How facial aging affects perceived gender: Insights from maximum likelihood conjoint measurement

Daniel Fitousi Ariel University, Israel

Conjoint measurement was used to investigate the joint influence of facial gender and facial age on perceived gender (Experiment 1) and perceived age (Experiment 2). A set of 25 faces was created, covarying independently five levels of gender (from feminine to masculine) and five levels of age (from young to old). Two independent groups of observers were presented with all possible pairs of faces from this set and compared which member of the pair appeared as more masculine (Experiment 1) or older (Experiment 2). Three nested models of the contribution of gender and age to judgment (i.e., independent, additive, and saturated) were fit to the data using maximum likelihood. The results showed that both gender and age contributed to the perceived gender and age of the faces according to a saturated observer model. In judgments of gender (Experiment 1), female faces were perceived as more masculine as they became older. In judgments of age (Experiment 2), young faces (age 20 and 30) were perceived as older as they became more masculine. Taken together, the results entail that: (a) observers integrate facial gender and age information when judging either of the dimensions, and that (b) cues for femininity and cues for aging are negatively correlated. This correlation exerts stronger influence on female faces, and can explain the success of cosmetics in concealing signs of aging and exaggerating sexually dimorphic features.

Introduction

Facial age and gender are crucial for face recognition. These dimensions are among the first to be extracted by observers (Young & Burton, 2017). They are easily estimated, even in unfamiliar faces (Burt & Perrett, 1995; George & Hole, 2000), and are perceived categorically (Allport et al., 1954; Beale & Keil, 1995), such that their corresponding sensory cues are transformed rapidly and automatically into discrete labels (Enlow & Moyers, 1982; O'Neil & Webster, 2011). Recently, there has been a resurgence of interest in these facial dimensions in the domains of psychology (Fitousi, 2017a, 2020; Kloth et al., 2015; O'Neil & Webster, 2011; Schweinberger et al., 2010; Wiese et al., 2013) and computer science (Eidinger et al., 2014; Levi & Hassner, 2015). This is primarily due to the increasing influence of social networks that, along with the cosmetic industry (Etcoff et al., 2011; Russell, 2009) reflects the tremendous importance individuals place on their own and others' facial age and gender. One question that stands out in psychological research concerns the extent to which facial age and gender interact in perception and cognition. Are cues for facial age affecting the way facial gender is perceived? Are cues for facial gender influencing the manner by which we estimate a face's age?

The present study addresses these questions by harnessing a rigorous psychophyscial technique known as maximum likelihood conjoint measurement (MLCM; Ho et al., 2008; Knoblauch & Maloney, 2012). The technique is based on the classic conjoint measurement methodology, originally developed by (Luce & Tukey 1964), and its subsequent elaborations (Krantz et al., 1971). This methodology allows researchers to construct the psychological scales for two (or more) tested dimensions based on a simple paired comparison task. It also allows testing the perceptual independence of the pertinent dimensions. In the current study, the MLCM was applied to the facial dimensions of age and gender. The results strongly suggested that these dimensions are not processed independent of each other. It is shown that cues to age and cues to gender have conjoint contributions to the perception of either of the dimensions. Moreover, the nature of these contributions can be characterized both qualitatively and quantitatively, and they shed light on theoretical and practical questions in social and cognitive psychology.

The perception of facial age relies on numerous shape and surface cues. In the course of human development, structural changes occur in the shape of the skull, facial features, and their configuration (Berry & McArthur, 1986). The configuration of the face continues to change throughout adult life and is characterized by many textural alterations of local and global aspects, such as the eyes and lips (Burt & Perrett, 1995). Both shape cues and texture cues have been shown to be important for age judgments (Burt & Perrett, 1995; George & Hole, 2000; O'Toole et al., 1997; Macke & Wichmann, 2010).

Citation: Fitousi, D. (2021). How facial aging affects perceived gender: Insights from maximum likelihood conjoint measurement. *Journal of Vision*, *21*(12):12, 1–16, https://doi.org/10.1167/jov.21.12.12.

https://doi.org/10.1167/jov.21.12.12 Received February 28, 2021; published November 23, 2021 ISSN 1534-7362 Copyright 2021 The Authors

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

Downloaded from abstracts.iovs.org on 04/25/2024

Shape cues for age consist of morphological changes, such as nose and ear cartilage growing, which increase the size of these features, and changes in other local features, such as the eyes and lips (Burt & Perrett, 1995). Surface cues for age consist of skin texture, such that young faces have soft, firm and smoother skin, and older faces (30 years old and older) are characterized by wrinkling (Berry & McArthur 1986; Burt & Perrett 1995; George & Hole 2000; O'Toole et al. 1997).

