Attard-Johnson Janice (Orcid ID: 0000-0002-0563-0197)

Your ID, please? The effect of Facemasks and Makeup on Perceptions of Age of Young

Adult Female Faces

Davis, H.[1], & Attard-Johnson, J.[1,2]

Word count: 5511 (including references, excluding abstract and figures)

[1]Faculty of Science and Technology, Bournemouth University, Poole, UK

[2]Correspondence for this article should be addressed to Janice Attard-Johnson, Department of

Psychology, Faculty of Science and Technology, Bournemouth University, Fernbarrow,

Poole, BH12 5BB, UK. Email: jattardjohnson@bournemouth.ac.uk

ABSTRACT

During the COVID-19 pandemic, wearing facemasks was mandatory in the United Kingdom except for individuals with medical exemptions. Facemasks cover the full lower half of the face, however the effect of facemasks on age perception is not yet known. The present study examined whether age estimation accuracy of unfamiliar young adult women is impaired when the target is wearing a facemask. This study also examined whether makeup, which has previously been shown to increase error bias, further impairs age estimation accuracy when paired with a facemask. The findings indicate that both facemasks and makeup tend to result in over-estimation of the young women's age compared to neutral faces, but the combination of both is not additive. Individual level analysis also revealed large individual differences in age estimation accuracy ranging from estimates within 1 year of the target's actual age, and age estimates which deviated by up to 20 years.

INTRODUCTION

Accurate age estimation of strangers is paramount in situations relating to forensic age identification in criminal investigations (Thorley et al., 2018) and age verification for age restricted sales (Willner & Rowe, 2001). Laboratory studies demonstrate that age estimates are moderately accurate with errors within the range of 3 to 5 years in adults (Sõrqvist et al., 2007). However, factors which impair age estimation accuracy by up to 8 years have also been identified (Clifford et al., 2018; Dehon & Brédart, 2001; Voelkle et al., 2012; for reviews see Moyse, 2014; Rhodes, 2009). One example is the use of sunglasses to obscure the eye region and disguise visual cues of age within this region (Thorley, 2019). However, the effect of obscuring other facial features on age perception is less known.

During the COVID-19 pandemic in the UK, individuals were mandated to wear a facemask when indoors with members of different households or with strangers, when on public transport, in shopping centres and supermarkets, and on hospitality premises (Department of Health and Social Care, 2021). Although use of facemasks potentially presents a challenge for salespeople selling alcohol, knives, or other age restricted items, the effect of facemasks on age estimation accuracy is not known. The challenge is further compounded by individuals who may unintentionally or deliberately attempt to appear older when wearing a facemask by also applying facial cosmetics to alter the visual cues of the visible facial features. Stores are required to use the '*Challenge 25*' Policy and should therefore request identification to confirm that the buyers are over the age of 18 when they perceive an individual to be under the age of 25. However, the extent to which facemasks can impair age estimation accuracy, if at all, is not clear. Although government mandates to wear facemasks will gradually lift as the pandemic subsides, some people will choose to continue to wear facemasks as evident from some cultures during pre-pandemic times (Lau et al.,

2010). Therefore, the question of whether facemasks, and makeup, distort age perception will likely continue to be relevant even after restrictions have been lifted.

The perceived age of women is typically found to be less accurate compared to men (Voelke et al., 2012), with younger faces frequently overestimated and older faces underestimated (Voelkle et al., 2012; Watson et al., 2016). The discrepancy between men and women may be attributed, at least in part, to the use of cosmetics. Women are more likely than men to apply cosmetics to alter their physical appearance usually with the desired effect of appearing more attractive and decreasing negative self-perception (Korichi et al., 2008). Previous studies revealed that faces wearing cosmetics are judged as healthier (Nash et al., 2006) and more attractive (Mulhern et al., 2003), and the latter correlates negatively with age perception (Henss, 1991). Although the effect of cosmetics on human age estimation has not been extensively researched, a small set of studies support the notion that cosmetics alter the perception of age (Russell, 2003; Russell et al., 2019).

