HHS Public Access
Author manuscript
Acta Psychol (Amst). Author manuscript; available in PMC 2022 September 01.

Published in final edited form as:
Acta Psychol (Amst). 2021 September ; 219: 103385. doi:10.1016/j.actpsy.2021.103385.

The effect of aging on facial attractiveness: An empirical and computational investigation

Dexian He[a,b], **Clifford I. Workman**[b,c], **Yoed N. Kenett**[d], **Xianyou He**[a,*], **Anjan Chatterjee**[b,c,**]

[a]Key Laboratory of Brain, Cognition and Education Sciences, Ministry of Education, China; School of Psychology, Center for Studies of Psychological Application, and Guangdong Key Laboratory of Mental Health and Cognitive Science, South China Normal University, China

[b]Penn Center for Neuroaesthetics, University of Pennsylvania, Philadelphia, PA 19104, USA

[c]Department of Neurology, University of Pennsylvania, Philadelphia, PA 19104, USA

[d]Faculty of Industrial Engineering and Management, Technion – Israel Institute of Technology, Haifa 3200003, Israel

Abstract

How does aging affect facial attractiveness? We tested the hypothesis that people find older faces less attractive than younger faces, and furthermore, that these aging effects are modulated by the age and sex of the perceiver and by the specific kind of attractiveness judgment being made. Using empirical and computational network science methods, we confirmed that with increasing age, faces are perceived as less attractive. This effect was less pronounced in judgments made by older than younger and middle-aged perceivers, and more pronounced by men (especially for female faces) than women. Attractive older faces were perceived as elegant more than beautiful or gorgeous. Furthermore, network analyses revealed that older faces were more similar in attractiveness and were segregated from younger faces. These results indicate that perceivers tend to process older faces categorically when making attractiveness judgments. Attractiveness is not a monolithic construct. It varies by age, sex, and the dimensions of attractiveness being judged.

Keywords

face perception; age; sex; attractiveness; network science

1. Introduction

Facial attractiveness has an important impact on social interactions. Studies investigating the “beautiful is good” and the complementary “ugly is bad”/“anomalous is bad” effects demonstrate how a person’s physical attractiveness affects beholders’ attitudes, judgments,

[*]**Correspondence:** X. He, South China Normal University, No.55, West of Zhongshan Avenue, Tianhe District, Guangzhou, 510631 China, xianyouhe@163.com. [**]**Correspondence:** A. Chatterjee, University of Pennsylvania, 3710 Hamilton Walk, Goddard Laboratories, Philadelphia, PA 19104 USA, anjan@pennmedicine.upenn.edu.

Declaration of competing interest
The authors declare no potential conflict of interest

behaviors, and brain functioning (Dion, Berscheid, & Walster, 1972; Eagly, Ashmore, Makhijani, & Longo, 1991; Griffin & Langlois, 2006; Hartung et al., 2019; Jamrozik, Oraa Ali, Sarwer, & Chatterjee, 2019; Workman et al., 2021). Attractive people are regarded as good people, while unattractive people are imbued with negative character traits. The age of a person, as a particularly salient feature of faces, influences the perception of attractiveness. The negative consequences of aging on perceived attractiveness are well-documented (Ebner et al., 2018; Foos & Clark, 2011; North & Fiske, 2015). However, are these negative attitudes modulated by the beholder's age and sex? The present study applies empirical and computational methods to further examine age and sex differences in aesthetic assessments of older faces. In addition, we investigated whether this effect varies depending on how attractiveness is queried.

1.1. Age and sex differences in the perception of facial attractiveness

The age and sex of the perceiver, and the age and sex of the face being viewed, play important roles in attractiveness judgments. Older people are generally perceived more negatively in physical terms than younger people (North & Fiske, 2015). For instance, older faces are rated as less attractive in physical appearance than young faces by young, middle-aged, and older perceivers (Ebner et al., 2018; Foos & Clark, 2011). Elderly faces, compared with younger faces, are also perceived as less likeable, distinctive, energetic, growth oriented (Ebner, 2008), and competent (Palumbo et al., 2017). Properties of faces that contribute to attractiveness (e.g., symmetry) are thought to signal desirable aspects of mate quality and "good genes" (Grammer et al., 2003; Rhodes, 2006). Youth is strongly associated with health and fertility, since younger people are expected to have more time to reproduce. These evolutionary accounts offer one reason why many regard young people as more attractive than older people.

However, previous observations on effects of sex and age of the perceiver, and of face sex, have been mixed. In an earlier study, younger and older participants rated younger and older faces on attractiveness and other dimensions (e.g., likeability; Ebner, 2008). Older faces were perceived as less attractive, especially by young raters. Older perceivers rated all the faces as more attractive than young raters. In contrast, Foos & Clark, (2011) reported young and middle-aged perceivers rated younger faces as more attractive than older faces, whereas older perceivers rated faces from three ages quite similarly. As face age increased, women more than men rated older faces as less attractive. Recently, Ebner et al. (2018) extended their earlier research and found that although older and middle-aged perceivers rated faces as more attractive than young perceivers, young perceivers rated young relatively higher than middle-aged and older faces. They also found that age influenced attractiveness more negatively for female than male faces.

These empirical studies do not consider different ways of processing faces: categorization and individuation (Fiske & Neuberg, 1990). Categorical processing is a core component of intergroup bias (Hughes et al., 2019) and older people are often stereotyped negatively (North & Fiske, 2015). For the present study, we applied a network science approach to examine the structural properties of face preference networks for the different groups, in part to determine if older faces are treated more categorically than if their attractiveness is distributed along a continuum.

Several hypotheses about the influence of the perceiver's age in judgments of attractiveness have been proposed. Ebner et al. (2018) suggest that age effects on attractiveness are influenced by different social goals and may reflect in-group biases (i.e., faces from own-age group are perceived as more attractive). Another account suggests that older perceivers have more experience in face processing and thus may develop a more nuanced appreciation of attractiveness for all ages (Foos & Clark, 2011). Individual face preferences are shaped by life experiences (Germine et al., 2015). On this view, older perceivers who, on average, have greater exposure to more and varied faces might harbor a richer set of face prototypes and preferences.

The powerful forces of sexual selection have forged different strategies to increase sex-specific mating and reproductive success in humans (Buss & Schmitt, 2018). Men tend to value physical attractiveness more than women do in choosing mates, whereas women prioritize male social status and resources (e.g., wealth, prestige) more than men (Li & Kenrick, 2006; Rhodes, 2006; Whyte et al., 2021). Women's fertility is limited by age. For men, women's physical features (e.g., youth) provide cues to health and reproductive value. In addition to physical features, resource-relevant information that could potentially provide economic resources and protection for women and their offspring, is also important when selecting mates. Thus, physical attractiveness is theorized to be less relevant to women than it is to men. These mating strategies may also be used in the evaluation of faces. We propose that male perceivers are more likely to be influenced by face age than female perceivers when making attractiveness judgments, especially for female faces.

Overall, we hypothesized that: First, perceived facial attractiveness is influenced by face age (H1). We expected younger faces to be rated more attractive than older faces; Second, the perceiver age affects attractiveness judgments (H2). We expected that older people would be less affected by face age than younger people; Third, perceiver sex affects how age modulates attractiveness judgments (H3). We predicted face age would affect men's ratings for facial attractiveness more robustly than women's, especially for female faces.

1.2. Effects of different aspects of attractiveness on aesthetic judgments

The general notion of attractiveness incorporates different dimensions. Recently, Menninghaus et al. (2019) acquired free association, questionnaire, and semantic differential data to examine the differences between four related concepts: beauty, elegance, grace, and sexiness, and found that the concept of elegance applied to older women and men more than the other descriptors. As a concept whose meaning is influenced by cultural norms, elegance may be less reliant on physical features. In their study, however, participants were asked to categorize the age ranges (e.g., 20–29) that women or men are likely to perceive as beautiful, elegant, and/or sexy, without using actual faces or judgments of specific stimuli, per se. Vision is one of the most important channels of aesthetic judgment and experience, especially for facial beauty. We sought to extend their observation by examining if different aspects of attractiveness judgments apply specifically to faces when these judgments are made by young, middle-aged, and older men and women.

We examined three concepts related to attractiveness: beauty, elegance, and gorgeousness. A recent study reported that beauty, the most used aesthetic concept, is closely related to ‘elegance’ and ‘gorgeous’ in semantic memory networks (Kenett, Ungar, & Chatterjee, 2021). The Oxford English Dictionary defines ‘gorgeous’ as ‘very beautiful and attractive; giving pleasure and enjoyment’. Though “gorgeousness” is not often used in studies of empirical aesthetics, given support from a large sample of participants in Kenett et al.’s research, we selected it as a concept of inquiry and as a contrast to elegance. In addition to these three terms, we also included liking judgments to further assess people’s preferences.

We hypothesized that aging differentially affects judgments of attractiveness, when framed as beauty, elegance, or gorgeousness (H4). We predicted that age effects would be attenuated or even reversed for elegance than for beauty or gorgeousness ratings.