Similarly, the perception of facial gender relies on numerous shape and texture cues. Shape cues for gender consist of eyes, eyebrows, the jaw and the face outline (Brown & Perrett, 1993; Bruce & Langton, 1994; Burton et al., 1993), whereas surface cues for gender include skin texture (Russell, 2009), luminance contrast between the eyes and the mouth (Russell, 2009), skin pigmentation (Bruce & Langton, 1994), and skin color (Tarr et al., 2001).

Given that many of the cues for age and gender rely on shared shape and texture cues, one may predict that these two dimensions are dependent in perception. However, the empirical evidence for this hypothesis is rather mixed. Some studies point to perceptual interactions (Cloutier et al., 2014; Fitousi, 2017a), whereas others argue for partial or complete independence (O'Neil & Webster 2011). Barrett and O'Toole (2009) deployed a face gender adaptation paradigm where a bias to classify the gender of a gender-neutral face is opposite to that of an adapting face. They found that gender adaptation effects transferred across age categories, a result that supports the independence of facial age and gender. Similarly, O'Neil and Webster (2011) recorded facial age adaption effects that transferred across gender categories, again pointing to perceptual independence. Quinn and Macrae (2005) applied Garner's speeded classification task (Algom & Fitousi, 2016; Garner, 1974b) to the dimensions of facial age and gender. They found that the classification of faces on gender were interfered by irrelevant variations on facial age, whereas categorization according to facial age was not affected by irrelevant variations on facial gender. Quinn and Macrae (2005) also found that classifications of gender were performed faster with young than old faces, but this difference was more pronounced for female than male faces. Fitousi (2020) reapplied the Garner task to the same dimensions and found them to be completely separable. However, he documented a bias in judgments of gender by which young females and old males were categorized faster than old females and young males. A similar bias has been recorded by Kloth et al. (2015). According to Kloth et al. (2015), the results reflect built-in correlations between the sensory cues for gender and age. In particular, observers took advantage of the presence or absence of smooth skin texture and used it as a clue for facial age.[1]

Taken together, the findings from the reviewed studies are somewhat inconsistent. At the dimensional level, facial age and gender appear as separable entities (Barrett & O'Toole, 2009; Fitousi & Wenger, 2013; Fitousi, 2020; O'Neil & Webster, 2011), but at the level of the constituting features, facial age and gender exhibit specific interactive patterns (Fitousi, 2020; Kloth et al., 2015), with faces of young females and old males being categorized more efficiently than faces of old females and young males. These findings suggest that cues for facial age and cues for facial gender are not represented independent of each other. The current study set to investigate the perceptual mechanisms that lead to such a bias, as well as the exact contributions of objective age and gender cues to perceived age and gender. Stimuli in previous studies were not tightly controlled, making it difficult to know how gradual changes within and across age and gender categories exactly influence each other.

The present study applied the MLCM methodology (Ho et al., 2008; Knoblauch & Maloney, 2012) to facial age and gender. The MLCM technique has been widely applied to classic psychophysical dimensions (for review see, Maloney & Knoblauch, 2020); such as surface gloss and bumpiness (Ho et al., 2008; Qi et al., 2015), color dimensions of hue, chroma, and lightness (Gerardin et al., 2018; Rogers et al., 2016); time (Lisi & Gorea, 2016); and recently to facial dimensions of lightness and ethnicity (Nichiporuk et al., 2018) and facial gender and voice (Abbatecola et al., 2021). The current study expands the scope of applications to additional facial dimensions. The goal is to shed light on the perception of facial age and gender. In Experiment 1, observers compared pairs of faces on perceived gender. In Experiment 2, a new group of observers judged the same set of stimuli with respect to perceived age. The goals were to a) exert tight control over face stimuli, altering their age and gender levels in a parametric fashion, b) investigate how gradual changes within and across age and gender physical categories are translated onto psychological scales, c) test a set of nested models that allow us to quantify the relative contributions of each dimension's physical cues to perceived gender (Experiment 1) and age (Experiment 2), and d) determine the best fitting model, separately for the gender and age tasks.

Methods

Observers

A total of 16 observers participated in this study. They were recruited from the Ariel University pool of participants, and were compensated with a course

Downloaded from abstracts.iovs.org on 04/25/2024

credit. Eight observers participated in Experiment 1 (four males and four females, mean age = 24.6, sd = 6.11 years). A new group of eight observers, who did not participate in Experiment 1, participated in Experiment 2 (four males and four females, mean age = 23.8, sd = 3.15 years). All observers had normal or corrected-to-normal vision. All observers were naive to the goals of the study.