Cosmetics alter the perception of specific facial features which act as visual cues for age. One cue is facial contrast, which comprises the colour differences and luminance in the skin and between the facial features. Facial contrast changes with age, and faces with increased facial contrast (e.g., darker eyebrows and lips) are perceived to be younger compared to those with decreased facial contrast (Porcheron et al., 2013). Consequently, makeup which is applied to darken the lips, increase skin luminance, and increase the colour contrast of the eyes and eyebrows, work to increase facial contrast and result in younger looking faces (Russell et al., 2019). Makeup applied to different facial features also have different effects on age perception. For example, makeup applied to the eyes and eyebrows has a stronger impact on age perception compared to application to the lips (Russell et al., 2019). Additionally, the effects of these manipulations to the appearance of the face are dependent on the target age (Egan & Cordan, 2009; Russell et al., 2019). Specifically,

photographs of faces depicting 40 and 50-year-old women are perceived on average 1.5 years younger when wearing full makeup, while 20-year-old women appear 1.4 years *older* with identical makeup (Russell et al., 2019). Although these estimation errors are relatively modest, other studies have reported a larger range of errors of up to 20 years older or younger than actual age (Dayan et al., 2015; Fink et al., 2006).

The difference in the strength of the effects when applying makeup to lips compared to the eye region suggests that some facial features may play a more prominent role in providing visual cues to age. Only a small set of studies have examined the effect of obscuring different facial features on age perception. In those studies, obscuring head shape and hair did not decrease accuracy (George & Hole, 1995), but obscuring the eye region led to a greater reduction in accuracy compared to disguising the hair and forehead or no disguise at all (Thorley, 2019). Furthermore, a recent eye-tracking study suggests that the central triangle (collectively eyes, nose, and mouth) may be important for age estimation (Liao et al., 2020). It is therefore not inconceivable that obscuring any visual information within the central triangle would disrupt the age estimation process. However, these studies did not specifically obscure the lower half of the face, and therefore the impact of wearing a facemask on age perception is not yet known.

Furthermore, individuals may also apply makeup which, in addition to the facemask, could decrease age estimation accuracy further. The combined effect of makeup and masks on age perception is yet to be examined. As makeup to the eye-region has been found to introduce more errors compared to other regions (Russell et al., 2019), makeup application to the eyes in combination with an absence of visible visual cues from the entire bottom half of the face, could result in further deterioration of age estimation accuracy. Our aim was to determine whether wearing a facemask or makeup affects age estimation accuracy, and whether pairing facemasks and makeup would further increase age estimation error.

Specifically, we expected that makeup would *increase* the perception of age for young adult women, and that facemasks would impair age estimation however we did not predict whether the impairment would be biased towards an over or underestimation of age. Furthermore, we expected that the combined use of facemasks and makeup would yield a greater increase in overestimation bias compared to makeup alone, due to a reduction in facial age cues and enhancing of visible cues (eye region) through cosmetics.

METHOD

Participants

A total of 68 participants (57 female, 11 male) aged 18-65 years (Mean = 36 years, SD = 13.81) volunteered in this experiment. Sixty-five indicated that they were Caucasian, one Asian, one Hispanic and one selected 'other'. Participants were recruited via social media and did not receive any monetary compensation for their participation. All reported normal or corrected-to-normal vision.

For the construction of the stimuli, 33 female participants volunteered. Of these, 27 were students from Bournemouth University aged 18-21, and 6 were recruited via social media aged 48 – 73. Participants received course credit or small monetary payment for their time. All volunteers were Caucasian except for three, two Chinese and one Mixed Caribbean. All the procedures were approved by the Ethics Committee at Bournemouth University.

Stimuli

Volunteers took passport-style photographs of their faces in the four following ways: 1) no makeup and no mask, 2) makeup and a mask, 3) makeup and no mask, and 4) no makeup and mask. Due to COVID-19 government restrictions, participants could not attend the laboratory to have their photographs taken, therefore they were provided with detailed instructions to take the photographs within their own home. Photographs were taken using a

digital camera or a phone camera and a digital timer or the assistance of a family member to steady the camera in front of the face. Participants were instructed not to take a 'selfie' with their arms extended to avoid distortion of facial features. Participants were instructed to stand in front of a white or cream background positioned facing towards a window with no shadows on the face, to remove headwear and glasses, to ensure hair was not covering the face, to take the photographs with a neutral facial expression, and to use a blue medical non-medical mask for the masked photographs. For the makeup photographs, participants were instructed to apply makeup as though they *'were attending an event'*, and to keep the same makeup for both makeup conditions (i.e., with and without mask). All participants wore foundation, and applied eyeshadow and mascara, and some, but not all, individuals applied eyeliner and eyebrow definition. Participants were provided with a checklist to self-check the quality of their photographs. Participants also provided a unique pseudonym ID which they were asked to keep so that they may withdraw their image from the database and future use of the images if they decided.