1.3. Network science

Network science provides quantitative methods to investigate complex systems as networks (Newman, 2003; Siew et al., 2019). Networks are comprised of nodes and edges. Nodes represent the basic unit of the system (e.g., persons, words, images) and edges between nodes represent correlations between the nodes (e.g., friendship, semantic relationships, similarity). This approach has been used fruitfully in aging and aesthetics research. Dubossarsky, De Yne, and Hills, (2017) used network methods to investigate how the mental lexicon changes across the life span. The analysis found a U-shaped pattern of development: networks begin rather sparse, move towards more density in midlife, and become sparser again in older age. Semantic networks of older adults are less connected and more segregated than younger adults. Similarly, Wulff, Hills, and Mata, (2018) examined the semantic networks of younger and older adults and found that the networks of older adults showed smaller average degree and longer path lengths than younger adults. Researchers suggest that the age-related differences in semantic network structure are a consequence of having more lived experiences (Siew et al., 2019; Wulff, Hills, & Mata, 2018; Wulff, De Deyne, Aeschbach, & Mata, 2021).

Hayn-Leichsenring et al. (2020) used network methods to investigate relations between verbal descriptors, global image properties and preferences for abstract art paintings. The analysis found that semantic descriptions play an important role in subjective preferences for paintings. Finally, Kenett, Ungar, and Chatterjee, (2021) used network methods to investigate the semantic neighborhoods of the concepts beauty and wellness, and how they change across age and sex. The analyses found both unique similarities and differences in these semantic neighborhoods across the various networks. Semantic neighborhoods become more segregated from each other, as determined by people’s associations, as they got older.

Here, we applied network science to complement our investigation and asked the following questions: 1) Do people process younger, middle-aged, and older faces differently when judging attractiveness? 2) How do the structural properties of face preference networks vary across perceivers’ age, sex, and across different dimensions of facial attractiveness? Network analyses offer us a different way to analyze behavioral data than traditional statistical analyses and provide potentially converging evidence for our claims. Additionally, these networks offer added insights based on the distribution of faces within a larger subjective attractiveness space.

We hypothesized that perceivers view older faces in a more categorical manner than middle-aged and younger faces (H5). We expected older faces would cluster together more closely (viewed as being more similar) than younger and middle-aged faces in the network space. Further, the structural properties of the face preference network vary by perceiver age (H6) and sex (H7), and by different dimensions of facial attractiveness (H8). Given previous studies on the effect of aging on semantic networks (Dubossarsky, De Yne, & Hills, 2017; Kenett, Ungar, & Chatterjee, 2021; Wulff, Hills, & Mata, 2018), we expected network would be more segregated for older people. Further, if age affects men's ratings more than women's, especially when looking at women, we expected faces would better cluster into groups by face age and sex in the male than the female network. Finally, if people view faces differently on three dimensions of attractiveness, we expected different ages and sexes of faces would cluster differently in three networks. For instance, same sex faces may be more tightly connected to same sex faces in the elegance than beauty and gorgeousness network.

2. Methods

2.1. Participants

A total of 191 participants was recruited via Amazon Mechanical Turk to complete an online survey administered through the Qualtrics platform (97 males; age: 46.57 ± 17.12 years; education: 15.28 ± 2.08 years). Using effect sizes of $\eta^2 = 0.137$ computed from data reported in Foos & Clark (2011) that investigated a similar research question, a minimum sample of $n = 32$ participants per age group was expected to provide sufficient power (80%) to detect the interaction between sex and age of face being rated, and sex of rater. Data were excluded from 30 participants: 9 for response times falling outside the mean reaction time ± 2 standard deviations, 12 for failing more than two of three attention catch trials embedded throughout the survey, 1 for choosing not to report their sex, and 6 who were aged between 58 and 59 years (to differentiate middle-aged and older groups more clearly). Finally, 2 were excluded because they acknowledged that their responses were of poor quality. The final sample consisted of $n = 161$ participants (race/ ethnicity: 128 white, 10 Asian, 8 black, 5 Hispanic or Latinx, 1 American Indian, and 9 multiracial), of which 57 were young (36 males; age: 27.32 ± 2.97 years; range 21–33 years; education: 15.23 ± 1.95 years), 43 middle-aged (19 males; age: 47.63 ± 6.82 years; range: 36–57 years; education: 15.37 ± 1.70 years), and 61 older (25 males; age: 65.77 ± 4.24 years; range: 60–76 years; education: 15.33 ± 2.47 years). The age categories used in this study are based on previous studies (Ebner, 2008; Ebner et al., 2018; Foos & Clark, 2011; Voelkle et al., 2012) and official definitions (United Nations, 2019). This study was approved by the Institutional Review Board at the University of Pennsylvania.

2.2. Face stimuli

The stimuli were comprised of 30 sets of faces, including 30 younger (range: 20–29 years), 30 middle-aged (range: 39–55 years), and 30 older faces (range: 60 years or older). We balanced the face sex and race/ ethnicity. Each set of faces consisted of three different ages for the same reference face (see Fig. 1 for sample stimuli and Table 1 for detailed information about the face stimuli).

Face stimuli were selected and generated in the following way: First, 80 middle-aged faces were selected from the Chicago Face Database (CFD; Ma et al., 2015; http://www.chicagofaces.org/), which also provides researchers with information about each face (e.g., race, age, attractiveness). We then used the FaceApp software (https://www.faceapp.com/) to generate 80 sets of younger and older faces based on the middle-aged faces from the CFD.

Second, in order to standardize the stimuli, face images were 1) normalized to inter-pupillary distance using algorithms provided by the OpenCV computer vision library (https://opencv.org/) and facial landmarks provided by the dlib machine learning toolkit (http://dlib.net/); 2) resized and cropped to 345 pixels (width) × 407 pixels (height); 3) placed onto a plain white background using the GIMP 2 software package (https://www.gimp.org/); 4) color corrected (Workman et al., 2021).

Third, an independent sample of $n = 129$ participants (race/ ethnicity: 102 white, 14 black, 6 Hispanic or Latinx, 3 Asian, 3 multiracial and 1 chose not to report), of which 33 were young (23 males; age: 28.82 ± 3.71 years; range: 20–34 years; education: 14.64 ± 2.56 years), 59 middle-aged (25 males; age: 47.05 ± 8.14 years; range: 35–59 years; education: 14.41 ± 2.71 years), and 37 older (11 males; age: 65.00 ± 4.22 years; range: 60–73 years; education: 14.92 ± 2.51 years), was recruited via Amazon Mechanical Turk to rate the computer-generated younger and older faces for attractiveness (how attractive do you find the person in the picture?) and realness (does the picture look like a real person?) on a scale from 1 to 7. Participants were also asked to indicate the age range of the faces (how old do you think the person in the picture is? e.g., 20–29 years). 43 sets of faces were selected based on the following criteria: 1) higher rates of being perceived as younger (20–29 years) and older (age 60 or older); 2) highest mean realness ratings.

Next, an independent sample of $n = 27$ participants (15 males; age: 26.81 ± 3.72 years; range: 22–36 years; education: 18.22 ± 2.64 years) was recruited via Amazon Mechanical Turk to judge whether each face from the three different ages belongs to the same person. Finally, the 30 sets of faces with the most accurate age group ratings were chosen for the main task (accuracy: 0.99 ± 0.005).

2.3. Procedure

In the main task, participants were asked to rate the faces for beauty (how beautiful is this face?), elegance (how elegant is this face?), gorgeousness (how gorgeous is this face?) and liking (how much do you like this face?) on a scale from 1 to 7. Images were presented in randomized order. There was no time limit, so ratings were acquired in a self-paced fashion. Finally, participants responded to a short socio-demographic questionnaire. The experiment lasted approximately 20 min.

2.4. Statistical analysis

We performed linear mixed-effects analyses using the *lme4* package (Bates et al., 2015) in R (R Core Team, 2019) to examine how age and sex of the perceiver, as well as age and sex of the faces being viewed interact to influence different dimensions of attractiveness. We obtained p values for the parameter estimates of each model using Satterthwaite's approximation as employed by the *lmerTest* package (Kuznetsova, Brockhoff, & Christensen, 2017). Post-hoc pairwise comparisons were conducted with the lsmeans package (Lenth, 2016). Below, regression coefficients (β), standard errors (SE), and t values are reported. Plots were generated with the *effects* package (Fox & Weisberg, 2018).

In the linear mixed-effects models, overall attractiveness was calculated by averaging the ratings of beauty, gorgeousness, and elegance. Direct ratings of attractiveness were not collected in the main task. The average value of the three judgments strongly correlated with the attractiveness ratings from the norming study ($r = 0.87$). Pearson correlations between the four items (beauty, gorgeousness, elegance, and liking), the overall attractiveness ratings, and the attractiveness ratings from the norming study (younger and older faces) and Chicago Face Database (middle-aged faces; Ma, Correll, & Wittenbrink., 2015) are reported in the supplementary information (Table S1).

