000 ms).

Figure 5. Graphs of Z-transformed EMG values at the corrugator, zygomatic, frontalis and levator sites in response to viewing of 6 makeup levels (M1: no makeup; M2: foundation; M3: lipstick; M4: mascara + pencil; M5: eyeliner; M6: eyeshadow). Solid curves show average Z-scores of all participants and shaded areas show confidence intervals for each 100 ms bin over the periods of pre-stimulus baseline (-500 to 0 ms), stimulus presentation (0 to 2000 ms) and interstimulus interval (ISI; 2000 to 4000 ms).

Figure 6. Graphs of Z-transformed EMG values at the corrugator and levator sites in male and female participants. Solid curves show average Z-scores of male and female participants and shaded areas show confidence intervals for each 100 ms bin over the periods of pre-stimulus baseline (-500 to 0 ms), stimulus presentation (0 to 2000 ms) and interstimulus interval (ISI; 2000 to 4000 ms).