The researcher screened all photographs, and five sets of images were excluded as they did not meet the criteria or were of insufficient quality. Therefore, a total of 28 photograph sets (112 images) were included in the experiment. All faces included in the final stimulus set were Caucasian except for three (two Chinese and one Mixed Caribbean). All images were cropped around the face to exclude the background and neck using GNU Image Manipulation Program (version 2.10) photo editing software (for an example, see Figure 1). All face images were resized to a width of 170 pixels and placed in the centre of the screen on a white background. The final stimulus set comprised 28 image sets, of these 22 sets included faces of women aged between 18 and 21, and six sets of women aged between 48 and 73. The focus of this study was to examine the effect of makeup and masks on young adults, however the six sets of faces of older adult women were included to prevent observers

from identifying a specific pattern (i.e., all young adults) and responding systematically during the experiment.

Procedure

Four versions of the experiment were created such that each observer only estimated the age of each given identity once, and all conditions of the identities (i.e., no makeup and no facemask / makeup and facemask / makeup and no facemask / no makeup and facemask) were counterbalanced across versions. Therefore, all observers estimated the ages of a total of 28 different identities comprising seven trials from each of the four condition. The experiment was hosted on an online testing suite Testable (Rezlescu et al., 2020) with a restriction enabling observers to run the experiment on a laptop or computer, but not their phone or tablet. Observers were assigned randomly to one of the four versions. The experiment was calibrated to the size of the screen of the device using a standard approach in Testable. When the study commenced, observers were provided with instructions to estimate age as accurately as possible. Each trial began with a centre fixation cross for 800ms, followed by the display of the face image for 2000ms. When this time elapsed, the image was removed from view and replaced with a response box, observers were prompted to enter a two-digit response using the number pad on the keyboard. There was no time pressure to enter their response, and observers pressed 'enter' when they were ready to proceed. All images were presented in random order which was computer generated and different for each participant.

<div align="center">

RESULTS

</div>

Data Preparation

Estimation Bias was measured by computing the estimation error by subtracting the estimated age from the actual age. These values were then averaged for each condition to create four average estimations for each observer. This was used to determine whether the faces were *overestimated* or *underestimated* for the four different conditions[1]. Only the 22 trials comprising age estimates for the young adults aged 18 to 21 were included in these calculations and in the following analysis.

Individual Differences in Estimation Bias

First, we explored the range of age estimation bias on an individual level. When combining all conditions, this revealed a range of bias from an underestimation of - 2.3 years and an overestimation of up to 15 years. These data show that only 12% of observers (8/68) were accurate within 1 year (+/-) of the actual age, while 78% (53/68) of observers overestimated by more than 1 year, and 10% (7/68) underestimated target age by more than a year. We also explored individual data for all four conditions separately. The distribution of scores are illustrated in Figure 2 and show a broad range of accuracy in age estimation. Of interest is the increase in number of observers overestimating age when women were presented with a facemask, makeup or both compared to the number of observers who underestimated the age of women *without* a facemask or makeup. Specifically, in the no mask/no makeup condition 16% (11/68) of observers underestimated the age of women by more than 1 year, compared to only 4% (3/68) in the mask/makeup condition and similarly in the mask/no makeup, and 6% (4/68) in the no mask/makeup. Furthermore, the largest average error made across all conditions was 20 years by two observers both in the mask and no-

---

[1] *Absolute Error*, which reflects average deviation from actual age without direction, was also calculated and analysed separately for completion (See Online Supplementary Materials).

makeup condition. Taken together, this indicates a greater tendency to overestimate the ages of faces wearing a facemask, makeup, or both.