2.5. Network analysis

We estimated three different types of networks: age-based face preference networks, sex-based face preference networks, and dimensions of attractiveness-based face preference networks, and then conducted global network analyses. These networks were estimated and compared using a computational approach developed by Kenett et al. (2014), and extended to analyze psychometric questionnaires (Christensen et al., 2018).

2.5.1. Network estimation—The estimated networks comprised 90 nodes. Each node represents a face. Edges between nodes represent the correlations between faces. The networks were estimated in the following way (Christensen et al., 2018): First, we used the face ratings to estimate age generation-based, sex-based and dimensions of attractiveness-based networks. Data matrices were structured with each column representing a stimulus (face), and each row containing all ratings for faces of a single participant. Thus, each cell represents a rating given by participant i for face j. Second, we computed the correlations between the face ratings using Pearson's correlation. This resulted in a 90 × 90 adjacency matrix where each cell represents the correlation between node i and node j. Finally, to minimize noise and possible spurious associations, we applied the triangulated maximally filtered graph (TMFG; Christensen et al., 2018; Massara, Di Matteo, & Aste, 2016) method. The TMFG method captures the most important relations between nodes and minimizes spurious associations. This approach filters the network by maintaining planarity and retaining a total of 3n - 6 edges. Thus, the resulting networks for the different groups, with 90 variables, can be compared directly because they have the same number of nodes (90 faces) and edges (264 edges). Although weighted and directed edges could provide more information (e.g., the strength and directionality of the relationships; Borodkin et al., 2016), this may also add noise to the structure of the network. Since we are interested in the network structure, our networks are analyzed as unweighted (all weights are treated as equal, i.e., 1) and undirected (symmetrical relations between nodes).

2.5.2. Network analysis—Analyses were performed with the Brain Connectivity Toolbox for Matlab (Rubinov & Sporns, 2010). The clustering coefficient (CC), the average shortest path length (ASPL) and the modularity (Q) measures were calculated for each of the networks (Fortunato, 2010; Newman, 2006; Siew et al., 2019).

The CC measures the extent in which two neighbors of a node will be neighbors themselves. Thus, higher CC reflects to higher overall connectivity in the network (Siew et al., 2019). Previous research shows that the CC denotes the similarity between concepts in semantic networks and the similarity between art paintings in image network (Hayn-Leichsenring et al., 2020). Thus, in face preference networks, the higher the CC, the more similar faces are in their rated attractiveness. The ASPL measures the average shortest distance (i.e., edges) between any pair of nodes in the network. The higher the ASPL, the more spread out a network is. In face preference networks, higher ASPL may denote the distinctness between faces in attractiveness. The CC and ASPL of all networks were evaluated qualitatively against the equivalent parameters in a random network with the same number of nodes and edges (CC_{rand} and $ASPL_{rand}$, respectively). The Q measures how much a network partitions into sub-communities or sub-networks (Fortunato, 2010; Newman, 2006). The higher the Q, the more the network breaks apart into sub-communities. In the face preference networks, sub-communities can be groups of faces similar attractiveness (e.g., attractive, moderate, or unattractive), or age, etc.

Further, to qualitatively identify whether same sex faces tend to be connected to same sex faces in the sex-based networks and dimensions of attractiveness-based networks, the EI homophily index (EI index; E-external, I-internal) was computed using the igraph package (Csardi & Nepusz, 2006) in R. A network demonstrates assortative mixing, also called homophily, if nodes that have many connections tend to associate with other nodes that are similar to them in the same way (Newman., 2002). The EI homophily index is a measure of in- and out-group preference. EI values that are less than zero suggest homophily and greater than zero suggest heterophily (−1 means complete homophily; 1 means complete heterophily). This analysis allowed us to examine whether the face sex influence men ratings for different aged faces more than women, and whether elegance is more closely related to face sex than beauty or/and gorgeousness.

2.5.3. Statistical analysis—To statistically analyze the networks, we applied a bootstrap method (Efron, 1979) using the *SemNeT* package (Christensen & Kenett, 2019) in R to simulate partial networks and then compared the measures across networks. The bootstrapping procedure involves iteratively resampling a dataset to create bootstrapped samples. For each network, half of the nodes were randomly selected to construct partial networks. This method is known as the without replacement bootstrap (Bertail, 1997). This process was iteratively repeated 1000 times. For each of these bootstrapped partial networks we computed its clustering coefficient (CC), average shortest path length (ASPL) and modularity (Q) measures. Finally, we applied independent samples *t*-test analyses for the sex-based networks, and one-way ANOVAs for the age generation-based networks and dimensions of attractiveness-based networks on each of the network measures. Such analyses allow us to examine the potential differences in CC, ASPL, and Q measures between different networks and to identify unique characteristics in one network with respect to the other.

3. Results

3.1. LMEMs Results

3.1.1. Hypothesis 1: perceived facial attractiveness is influenced by face age—Linear mixed models examined the effect of age on facial attractiveness, with overall attractiveness as the dependent variable and face age (Younger | Middle-aged | Older) as a fixed factor. We included random intercepts for stimulus and subject. Face age significantly affected facial attractiveness judgments, with younger faces rated as more attractive than middle-aged faces ($\beta = 0.80$, SE = 0.16, $t(87) = 5.18$, $p < 0.001$), and middle-aged faces as more attractive than older faces ($\beta = 0.73$, SE = 0.16, $t(87) = 4.71$, $p < 0.001$). In addition, younger faces were liked more than middle-aged faces ($\beta = 0.67$, SE = 0.15, $t(94.85) = 4.55$, $p < 0.001$) and middle-aged faces were liked more than older faces ($\beta = 0.51$, SE = 0.15, $t(99.46) = 3.42$, $p < 0.001$).

3.1.2. Hypothesis 2: the perceiver age affects attractiveness judgments—A linear mixed model examined how the aging effect varies cross-generationally, with overall attractiveness as the dependent variable, face age, and perceiver age (Younger | Middle-aged | Older) as fixed factors, random intercepts for stimulus and subject. This model revealed a significant interaction between face and perceiver ages (Table 2; Fig. 2A). Post-hoc pairwise comparisons indicated that, for middle-aged and older faces, a significant difference between age groups was not detected ($p > 0.05$; Table 3). Younger faces were seen as more attractive by younger than by older perceivers, $t(167.2) = 2.39$, $p < 0.05$. Older perceivers were less influenced by the age of the viewed face than middle-aged and younger perceivers.

3.1.3. Hypothesis 3: the perceiver sex affects how age modulates attractiveness judgments—A linear mixed model examined how the aging effect varies as functions of perceiver and face sex, with overall attractiveness as the dependent variable, face age (Younger | Middle-aged | Older), and face and perceiver sex (Female | Male) as fixed factors. Random intercepts for stimulus and subject were included and, for subject, slopes were allowed to vary according to face age. There was a significant interaction between face age, face sex, and perceiver sex (Table 4; Fig. 2B). To better understand this interaction, we conducted post-hoc pairwise comparisons. For older male faces, there was no significant effect of perceiver sex ($p > 0.05$; Table 5). For older female faces, however, ratings from men were significantly lower than those from women raters, $t(168.7) = -2.68$, $p < 0.01$. Age had a stronger influence on men's ratings than women's ratings for female faces.

3.1.4. Hypothesis 4: aging differentially affects judgments of attractiveness, when framed as beauty, elegance, or gorgeousness—A linear mixed model characterized differences in the aging effect as a function of dimensions of attractiveness, with rating score as the dependent variable, dimension of attractiveness (Beauty | Elegance | Gorgeousness) as a fixed factor, random intercepts for stimulus and subject and, for subject, slopes were allowed to vary by dimension of attractiveness. There was a significant interaction between face age and dimension of attractiveness (Table 6 and 7; Fig. 2C). Post-hoc pairwise comparisons indicated that older faces were rated higher on elegance than on beauty ($t(285) = 3.53$, $p < 0.001$) and gorgeousness ($t(285.3) = 9.59$, $p < 0.001$). Additionally, younger faces were rated higher on beauty than on elegance ($t(285) = 6.02$, $p < 0.001$) and gorgeousness ($t(360.9) = 7.74$, $p < 0.001$). In other words, as age increases, elegance in faces seems more apt as a description of attractiveness than beautiful or gorgeous.

3.2. Network analysis

3.2.1. Age generation-based face preference networks—

We estimated the face preference networks for the younger, middle-aged, and older perceivers and computed network measures (Table 8). Networks were visualized with the Cytoscape software (Shannon et al., 2003).

On visual inspection, older faces clustered together more closely (greater similarity in attractiveness) than younger and middle-aged faces and were segregated (more distinct) from younger faces in all three networks (Fig. 3). The cluster of older faces indicates the perceivers tend to judge older faces similarly when making attractiveness judgments. Further, in older perceivers network (Fig. 3C), older faces were segregated into two main communities and farther away from each other in the network space. Older viewers may distinguish older faces with more acuity than younger and middle-aged viewers.