Stimuli and apparatus

The stimuli were two-dimensional color face images positioned at the center of the screen over a black background. The stimuli were created with Singular Inversions FaceGen Modeller 3.2 (Inversions, 2008). This software incorporates a three-dimensional morphable model of faces to allow the generation and variation of face images along several dimensions such as identity, gender, and emotion (Blanz & Vetter, 1999). Faces generated by FaceGen have been widely used in recent studies (Johnson et al., 2012; Oosterhof & Todorov, 2008). The FaceGen model is based on three-dimensional imaging of 273 males and females from 12 to 67 years old, and of various ethnicities and appropriate controls for age and gender.

The stimuli in the current study were created by first generating a single European identity by moving the "race morphing" slider toward the "European" end of scale. I then changed the desired parameters of age and gender of this identity to produce the faces in the set. All faces in the set were created using the same angle and default setting for lighting, without hair or other external features. In both age and gender scaling, the 'Sync Lock' option was checked to allow synchronized contributions of texture and shape. The

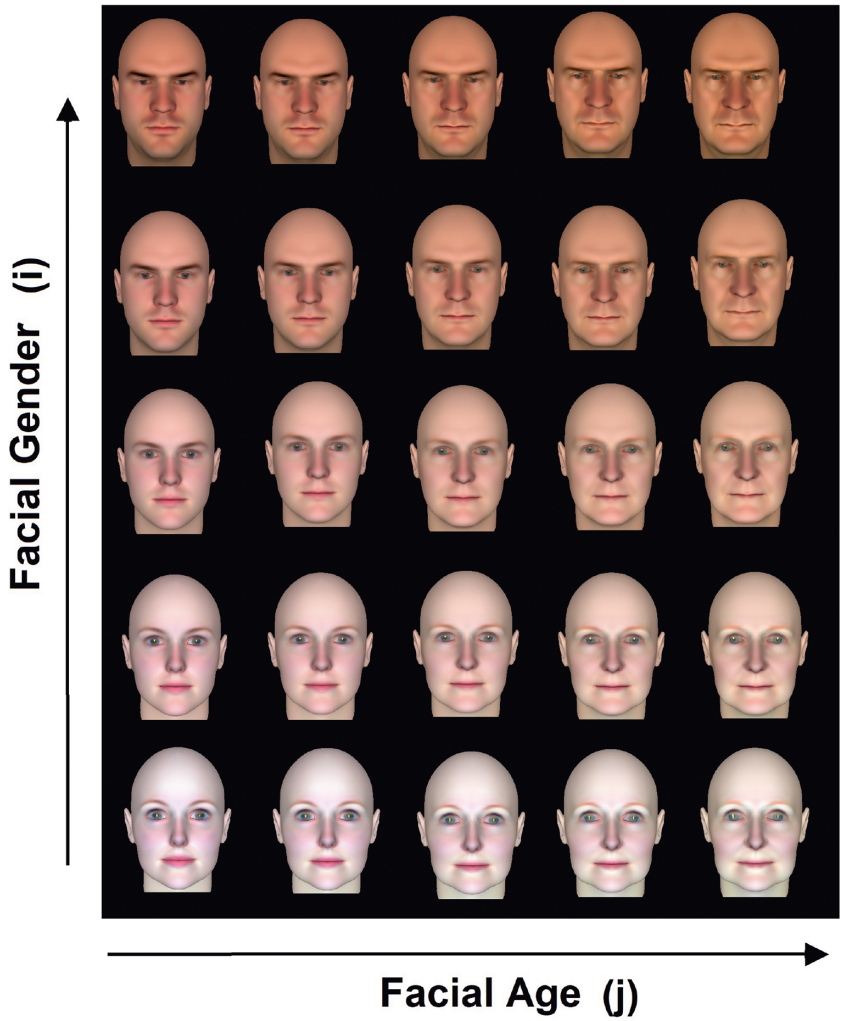

Figure 1. The stimuli set in Experiment 1 and Experiment 2. The faces were created with the FaceGen software by combining five levels of gender and five levels of age. The *i* stand for levels of gender and *j* for levels of age. Note that for a given level of age *j*, physical gender becomes more masculine up the *i*th row. For a given level of gender *i*, physical age becomes older as one moves to the right *j*th columns.