Estimation Bias

Next, we compared age estimation bias for the four conditions. Data screening revealed four participants whose scores on one condition was identified as an outlier because their value was 3 SDs above or below the mean. Inspection of their data did not suggest that these extreme scores were not due to in adherence of task instruction but that the participants were generally poorer at estimating the ages of the faces. While these data points are still valid, such extreme scores can have an undue influence on the analysis, therefore we decided to winsorize the values to the next highest estimate within the condition (Reifman & Keyton, 2010). These data are illustrated in Figure 2 and show that the ages of faces in all conditions were generally overestimated, however faces without any makeup or a facemask produced the least deviation from actual age (+ 2 years) compared to faces wearing a mask and makeup (+ 4 years). Furthermore, Normal QQ-Plots suggested violation of normality in all conditions due to a slight positive skew most pronounced in the mask and makeup condition. Consequently, the data were transformed using an Aligned Rank Transformation (Kay & Wobbrock, 2019; Wobbrock et al., 2011;) using ARTool package in R (version 4.0.3). The Aligned Rank Transformation provides an effective non-parametric approach for multifactorial analysis which, in contrast to other commonly used non-parametric approaches, can effectively handle interactions (for a detailed comparison of approaches see Wobbrock et al., 2011).

To formally analyse these differences, a 2 (Face covering: mask, no mask) x 2 (Face Cosmetics: makeup, no makeup) non-parametric repeated-measures factorial ANOVA using the Aligned Rank Transformed data revealed a main effect of Face Covering, $F(1, 201) =$

22.06, $p < 0.001$, $\eta p^2 = 0.10$, whereby the ages of faces without a facemask were estimated more closely to their actual age. The analysis also revealed a main effect of Face Cosmetics, $F(1, 201) = 8.05$, $p = 0.005$, $\eta p^2 = 0.04$, such that the ages of faces without makeup were estimated more accurately. An interaction between the two factors was also found, $F(1, 201) = 3.92$, $p = 0.049$, $\eta p^2 = 0.02$. To analyse this interaction, non-parametric Wilcoxon signed-rank t-tests were performed with Rank-Biserial Correlation Coefficients reported as effect sizes. These results are summarised in Table 1 and suggest that both facemasks and makeup have an impact on estimation bias with small to medium effect sizes, however a combination of the two is not additive.

Correlations for Estimation Bias

We also explored whether the age of the observer was correlated with the estimated age of the face for the four different conditions. Due to non-normal distribution, a non-parametric spearman's correlation was performed for all four conditions. This analysis revealed a positive correlation for all conditions, all $r_s \geq 0.34$, all $ps \leq 0.005$, with the exception of faces with no mask and no makeup, $r_s(67) = 0.21$, $p = 0.09$. In summary, older observers overestimated faces in the three manipulation conditions to a greater degree than younger observers, but this pattern did not persist for unmasked faces with no makeup.

DISCUSSION

This is the first study to investigate the effect of facemasks, and the combined effect of wearing a facemask with makeup, on the perception of age of young adult female faces. Faces across all four conditions were perceived as older than their actual age. Overestimation was greater for faces wearing a facemask or wearing makeup, but contrary to our prediction, a combination of both did not impair performance further. The findings also revealed a large range of individual differences in age estimation ability across all conditions.

In line with previous findings, 18-to 21-year-olds were perceived to be older than their actual age by at least 1 year (Thorley, 2019; Willner & Rowe, 2001) and this upward bias increased with the application of makeup (Russell et al., 2019). The reason for this upward bias effect of makeup on young faces is not fully understood, and contrasts with the more youthful appearance of older people wearing makeup (Russell et al., 2019). Although older faces appear more youthful because of feature size manipulation (e.g., making eyes appear larger), increasing feature contrast, and enhancing skin homogeneity (Russell et al, 2019), enhancing these features in younger women does not reduce their perceived age. Rather, it has been suggested that makeup in younger women acts as a visual contextual cue activating judgements influenced by beliefs about social norms, specifically that makeup use is associated with adulthood (Russell et al., 2019).

In the present study, we also examined the effect of facemasks on age estimation accuracy. Masked faces were estimated to be approximately 4 years older than their actual age compared to 2 years older when unmasked. We note that the average error of 4 years falls within the range of error found in some other studies (e.g., Clifford et al., 2018; Dehon & Brédart, 2001; Sõrqvist et al., 2007), and by comparison the effect of facemasks on age estimation may not appear large. However, these studies differ methodologically (for example, use of time pressure versus self-paced, accessories on face and hair, average age of perceivers) and consequently the tasks also differ in difficulty. Therefore, such direct comparisons of error across studies offer limited insight.