To statistically compare the clustering coefficient (CC), the average shortest path length (ASPL) and the modularity (Q) measures across the networks, we applied a bootstrapping method (Table 9) and then conducted a one-way ANOVA. Results revealed a significant main effect of perceiver age on clustering coefficient (CC), $F(2, 2997) = 45.57$, $p < 0.001$, $\eta^2 = 0.030$. Post-hoc t-test analyses indicated that younger ($t(1998) = 7.31$, $p < 0.001$) and older perceivers ($t(1998) = 6.23$, $p < 0.001$) had significantly higher CC scores than middle-aged perceivers. For younger and older perceivers, a significant difference in CC scores was not detected ($p > 0.05$).

There was a significant main effect of perceiver age on average shortest path length (ASPL), $F(2, 2997) = 18.99$, $p < 0.001$, $\eta^2 = 0.013$. Post-hoc t-test analyses indicated that older perceivers had significantly higher ASPL scores than younger and middle-aged perceivers, $t(1998) = 5.84$, $p < 0.001$; $t(1998) = 6.05$, $p < 0.001$. For younger and middle-aged perceivers, a significant difference in ASPL scores was not detected ($p > 0.05$).

There was a significant main effect of perceiver age on modularity (Q), $F(2, 2997) = 75.61$, $p < 0.001$, $\eta^2 = 0.048$. Post-hoc t-test analyses indicated that middle-aged perceivers had significantly higher Q scores than younger and older perceivers, $t(1998) = 11.08$, $p < 0.001$; $t(1998) = 7.86$, $p < 0.001$. Older perceivers had significantly higher Q scores than younger perceivers, $t(1998) = 2,31$, $p < 0.05$.

These findings are similar to previous studies (Dubossarsky, De Yne, & Hills, 2017; Kenett, Ungar, & Chatterjee, 2021; Wulff, Hills, & Mata, 2018), demonstrating that older perceivers have a more segregated (higher ASPL and Q scores) network than younger perceivers.

3.2.2. Sex-based face preference networks—Next, we estimated sex-based face preference networks and computed the network measures (Table 8). Similarly, on visual inspection, older faces clustered more closely than younger and middle-aged faces (Fig. 4), suggesting they were rated similarly in facial beauty. Moreover, older faces were segregated from younger faces in the networks.

We observed significant differences in how the faces were distributed between the two networks. Faces were better segregated into clusters by age and sex in the male network (Fig. 4B). To qualitatively examine whether faces belonging to the same sex were more closely connected in the male network than female, we computed the homophily measure. Results show that males (EI = − 0.51; 75.38% female faces go to the in-group) were more likely to associate same sex faces than females (EI = − 0.39; 69.70% female faces go to the in-group). This distribution pattern is in line with our LMEM results showing that face age and sex influenced men's ratings more than women's when judging facial beauty.

To statistically compare the networks, we applied a bootstrap method (Table 9) and conducted independent samples *t*-test on each of the network measures. Results revealed significant differences in the clustering coefficient (CC), the average shortest path length (ASPL) and the modularity (Q) measures between female and male networks. The male network had a significantly lower CC scores ($t(1998) = -21.24$, $p < 0.001$, Cohen's $d = 0.95$), higher ASPL scores ($t(1998) = 25.51$, $p < 0.001$, Cohen's $d = 1.14$) and Q ($t(1998) = 4.52$, $p < 0.001$, Cohen's $d = 0.20$) than the female network.

The male network was more organized (lower CC scores) and more segregated (higher ASPL and Q scores) than the female network, suggesting that faces from different ages and sexes were more distinct from each other in terms of facial beauty for men.

3.2.3. Dimensions of attractiveness-based face preference networks—We estimated dimensions of attractiveness-based face preference networks and computed the network measures (Table 8). Again, older faces clustered together and were segregated from younger faces in all three networks (Fig. 5).

We found that older faces were better segregated by sex in the elegance network. To examine whether faces belonging to the same sex were more closely connected in the elegance network than beauty or/ and gorgeousness, the homophily measure was computed. Results show that perceivers were more likely to associate female faces on elegance (EI = − 0.71; 85.61% female faces go to the in-group) than on beauty (EI = − 0.57; 78.41% female faces go to the in-group) and gorgeousness (EI = − 0.44; 71.97% female faces go to the in-group).

Next, we applied the bootstrapping method (Table 9) and then conducted a one-way ANOVA on the three network measures. The results revealed a significant main effect of dimension of attractiveness on the clustering coefficient (CC), $F(2, 2997) = 453.93$, $p < 0.001$, $\eta^2 = 0.23$. Post-hoc *t*-test analyses indicated the elegance network had significantly lower CC scores than beauty and gorgeousness, $t(1998) = -18.89$, $p < 0.001$; $t(1998) = -28.82$, $p < 0.001$. The gorgeousness network had significantly higher CC scores than beauty, $t(1998) = 11.45$, $p < 0.001$.

There was a significant main effect of dimensions of attractiveness on average shortest path length (ASPL), $F(2, 2997) = 538.84$, $p < 0.001$, $\eta^2 = 0.26$. Post-hoc *t*-test analyses indicated the elegance network had significantly higher ASPL scores than beauty and gorgeousness, $t(1998) = 20.30$, $p < 0.001$; $t(1998) = 31.69$, $p < 0.001$. The beauty network had significantly higher ASPL scores than gorgeousness, $t(1998) = 11.00$, $p < 0.001$.

There was a significant main effect of dimensions of attractiveness on modularity (Q), $F(2, 2997) = 318.79$, $p < 0.001$, $\eta^2 = 0.18$. Post-hoc *t*-test analyses indicated the elegance network had significantly higher Q scores than beauty and gorgeousness, $t(1998) = 13.98$, $p < 0.001$; $t(1998) = 24.72$, $p < 0.001$. The beauty network had significantly higher Q scores than gorgeousness, $t(1998) = 10.84$, $p < 0.001$.

The elegance network was less connected (lower CC scores) and more segregated (higher ASPL and Q scores) than the beauty and gorgeousness networks.

4. Discussion

The present study used empirical and computational network science methods to investigate the effect of aging on attractiveness and to examine how this effect is modulated by the perceiver's age, sex, and dimensions used to make attractiveness judgments. Using highly controlled stimuli, and replicating earlier observations, we found that older faces were perceived as less beautiful, elegant, and gorgeous, and they were liked less. Further, young people rated young faces as more attractive than did older perceivers. Older female faces received lower ratings from male perceivers than female perceivers, suggesting that the age of faces influenced men's ratings for attractiveness more robustly than it does for women making ratings; Finally, beauty, elegance, and gorgeousness ratings were affected differently by age. While the ratings for all these attractiveness descriptors diminished with age, elegance was affected least.

We also observed a relative categorical perception of older faces in that they were viewed more similarly to each other (i.e., they clustered closer together) than the other two groups of faces in face preference networks, which could make it easier for older faces to be subject to negative stereotyping. Alternatively, it's also possible that negative biases towards older individuals make people less inclined to distinguish them. Consistent with these interpretations, older faces were more segregated from and located further away from younger faces compared to middle-aged faces in the networks, again suggesting older faces were more distinct from younger faces in facial beauty.

Perceivers showed negative biases towards older faces, rating them as less beautiful, gorgeous, elegant, and liked. Face preferences are regarded as adaptations for mate choice since attractive traits signal mate quality (Grammer et al., 2003; Rhodes, 2006). The human brain may have evolved to favor these traits (Chatterjee, Thomas, Smith, & Aguirre, 2009; Rellecke et al., 2011). Thus, an evolutionary mechanism might enhance perceptual sensitivity towards younger faces. Alternatively, younger people may simply have less exposure to and experience with older faces. Faces of one's own age group are better recognized and remembered than faces of another age group (own-age bias, OAB; Bartlett & Leslie, 1986; Ebner et al., 2013). Either way, older faces were judged as less distinct from each other and treated more categorically when making attractiveness judgments.

Despite commonalities, the structural properties of the networks varied across perceiver age, sex, and dimension of attractiveness. Faces in the older perceivers face preference network were more segregated than those of younger perceivers. As perceiver age increased, older faces were seen as more distinct in attractiveness. These dynamic changes may reflect that our face preferences are updated by experiences and exposures to faces across the lifespan.

Considerable research has demonstrated that environmental factors, including cumulative environmental exposure and different environments, contribute to age differences in human cognition (Siew et al., 2019; Wulff, De Deyne, Aeschbach, & Mata, 2021; Wulff, De Deyne, Jones, & Mata, 2019). Individuals continue to learn as they get older. Older people are assumed to have acquired more knowledge (e.g., broader vocabulary) than younger people, which subsequently leads to the concepts becoming more distant and further apart from each other in their mental representation (Cosgrove et al., 2021; Wulff, De Deyne, Aeschbach, & Mata, 2021). This may account for the pattern observed in the older adults' semantic network and the similar segregated effect in face preference networks. Research on face preferences also emphasizes the substantial role of experience/environmental factors in shaping our notions of attractiveness (Germine et al., 2015). The cumulative exposure to faces has important implications for individual face preferences. Older people have been generally exposed to more faces and have more diverse experiences compared to younger and middle-aged perceivers. Regarding different environments, people interact more with peers in daily life. These cohort effects may contribute to older viewers being less influenced by the age of the viewed face and more discriminating with older faces in attractiveness. Taken together, we propose that differences in face experience may account for the age-related changes in perception of attractiveness that we report. Older people's experiences and preferences cover a greater span of time.

Men, more than women, segregated faces into clusters by age and sex. The homophily analysis also showed that men more than women were likely to associate same sex faces together. Finally, men viewed faces from different ages and sexes as more organized and more segregated, suggesting they make more distinctions between faces when judging facial beauty. These observations confirm the hypothesis that men are more sensitive to features of physical attractiveness than women, they are more likely to treat face attractiveness categorically, and their sensitivity is further pronounced when judging women's faces.

Sex-specific mating strategies might be reflected in these perceptions of facial attractiveness. Men tend to prioritize women's physical attractiveness, healthiness, and youth, which are theorized to ultimately increase reproductive success and off-spring quality. In contrast, women are thought to value men's status and resources more than attractiveness (Li & Kenrick, 2006; Rhodes, 2006). Empirical data also corroborate that these mate preferences translate into actual mating behavior (Conroy-Beam & Buss, 2018; see also Buss & Schmitt, 2018). Such sex differences in preferences for physical appearance are likely important drivers of differences in perceptions of attractiveness between men and women. However, these strategies are confined to theorizing about heterosexual mating contexts. We do not know if these results would generalize to non-heterosexual individuals.

Finally, there was a stronger association of the dimension of elegance with older than younger and middle-aged faces, and with female than male faces. Elegance, as a descriptor of attractiveness, seems to alert people to finer distinctions in attractiveness for older than younger faces. The overall decrease in attractiveness judgments by age is muted for elegance compared to beauty or gorgeousness is consistent with the view that the notion of elegance goes beyond physical attractiveness, and signals non-physical properties (Menninghaus et al., 2019). We speculate that elegance incorporates cultural norms of attractiveness that are not tethered to physical features as tightly as for beauty and gorgeousness.

We extend previous findings for aging effects to different aspects of attractiveness and revealed differences in the processes people use when judging attractiveness of older faces. However, our study has a few limitations. Different effect sizes were observed for the three network measures in face preferences networks. This probably indicates that one data source is better than the other for these psychometric networks. Future studies are needed to replicate and strengthen our findings. In addition, age-related differences may result from generational or/and developmental differences. We suggest that face preference is influenced by face experiences across the lifespan. But it is hard to quantitatively measure individual difference in face experiences. Whether our findings are the effect of specific generational cohorts or actual aging and accumulation of experience is difficult to determine. Our study was also conducted in the US. American culture may disproportionately value youth. Perhaps these aging effects would be mitigated in cultures with different attitudes towards the elderly.

5. Conclusions

In summary, we replicate and extend the basic finding that people are judged to be less attractive as they age. However, attractiveness judgments are modulated by the age of the perceiver, the sex of the perceiver, and the dimensions of attractiveness judgments being made. Older perceivers are less influenced by the age of the viewed face than are the other groups. Men, more than women, distinguish between faces when judging attractiveness, especially when looking at women. Finally, attractiveness is not a monolithic construct. Aging has less of an effect on judgments of elegance compared to beauty and gorgeousness.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding

This work was supported by the China Scholarship Council (D.H.), the SCNU Study Abroad Program for Elite Postgraduate Students and the Innovation Project of Graduate School of South China Normal University (D.H.), the National Institute of Dental & Craniofacial Research of the National Institutes of Health (F32DE029407 awarded to C.I.W.), the National Natural Science Foundation of China (31970984 awarded to X.H.), and the Edwin and Fannie

Gray Hall Center for Human Appearance (A.C.). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Data and materials availability

The data, code, study, and supplementary materials for this article are publicly available at https://osf.io/3qg6v/.

References

Bartlett JC, & Leslie JE (1986). Aging and memory for faces versus single views of faces. Memory & Cognition, 14(5), 371–381. [PubMed: 3773703]

Bates D, Maechler M, Bolker B, & Walker S (2015). Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software, 67(1), 1–48.

Bertail P (1997). Second-order properties of an extrapolated bootstrap without replacement under weak assumptions. Bernoulli, 3(2), 149–179.

Borodkin K, Kenett YN, Faust M, & Mashal N (2016). When pumpkin is closer to onion than to squash: The structure of the second language lexicon. Cognition, 156, 60–70. [PubMed: 27513870]

Buss DM, & Schmitt DP (2018). Mate preferences and their behavioral manifestations. Annual Review of Psychology, 70(1).

Chatterjee A, Thomas A, Smith SE, & Aguirre GK (2009). The neural response to facial attractiveness. Neuropsychology, 23(2), 135–143. [PubMed: 19254086]

Christensen AP, & Kenett YN (2019). Semantic network analysis (SemNA): A tutorial on preprocessing, estimating, and analyzing semantic networks. PsyArXiv. 10.31234/osf.io/eht87

Christensen AP, Kenett YN, Aste T, Silvia PJ, & Kwapil TR (2018). Network structure of the Wisconsin schizotypy scales-short forms: Examining psychometric network filtering approaches. Behavior Research Methods, 50(6), 2531–2550. [PubMed: 29520631]

Conroy-Beam D, & Buss DM (2018). Why is age so important in human mating? Evolved age preferences and their influences on multiple mating behaviors. Evolutionary Behavioral Sciences, 13(2), 127–157.

Cosgrove AL, Kenett YN, Beaty R, & Diaz MT (2021). Quantifying flexibility in thought: The resiliency of semantic networks differs across the lifespan. Cognition, 211.

Csardi G, & Nepusz T (2006). The igraph software package for complex network research. InterJournal Complex Systems, 1695(5), 1–9.

Dion K, Berscheid E, & Walster E (1972). What is beautiful is good. Journal of Personality and Social Psychology, 24(3), 285–290. [PubMed: 4655540]

Dubossarsky H, De Yne S, & Hills TT (2017). Quantifying the structure of free association networks across the life span. Developmental Psychology, 53(8), 1560–1570. [PubMed: 28569517]

Eagly AH, Ashmore RD, Makhijani MG, & Longo LC (1991). What is beautiful is good, but…: A meta-analytic review of research on the physical attractiveness stereotype. Psychological Bulletin, 110(1), 109–128.

Ebner NC, Johnson MR, Rieckmann A, Durbin KA, Johnson MK, & Fischer H (2013). Processing own-age vs. other-age faces: Neuro-behavioral correlates and effects of emotion. NeuroImage, 78, 363–371. [PubMed: 23602923]

Ebner NC (2008). Age of face matters: Age-group differences in ratings of young and old faces. Behavior Research Methods, 40(1), 130–136. [PubMed: 18411535]

Ebner NC, Luedicke J, Voelkle MC, Riediger M, Lin T, & Lindenberger U (2018). An adult developmental approach to perceived facial attractiveness and distinctiveness. Frontiers in Psychology, 9, 561. [PubMed: 29867620]

Efron B (1979). Bootstrap methods: another look at the jackknife. The Annals of Statistics, 7(1), 1–26.

Fiske ST, & Neuberg SL (1990). A continuum of impression formation, from category-based to individuating processes: influences of information and motivation on attention and interpretation. Advances in Experimental Social Psychology, 23(1), 1–74.

Foos PW, & Clark MC (2011). Adult age and gender differences in perceptions of facial attractiveness: beauty is in the eye of the older beholder. The Journal of Genetic Psychology, 172(2), 162–175. [PubMed: 21675545]

Fortunato S (2010). Community detection in graphs. Physics Reports, 486(3–5), 75–174.

Fox J, & Weisberg S (2018). Visualizing Fit and Lack of Fit in Complex Regression Models with Predictor Effect Plots and Partial Residuals. Journal of Statistical Software, 87(9), 1–27.

Germine L, Russell R, Bronstad PM, Blokland GA, Smoller JW, Kwok H, et al. (2015). Individual aesthetic preferences for faces are shaped mostly by environments, not genes. Current Biology, 25(20), 2684–2689. [PubMed: 26441352]

Grammer K, Fink B, Møller AP, & Thornhill R (2003). Darwinian aesthetics: sexual selection and the biology of beauty. Biological Reviews, 78(3), 385–407. [PubMed: 14558590]

Griffin AM, & Langlois JH (2006). Stereotype directionality and attractiveness stereotyping: is beauty good or is ugly bad? Social Cognition, 24(2), 187–206. [PubMed: 17016544]

Hartung F, Jamrozik A, Rosen ME, Aguirre G, Sarwer DB, & Chatterjee A (2019). Behavioural and neural responses to facial disfigurement. Scientific Reports, 9(1).

Hayn-Leichsenring GU, Kenett YN, Schulz K, & Chatterjee A (2020). Abstract art paintings, global image properties, and verbal descriptions: an empirical and computational investigation. Acta Psychologica, 202, Article 102936.

Hughes BL, Camp NP, Gomez J, Natu VS, Grill-Spector K, & Eberhardt JL (2019). Neural adaptation to faces reveals racial outgroup homogeneity effects in early perception. Proceedings of the National Academy of Sciences, 116(29).