Although previous research suggests that the eye region is an important facial feature for accurate age judgements (Jones & Smith, 1984), disrupting the processing of other facial features while keeping the eye region intact also impaired age estimation accuracy in this study. However, this raises the question of whether the pattern found here is driven by an individual facial feature obscured by the mask, a combination of features, or the disruption of

the central triangle of the face. Few studies have examined the role of individual features in age perception (Jones & Smith, 1984; Liao et al., 2020). Of these, it was found that children aged 3 to 9 made more errors during the ranking of adult faces by age when the eye region was masked, compared to when the faces were unmasked, or when a combination of features were masked (mouth and chin, or face shape) (Jones & Smith, 1984). The masked nose and cheek condition also produced more errors than the other conditions, though not to the same extent as the masked eye region. In a recent eye-tracking study, observers directed their gaze towards the central triangle of the face both when viewing adult faces freely and when tasked to estimate ages, however in the tasked condition, eye gaze was also redirected towards the lower part of the face (Liao et al. 2020). Both these studies suggest that the bottom half of the face is important for age estimation, and our findings add further support. The present study is a first step for demonstrating the impact of facemasks on age perception, however the theoretical mechanism underlying this effect is beyond the scope of this paper and a systematic examination of different facial features on age perception is an important avenue for further investigation.

We also examined the combined effect of makeup and facemasks. Results indicate that wearing makeup or a facemask individually increased the upward bias by an average of 1.2 and 1.7 years, respectively. However, estimation bias did not increase further when facemasks and makeup were combined. This finding is contrary to our prediction that wearing a facemask *and* applying makeup would increase error bias even further given the pertinent role of makeup to the eye-region compared to other facial features (Russell et al., 2019). To increase ecological validity, the style and intensity of makeup was left to participants to apply in a manner that was natural for them to wear to a formal event. For this reason, we did not control for the intensity of makeup applied to faces. Although research has not found a difference in age estimation bias when using different makeup intensities (Russell

et al., 2019), it is plausible that makeup intensity to the eye region, when paired with a facemask, could influence age estimation differently due to a reduction on reliance from other visual cues for age judgements.

This study also explored age estimation ability at an individual level. Findings demonstrate that there is a broad range in ability, while some individuals were fairly accurate in their estimates (within =/- 1 year), others provided estimates with a deviation of up to 20 years over the target's actual age. Furthermore, older observers overestimated faces to a greater degree when the faces were wearing a mask, makeup, or both, but this pattern did not reach significance when the face was without makeup or mask. These findings support previous research suggesting that the perceiver's age influences age estimation from faces in two ways. First, studies have reported a general decline in age estimation accuracy linked to cognitive ability for older participants compared to younger participants (Voelkle et al., 2012). Secondly, an own-age bias has also been previously recorded, demonstrating an advantage for estimating ages of faces similar to that of the observer (Moyse et al, 2012; Willner & Rowe, 2001) but this pattern is not consistent across studies (Burt & Perrett, 1995). Further exploratory analysis for perceiver age bias in the present study is provided in the supplementary materials, this analysis indicates that perceivers over the age of 34 produced greater error in their estimates of 4.4 years compared to 1.8 years by perceivers under the age of 34 across all conditions. As this study did not set out to investigate own-age bias and therefore did not include the older target faces for a comparison group, it is not possible to conclude whether perceiver age bias found here is due to age related decline or an effect of own-age bias.

Other group characteristics may also account for some of the variance recorded in the present study, namely gender and ethnicity of stimulus and participants. Although these characteristics have not been extensively documented, existing findings suggest that

Caucasian participants perform better when evaluating Caucasian faces (Dehon & Brédart, 2011; Thorley, 2019), but this own-ethnicity bias does not persist for African participants (Dehon & Brédart, 2001). Few studies have examined own-gender bias in age estimation from faces, and those which have did not find evidence to support this (Dehon & Brédart, 2001; Voelkle et al., 2021). The present study did not examine gender and ethnicity bias, however this may warrant further investigation.