Jamrozik A, Oraa Ali M, Sarwer DB, & Chatterjee A (2019). More than skin deep: Judgments of individuals with facial disfigurement. Psychology of Aesthetics, Creativity, and the Arts, 13(1), 117–129.

Kenett YN, Anaki D, & Faust M (2014). Investigating the structure of semantic networks in low and high creative persons. Frontiers in Human Neuroscience, 8(407), 1–16. [PubMed: 24474914]

Kenett YN, Ungar L, & Chatterjee A (2021). Beauty and wellness in the semantic memory of the beholder. Frontiers in Psychology, 12, 696507. [PubMed: 34421747]

Kuznetsova A, Brockhoff PB, & Christensen RHB (2017). lmerTest package: Tests in linear mixed effects models. Journal of Statistical Software, 82(13).

Lenth R (2016). Least-Squares Means: The R Package lsmeans. Journal of Statistical Software, 69(1), 1–33.

Li NP, & Kenrick DT (2006). Sex similarities and differences in preferences for short-term mates: What, whether, and why. Journal of Personality and Social Psychology, 90(3), 468–489. [PubMed: 16594832]

Ma DS, Correll J, & Wittenbrink B (2015). The Chicago face database: A free stimulus set of faces and norming data. Behavior Research Methods, 47(4), 1122–1135. [PubMed: 25582810]

Massara GP, Di Matteo T, & Aste T (2016). Network filtering for big data: Triangulated maximally filtered graph. Journal of Complex Networks, 5(2), 161–178.

Menninghaus W, Wagner V, Kegel V, Knoop CA, & Schlotz W (2019). Beauty, elegance, grace, and sexiness compared. PLoS ONE, 14(6), Article e0218728.

Newman MEJ (2002). Assortative mixing in networks. Physical Review Letters, 89(20), Article 208701.

Newman MEJ (2003). The structure and function of complex networks. Computer Physics Communications, 147(1–2), 40–45.

Newman MEJ (2006). Modularity and community structure in networks. Proceedings of the National Academy of Sciences, 103(23), 8577–8582.

North MS, & Fiske ST (2015). Modern attitudes toward older adults in the aging world: A cross-cultural meta-analysis. Psychological Bulletin, 141(5), 993–1021. [PubMed: 26191955]

Palumbo R, Adams RB Jr., Hess U, Kleck RE, & Zebrowitz L (2017). Age and gender differences in facial attractiveness, but not emotion resemblance, contribute to age and gender stereotypes. Frontiers in Psychology, 8, 1704. [PubMed: 29033881]

R Core Team. (2019). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Version 3.6.2.

Rellecke J, Bakirtas AM, Sommer W, & Schacht A (2011). Automaticity in attractive face processing: Brain potentials from a dual task. Neuroreport, 22(14), 706–710. [PubMed: 21849915]

Rhodes G (2006). The evolutionary psychology of facial beauty. Annual Review of Psychology, 57(1), 199–226.

Rubinov M, & Sporns O (2010). Complex network measures of brain connectivity: Uses and interpretations. NeuroImage, 52(3), 1059–1069. [PubMed: 19819337]

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. Genome Research, 13(11), 2498–2504. [PubMed: 14597658]

Siew CSQ, Wulff DU, Beckage NM, & Kenett YN (2019). Cognitive network science: A review of research on cognition through the lens of network representations, processes, and dynamics. Complexity, 2019, 24.

United Nations. (2019). World Population Ageing. https://www.un.org/en/development/desa/population/publications/pdf/ageing/WorldPopulationAgeing2019-Highlights.pdf.

Voelkle MC, Ebner NC, 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(2), 265–277. [PubMed: 21895379]

Whyte S, Brooks RC, Chan HF, & Torgler B (2021). Sex differences in sexual attraction for aesthetics, resources and personality across age. PLoS ONE, 16(5), Article e0250151.

Workman CI, Humphries S, Hartung F, Aguirre GK, Kable JW, & Chatterjee A (2021). Morality is in the eye of the beholder: The neurocognitive basis of the “anomalous-is-bad” stereotype. Annals of the New York Academy of Sciences, 1494(1), 3–17. [PubMed: 33565114]

Wulff DU, De Deyne S, Aeschbach S, & Mata R (2021). Understanding the aging lexicon by linking individuals’ experience, semantic networks, and cognitive performance. PsyArXiv. 10.31234/osf.io/z3ebt

Wulff DU, De Deyne S, Jones MN, Mata R, & The Aging Lexicon Consortium. (2019). New perspectives on the aging lexicon. Trends in Cognitive Sciences, 23(8), 686–698. [PubMed: 31288976]

Wulff DU, Hills T, & Mata R (2018). Structural differences in the semantic networks of younger and older adults. PsyArXiv. 10.31234/osf.io/s73dp

Fig. 1.

Sample stimuli. Middle-aged faces selected from the Chicago Face Database (Ma et al., 2015) were morphed to appear either younger or older using the FaceApp software package (https://www.faceapp.com/).

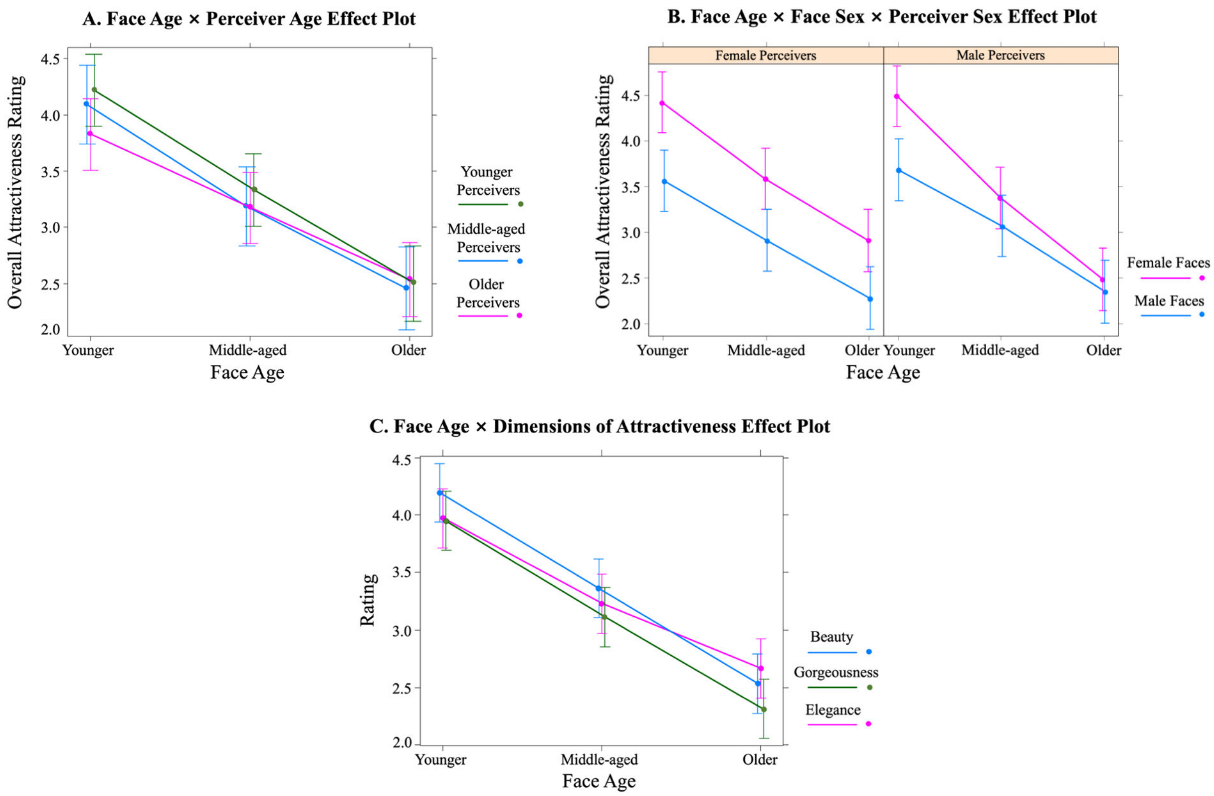

Fig. 2.

A. Effects of face age on overall attractiveness ratings (y-axis) as a function of perceiver age. The dots represent means. The error bars represent 95% confidence intervals. There was a significant interaction between face (x-axis: younger, middle-aged, and older) and perceiver ages (younger perceivers, green line; middle-aged perceivers, blue line; older perceivers, pink line). Older perceivers were less affected by face age than younger and middle-aged perceivers. **B.** Effects of face age on overall attractiveness ratings (y-axis) for female and male faces as a function of perceiver sex. There was a significant interaction between face age, perceiver sex, and face sex (female faces, pink line; male faces, blue line). Men rated older female faces as less attractive than women raters. **C.** Effects of face age on ratings (y-axis) as a function of the dimensions of attractiveness. There was a significant interaction between face age (x-axis) and the dimensions of attractiveness (beauty, blue line; elegance, pink line; gorgeousness, green line). Older faces were rated higher on elegance than on beauty and gorgeousness.