Despite these findings, there are some limitations to consider. Only female faces were used to examine the interaction of makeup and facemasks. Some studies report an advantage for accuracy of male faces compared to female faces (Dehon & Brédart, 2001; Voelkle et al., 2012) which may be due to women being more likely to apply cosmetics and makeup to their faces (87% of women compared to 7% of men reported wearing makeup at least once in a government survey; Waldersee, 2019). It is therefore not known whether facemasks and makeup affect the perceived age of men and women differently. Additionally, to control for extraneous variables we did not include the rest of the body in this experiment. While faces have been found to provide the strongest cue for age leading to least errors (Cattaneo et al., 2009), additional visual cues from the body, such as body height and shape, could be used by the assessor to provide a more accurate estimate.

Finally, this study focussed on young women aged 18 to 21, although it is not inconceivable to expect similar patterns with 17-year-old women considering the ages of 13- and 16-year-old girls also tend to be overestimated under 'normal' conditions (Willner & Rowe, 2001). However, face age estimation research with adolescence remains scarce and further work examining the effect of facemasks and makeup with adolescent faces is warranted. The findings with the 18-to 21-year age group are relevant still. Although the most common minimum purchasing age in the European region is 18 years old, minimum age restrictions vary widely across countries and in some cases also depend on the item of sale

(for example, beverage type for alcohol sales). Specifically, for example, in Norway and Sweden, the age limit for beer and wine is 18, but rises to 20 for sales of spirits (Kadiri, 2014), and in some states of the United States of America, Sri Lanka, and Egypt, the law restricts sale of alcohol to individuals aged 21 and over (International Alliance for Responsible Drinking, 2020).

In conclusion, these findings offer some practical implications, particularly in the context of age estimation in the sales of restricted items. Although the *'Challenge 25'* policy is in place in the United Kingdom to encourage sales personnel to verify identification for individuals appearing under the age of 25, many countries do not implement a similar retail sales strategy. Even with such a policy in place, the wide range in estimation ability with some individuals (specifically, 12% on average for neutral faces and between 13% and 20% on average for faces wearing a makeup, a facemask, or both) who over-estimated age by 7 or more years, suggests that a percentage of young women may still be incorrectly perceived to be older than 25. The *'Challenge 25'* policy therefore goes some way to mitigate risk of sales to minors but does not eliminate the potential for decisional errors among those whose age estimation skills are poor, and the potential for error increases further when faces are partially obscured or transformed with the application of makeup.

Finally, the range of individual ability found here also indicates broader implications extending to settings which require security personnel to make age judgements. For example, for the classification of children and adults of undocumented refugees by immigration officers or for the identification of minors in the assessment of online illicit content by internet safety officers. In such situations, it may be beneficial to delegate such tasks to groups of people who are proficient in the task, though, further work to understand individual differences is needed and we hope that these findings will prompt further research in this field.

REFERENCES

Burt, D. M., & Perrett, D. I. (1995). Perception of age in adult Caucasian male faces: computer graphic manipulation of shape and colour information. *Proceedings of the Royal Society of London. Series B: Biological Sciences, 259,* 137-143.

Cattaneo, C., Ritz-Timme, S., Gabriel, P., Gibelli, D., Giudici, E., Poppa, P., Nohrden, D., Assmann, S., Schmitt, R., & Grandi, M. (2009). The difficult issue of age assessment on pedo-pornographic material. *Forensic Science International, 183,* e21– e24. https://doi.org/10.1016/j.forsciint.2008.09.005

Clifford, C., Watson, T. L., & White, D. (2018). Two sources of bias explain errors in facial age estimation. *Royal Society Open Science*, *5*, 180841. https://doi.org/10.1098/rsos.180841

Dayan, S. H., Cho, K., Siracusa, M., & Gutierrez-Borst, S. (2015). Quantifying the impact cosmetic make-up has on age perception and the first impression projected. *Journal of drugs in dermatology: JDD, 14,* 366-374.