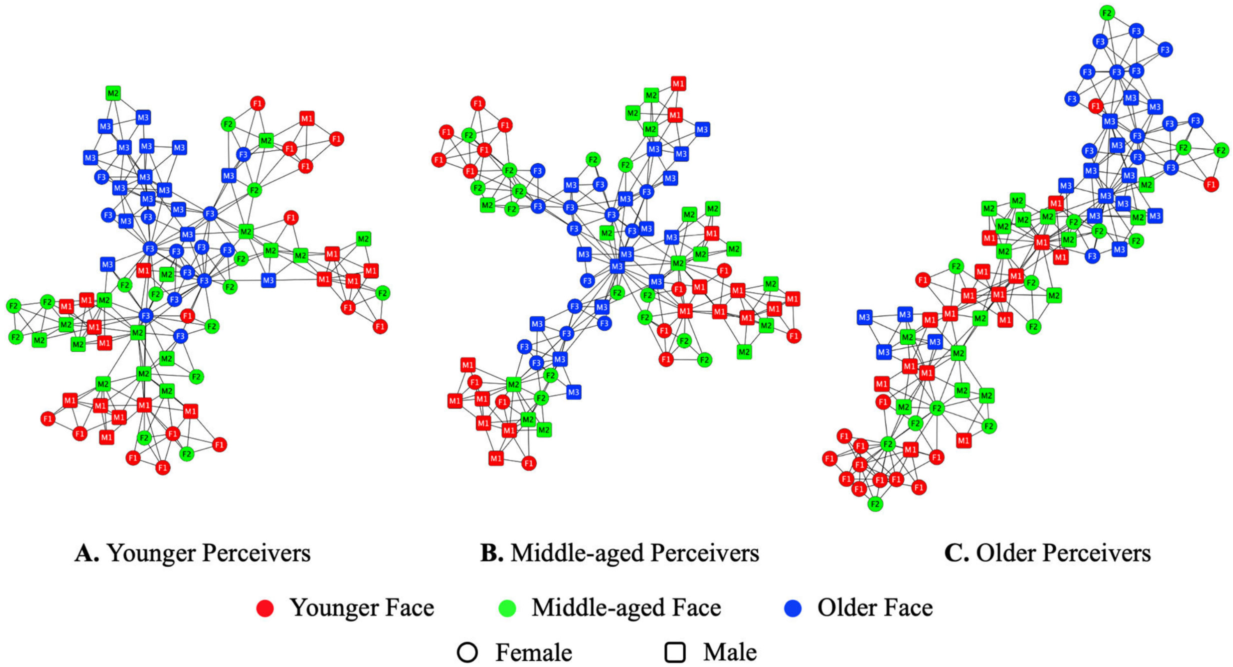

Fig. 3.

2D visualization of the age generation-based face preference networks. Nodes represent the 90 faces. Edges represent symmetrical, binary relations between nodes. Colors represent face age (younger, red; middle-aged, green; older, blue). Shapes represent face sex (female, ellipse; male, round rectangle). Labels denote face sex and age (F1, younger female face; M1, younger male face; F2, middle-aged female face; M2, middle-aged male face; F3, older female face; M3, older male face).

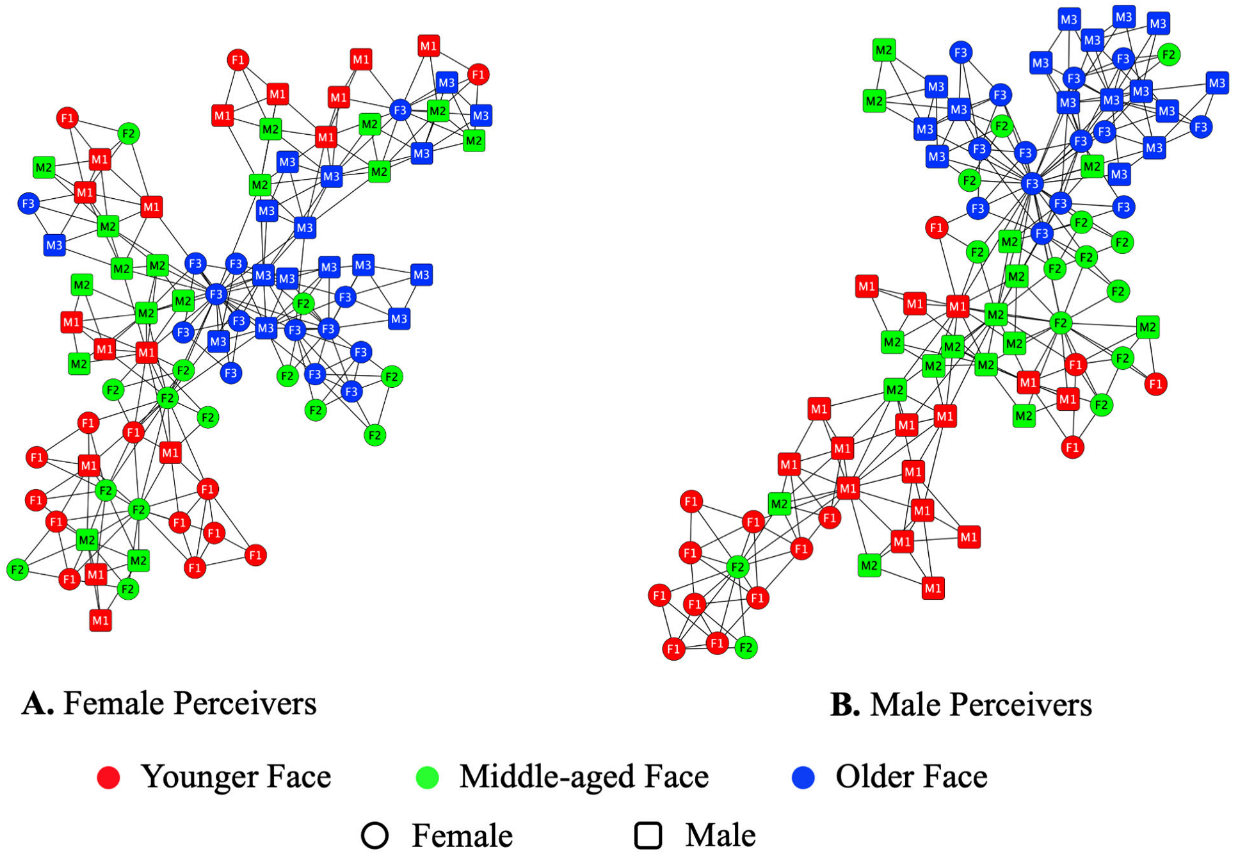

Fig. 4.

2D visualization of the sex-based face preference networks. Nodes represent the 90 faces. Edges represent symmetrical, binary relations between nodes. Colors represent face age. Shapes represent face sex. Labels denote face sex and age (F1, younger female face; M1, younger male face; F2, middle-aged female face; M2, middle-aged male face; F3, older female face; M3, older male face).

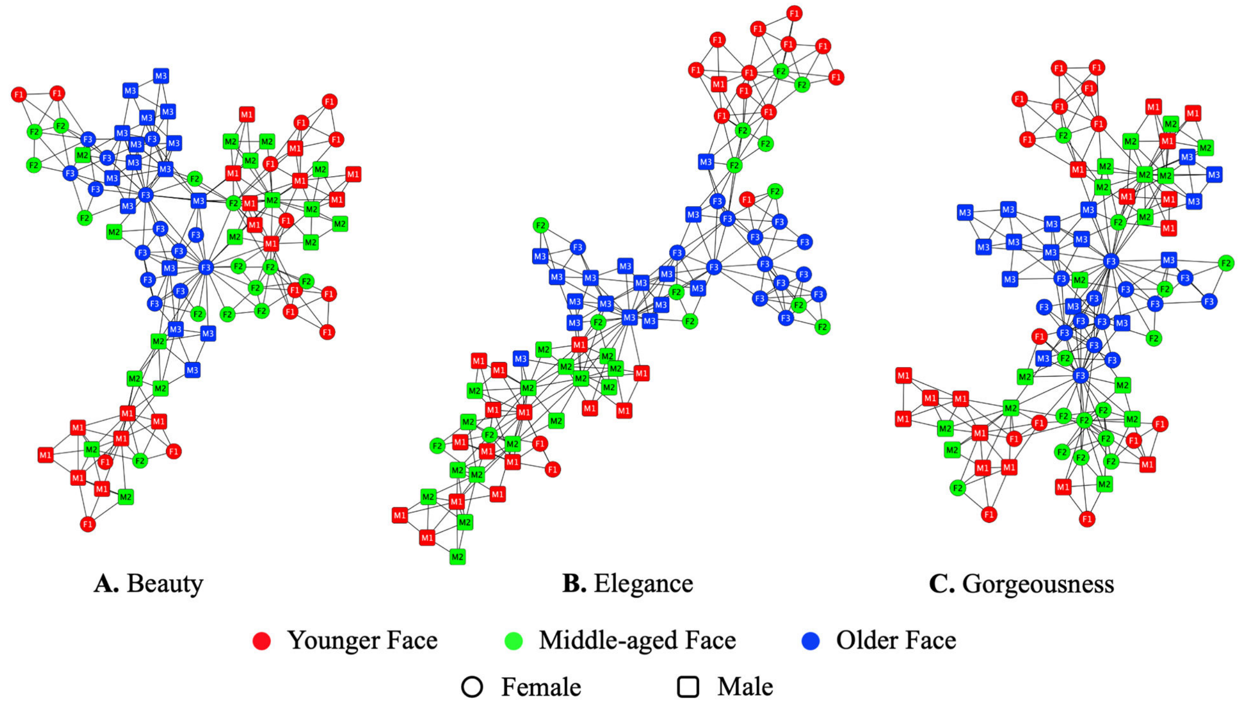

Fig. 5.