Dehon, H., & Brédart, S. (2001). An "Other-Race" Effect in Age Estimation from Faces. *Perception, 30,* 1107–1113. https://doi.org/10.1068/p3122

Department of Health and Social Care. (2020, July 14). Face coverings: when to wear one, exemptions, and how to make your own. *GOV.UK.* Retrieved November 9, 2021, from https://www.gov.uk/government/publications/face-coverings-when-to-wear-one-and-how-to-make-your-own/face-coverings-when-to-wear-one-and-how-to-make-your-own

Egan, V., & Cordan, G. (2009). Barely legal: Is attraction and estimated age of young female

faces disrupted by alcohol use, make up, and the sex of the observer? *British Journal*

of Psychology, 100, 415–427. https://doi.org/10.1348/000712608x357858

Fink, B., Grammer, K., & Matts, P. J. (2006). Visible skin color distribution plays a role in

the perception of age, attractiveness, and health in female faces. *Evolution and*

Human Behavior, 27, 433-422. https://doi.org/10.1016/j.evolhumbehav.2006.08.007

George, P. A., & Hole, G. J. (1995). Factors influencing the accuracy of age estimates of

unfamiliar faces. *Perception*, *24*, 1059-1073. https://doi.org/10.1068/p241059

Henss, R. (1991). Perceiving Age and Attractiveness in Facial Photographs. *Journal of*

Applied Social Psychology, 21, 933– 946. https://doi.org/10.1111/j.1559-

1816.1991.tb00451.x

International Alliance for Responsible Drinking (IARD). (2020). *Minimum legal age*

limits. Retrieved November 9, 2021, from https://iard.org/science-

resources/detail/Minimum-Legal-Age-Limits.

Jones, G., & Smith, P. K. (1984). The eyes have it: Young children's discrimination of age in

masked and unmasked facial photographs. *Journal of Experimental Child Psychology,*

38, 328–337. https://doi.org/10.1016/0022-0965(84)90129-2

Kadiri, H. (2014). *Alcohol purchase age limits in Europe.* Europeans Alcohol Policy

Alliance. Retrieved November 9, 2021, from

https://www.eurocare.org/media/GENERAL/docs/reports/2014alcoholpurchaseagelim

itsineurope.pdf

Kay, M., & Wobbrock, J. (2019). ARTool: Aligned Rank Transform for Nonparametric

Factorial ANOVAs. R package version 0.10. 6.9000 (2019).

Korichi, R., & Pelle-de-Queral, D. (2008). Why women use makeup: Implication of psychological. *Journal of Cosmetic Science, 59,* 127-137. https://doi.org/10.1111/j.1468-2494.2008.00452_3.x

Lau, J.T., Griffiths, S., Choi, K.-C. & Lin, C. (2010) Prevalence of preventive behaviors and associated factors during early phase of the H1N1 influenza epidemic. *American Journal of Infection Control, 38,* 374-380. https://doi.org/10.1016/j.ajic.2010.03.002

Liao, D., Ishii, L. E., Chen, J., Chen, L. W., Kumar, A., Papel, I. D., … & Ishii, M. (2020). How Old Do I Look? Exploring the Facial Cues of Age in a Tasked Eye-Tracking Study. *Facial Plastic Surgery & Aesthetic Medicine, 22,* 36– 41. https://doi.org/10.1089/fpsam.2019.29001.lia

Moyse, E. (2014). Age estimation from faces and voices: a review. *Psychologica Belgica, 54*, 255-265. https://doi.org/10.5334/pb.aq

Moyse, E., Beaufort, A., & Brédart, S. (2014). Evidence for an own-age bias in age estimation from voices in older persons. *European Journal of Ageing*, *11*, 241–247. https://doi.org/10.1007/s10433-014-0305-0

Mulhern, R., Fieldman, G., Hussey, T., Leveque, J.-L., & Pineau, P. (2003). Do cosmetics enhance female Caucasian facial attractiveness? *International Journal of Cosmetic Science, 25*, 199–205. https://doi.org/10.1046/j.1467-2494.2003.00188.x

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*, 493–504. https://doi.org/10.1111/j.0021- 9029.2006.00016.x

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

Reifman, A., & Keyton, K. (2010). Winsorize. In N. J. Salkind (Eds.), *Encyclopaedia of Research Design* (pp. 1636-1638). Sage.