2D visualization of the dimension of attractiveness-based face preference networks. Nodes represent the 90 faces. Edges represent symmetrical, binary relations between nodes. Colors represent face age. Shapes represent face sex. Labels denote face sex and age (F1, younger female face; M1, younger male face; F2, middle-aged female face; M2, middle-aged male face; F3, older female face; M3, older male face).

Table 1.

Information about the face stimuli

	Younger faces	Middle-aged faces	Older faces
N	30	30	30
M/F	15/15	15/15	15/15
Age	20–29 *(67.3%)	39.69 (±3.88)	60+ *(79.4%)
Attractiveness	4.50 (0.66)	3.06 (0.58)	3.14 (0.39)
Realness	5.24 (0.43)	Real face	5.57 (0.33)

Note. M - Male; F - Female; Ratings for the middle-aged faces (3.06 ± 0.58) and age information (39.69 ± 3.88 years) were provided by the Chicago Face Database (Ma, Correll, & Wittenbrink., 2015). Information of younger and older faces derives from the results of our face norming tasks.

* On average, 67.3% participants rated the 30 computer-generated younger faces as 20–29 years; 79.4% participants rated the 30 computer-generated older faces as age 60 or older.

Author Manuscript
Author Manuscript
Author Manuscript

Table 2.

Fixed effects from the linear mixed model constructed to examine how the aging effect varies cross-generationally

Fixed effects	β	SE	t value	p value
Intercept	3.19	0.17	18.30	< 0.001
Face age (older)	−0.73	0.16	−4.58	< 0.001
Face age (younger)	0.91	0.16	5.72	< 0.001
Perceiver age (older)	−0.01	0.18	−0.06	0.951
Perceiver age (younger)	0.15	0.18	0.83	0.410
Face age (older) : Perceiver age (older)	0.09	0.05	1.70	0.090
Face age (younger) : Perceiver age (older)	−0.26	0.05	−4.98	< 0.001
Face age (older) : Perceiver age (younger)	−0.10	0.05	−1.95	0.051
Face age (younger) : Perceiver age (younger)	−0.02	0.05	−0.46	0.647

Note. SE, standard error.

Table 3.

Means and standard deviations for overall attractiveness ratings according to face and perceiver age

	Younger face	Middle-aged face	Older face	Overall
Younger perceiver	4.22 (1.55)	3.34 (1.45)	2.51 (1.37)	3.35 (1.62)
Middle-aged perceiver	4.10 (1.74)	3.19 (1.59)	2.46(1.44)	3.25 (1.73)
Older perceiver	3.83 (1.42)	3.18 (1.31)	2.54 (1.25)	3.18 (1.43)
Overall	4.04 (1.56)	3.24 (1.44)	2.51 (1.34)	

Table 4.

Fixed effects from the linear mixed model constructed to examine how the aging effect varies as functions of perceiver and face sex

Fixed effects	β	SE	*t* value	*p* value
Intercept	3.58	0.17	20.97	< 0.001
Face age (older)	−0.67	0.20	−3.36	< 0.010
Face age (younger)	0.84	0.20	4.20	< 0.001
Face sex (male)	−0.67	0.19	−3.46	< 0.001
Perceiver sex (male)	−0.21	0.15	−1.38	0.168
Face age (older) : Face sex (male)	0.04	0.27	0.14	0.889
Face age (younger) : Face sex (male)	−0.19	0.27	−0.68	0.496
Face age (older) : Perceiver sex (male)	−0.22	0.90	−2.44	< 0.050
Face age (younger) : Perceiver sex (male)	0.28	0.08	3.27	< 0.010
Face sex (male) : Perceiver sex (male)	0.36	0.05	6.70	< 0.001
Face age (older) : Face sex (male): Perceiver sex (male)	0.14	0.08	1.76	0.078
Face age (younger) : Face sex (male): Perceiver sex (male)	−0.31	0.08	−4.09	< 0.001

Note. SE, standard error.

Table 5.

Means and standard deviations for overall attractiveness ratings according to face age and to face and perceiver sex

	Female face			Male face			
	Younger face	**Middle-aged face**	**Older face**	**Younger face**	**Middle-aged face**	**Older face**	**Overall**
Female perceiver	4.42(1.58)	3.58(1.54)	2.91(1.49)	3.56(1.63)	2.91(1.51)	2.28(1.34)	2.28(1.66)
Male perceiver	4.49(1.42)	3.38(1.33)	2.48(1.22)	3.68(1.37)	3.07(1.27)	2.35(1.21)	3.24(1.49)
Overall	4.46(1.50)	3.48(1.44)	2.70(1.38)	3.62(1.51)	2.99(1.40)	2.31(1.28)	

Table 6.

Fixed effects from the linear mixed model constructed to examine how the aging effect varies as a function of dimension of attractiveness

Fixed effects	β	SE	*t* value	*p* value
Intercept	3.36	0.13	25.69	< 0.001
Face age (older)	−0.83	0.16	−5.29	< 0.001
Face age (younger)	0.83	0.16	5.32	< 0.001
Dimensions of attractiveness (elegance)	−0.13	0.04	−3.55	< 0.001
Dimensions of attractiveness (gorgeousness)	−0.25	0.03	−7.87	< 0.001
Face age (older) : Dimension of attractiveness (elegance)	0.26	0.03	8.16	< 0.001
Face age (younger) : Dimension of attractiveness (elegance)	−0.09	0.03	−2.85	0.004
Face age (older) : Dimension of attractiveness (gorgeousness)	0.03	0.03	0.88	0.380
Face age (younger) : Dimension of attractiveness (gorgeousness)	0.004	0.03	0.12	0.902

Note. SE, standard error.

Author Manuscript

Author Manuscript

Author Manuscript

Table 7.

Means and standard deviations of attractiveness ratings according to face age and dimension of attractiveness

	Younger face	Middle-aged face	Older face	Overall
Beauty	4.19 (1.61)	3.36 (1.51)	2.54 (1.44)	3.36 (1.66)
Elegance	3.97 (1.67)	3.23 (1.57)	2.67 (1.54)	3.29 (1.68)
Gorgeousness	3.95 (1.66)	3.11 (1.50)	2.31 (1.34)	3.13 (1.65)

Table 8.

Network measures for the sex-based, age generation-based and dimensions of attractiveness-based networks

	Age generation-based networks			Sex-based networks		Dimensions of attractiveness-based networks		
	Younger	Middle-aged	Older	Female	Male	Beauty	Elegance	Gorgeous-ness
CC	0.716	0.711	0.717	0.716	0.725	0.722	0.710	0.722
ASPL	3.471	3.765	4.026	3.457	3.457	3.516	4.111	3.292
Q	0.647	0.670	0.656	0.650	0.635	0.641	0.656	0.651
CC_{rand}	0.046	0.049	0.055	0.046	0.049	0.046	0.049	0.055
$ASPL_{rand}$	2.720	2.735	2.703	2.720	2.734	2.720	2.735	2.703

Note. CC - clustering coefficient; ASPL - average shortest path length; Q - modularity index; CC_{rand} - clustering coefficient of random graph; $ASPL_{rand}$ - average shortest path length of random graph.

Table 9.

Bootstrapping analysis of CC, ASPL, and Q for the sex-based, age generation-based and dimensions of attractiveness-based networks

	Age generation-based networks			Sex-based networks		Dimensions of attractiveness-based networks		
	Younger	Middle-aged	Older	Female	Male	Beauty	Elegance	Gorgeousness
M_{cc}	0.719[a]	0.716[b]	0.719[a]	0.728[a]	0.719[b]	0.717[b]	0.709[c]	0.722[a]
M_{ASPL}	2.766[b]	2.782[b]	2.813[a]	2.641[b]	2.832[a]	2.783[b]	2.947[a]	2.682[c]
M_{Q}	0.535[c]	0.547[a]	0.537[b]	0.527[b]	0.532[a]	0.540[b]	0.553[a]	0.528[c]

Note. M_{CC} - mean clustering coefficient of bootstrapping analysis; M_{ASPL} - mean average shortest path length of bootstrapping analysis; M_{Q} - mean modularity measure of bootstrapping analysis. In each group of networks, identical superscripts are placed next to means that do not significantly differ from one another. Likewise, different superscripts are placed next to means that significantly differed.