Rezlescu, C., Danaila, I., Miron, A., & Amariei, C. (2020). More time for science: Using Testable to create and share behavioral experiments faster, recruit better participants, and engage students in hands-on research. *Progress in Brain Research, 253,* 243- 262. https://doi.org/10.1016/bs.pbr.2020.06.005

Rhodes, M. G. (2009). Age estimation of faces: a review. *Applied Cognitive Psychology, 23,* 1-12. https://doi.org/10.1002/acp.1442

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

Russell, R., Batres, C., Courrèges, S., Kaminski, G., Soppelsa, F., Morizot, F., & Porcheron, A. (2019). Differential effects of makeup on perceived age. *British Journal of Psychology, 110,* 87–100. https://doi.org/10.1111/bjop.12337

Thorley, C. (2018). Enhancing individual and collaborative eyewitness memory with category clustering recall. *Memory*, *26*, 1128- 1139. https://doi.org/10.1080/09658211.2018.1432058

Thorley, C. (2020). How old was he? Disguises, age, and race impact upon age estimation accuracy. *Applied Cognitive Psychology, 35,* 460-472. https://doi.org/10.1002/acp.3744

Voelkle, M. C., Ebner, N. C., Lindenberger, U., & Riediger, M. (2012). Let me guess how old you are: Effects of age, gender, and facial expression on perceptions of age. *Psychology and Aging, 27,* 265–277. https://doi.org/10.1037/a0025065

Waldersee, V. (2019, February 18). *One in twenty British men now wear makeup*. YouGov. Retrieved March 12, 2021, from https://yougov.co.uk/topics/lifestyle/articles-reports/2019/02/18/one-twenty-british-men-now-wear-makeup

Watson, T. L., Otsuka, Y., & Clifford, C. W. G. (2016). Who are you expecting? Biases in face perception reveal prior expectations for sex and age. *Journal of Vision*, *16*, 1-9. https://doi.org/10.1167/16.3.5

Willner, P., & Rowe, G. (2001). Alcohol servers' estimates of young people's ages. *Drugs: Education, Prevention & Policy, 8,* 375–383. https://doi.org/10.1080/09687630010019299

Wobbrock, J. (2011). The aligned rank transform for nonparametric factorial analyses using only anova procedures. *CHI '11: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 143–146. https://doi.org/10.1145/1978942.1978963

APPENDIX

Figure 1. An example of one of the faces depicted with the four conditions of facemask (bottom) and makeup (right). The woman in the picture is twenty years old.
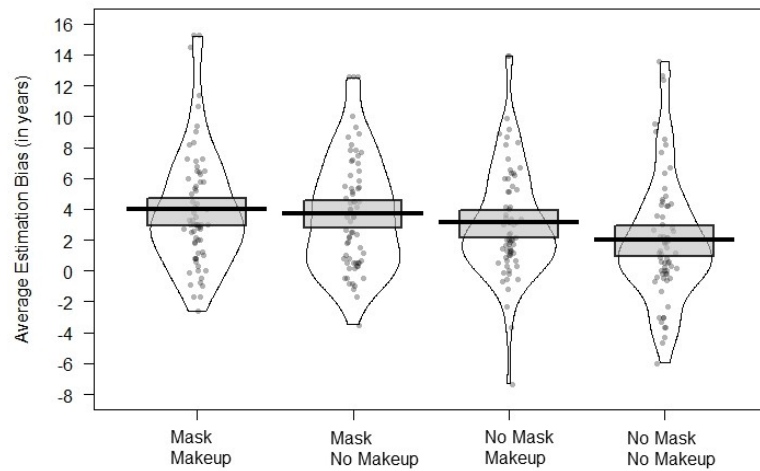
Figure 2. Observers' mean age estimation bias in years and distribution of raw scores when estimating the age of faces with or without a facemask, and with or without makeup. Positive scores reflect an overestimation of age and zero reflects no deviation from actual age. Band around the mean represents Confidence Intervals. Note that the data here represents the non-transformed means and data after outliers have been winsorized.

Table 1.

Summary of Wilcoxon Signed-Rank Post Hoc Comparisons with Rank-Biserial Correlation

Coefficients by Participant Reported as Effect Sizes

		W	*p*	Rank-Biserial Correlation
Mask	Makeup vs. No Makeup	1276.5	0.184	0.19
No Mask	**Makeup vs. No Makeup**	**1622**	**0.006**	**0.38**
Makeup	Mask vs. No Mask	1549	0.022	0.32
No Makeup	**Mask vs. No Mask**	**1750**	**<0.001**	**0.49**

Note. Bonferroni adjustment was used for multiple comparisons (α is 0.05/4 = 0.0125).
Comparisons in bold denote significance (p < 0.0125)