Downloaded from abstracts.iovs.org on 04/25/2024

age and gender group sliders are linear regressions on the model data set. The facial gender and age of each face were altered in a parametric fashion. This was done by first dividing the corresponding age and gender sliders into five equally distanced points, and then, by moving the handles of these sliders along those points to determine the exact age and gender levels for a given face. Each face was assigned one of those five possible gender levels (from very feminine to very masculine), and one of five possible age levels (from very young to very old). The five age values corresponded with the objective ages of 20, 30, 40, 50, and 60 years old. This procedure resulted in 25 possible faces (5 levels of gender $\times$ 5 levels of age = 25). Figure 1 presents the faces created according to this method. The figure shows the structure of the face stimuli on a Cartesian grid. Let i denote the rows of this grid and j denote its columns. The stimulus matrix represents a full factorial combination of i levels of gender and j levels of age. For a given level of facial age j, physical gender becomes more masculine as we move up the ith rows. For a given gender level i, physical age becomes older as we move to the right along the jth columns.

It should be noted that the random generator tool in FaceGen uses a face space of features that is based on a large sample of real-world faces. Any phenotypic correlations between cues for age and gender may represent the structure of cues in real-world faces; and is therefore ecologically valid. For example, a quick look at Figure 1 reveals that skin tone gets lighter as the face becomes feminine. Rahrovan et al. (2018) have noted that: "several spectrophotometric studies have shown that in diverse populations in Europe, Asia, Africa, and North and South America, female skin reflectance is 2 to 3 percentage points above that of male skin (having a higher reflectance means having paler skin.)" (p.127).

Procedure and design

Two groups of observers participated. The first group judged facial gender (Experiment 1), and the second group judged facial age (Experiment 2). Faces were presented at the center of the screen as color images over black background. Viewed from a distance of 56 cm, each face subtended 4.1° horizontally and 9.6° vertically. On each trial, observers viewed 1 of the 325 possible pairs of faces from those illustrated in Figure 1 (including self-comparisons[2]). The two faces appeared in succession. The observer's task was to judge which of the two faces looked more masculine (Experiment 1) or older (Experiment 2). Each pair was presented three times, making a total of 975 trials.

The sequence of events was similar to that administrated in previous MLCM studies (Ho et al., 2008). Each trial commenced with a central fixation point presented for 200 ms, then the first face was presented for 400 ms and followed by a 200-ms interstimulus interval. Then the second face was presented for 400 ms. The observer judged whether the first or second faces appeared to him or her more masculine or older, by pressing one of two keys on the keyboard ('z' and 'm'). The observer's response initiated the next trial.

Model

The psychophysical task required for application of MLCM is rather simple (Knoblauch & Maloney, 2012; Luce & Tukey, 1964). Observes viewed all possible pairs of the 25 faces in Figure 1, and on each trial judged which member of the pair was more masculine. In a second experiment, a different group of observers viewed the same stimuli and judged which member of a pair was older. The stimuli in the two experiments were identical, but observers in each experiment evaluated the perceptual effects of one dimension while ignoring the other dimension.

Now, let ϕ^g denote values on the physical gender scale and let ϕ^a denote values on the physical age scale. The value of gender is constant across the ith row in Figure 1, and is denoted ϕ_i^g and the value of age is similarly constant in the jth column and is denoted ϕ_j^a. Comparison between the two faces can be represented as a comparison between two ordered pairs (ϕ_i^g, ϕ_j^a) and (ϕ_k^g, ϕ_l^a). MLCM assumes that the visual system computes an estimate of perceived gender that is based on both its physical gender ϕ^g and physical age ϕ^a. The estimate of this quantity is $\psi^G(\phi_k^g, \phi_l^a)$ with uppercase letters for perceptual measures and lowercase letters for physical variables. The resulting estimates of perceived gender amount to the decision variable:

$$\Delta(i, j, k, l) = \psi^G(\phi_k^g, \phi_l^a) - \psi^G(\phi_i^g, \phi_j^a) + \varepsilon \qquad (1)$$

where the random variable $\varepsilon \sim \mathcal{N}(\mu, \sigma_G^2)$ is a judgment error. The model for perceived age was constructed analogously, but based on the perceived age $\psi^A(\phi_k^g, \phi_l^a)$ and $\psi^A(\phi_i^g, \phi_j^a)$ of the two faces compared and a different magnitude of judgment error for age comparisons σ_A^2.

Maximum likelihood conjoint measurement (Ho et al., 2008; Knoblauch & Maloney, 2012) considers three nested models of the decision process: (a) the independent observer, (b) the additive observer, and (c) the saturated observer. These are applied separately for each observer's data, to obtain the best prediction. Each model provides estimates of perceptual scale values or internal response. These have the property that equal differences in response are perceptually equal. The independence model assumes that decisions are made based only on one dimension, with no

Downloaded from abstracts.iovs.org on 04/25/2024

contribution from the other dimension. The additive model assumes that decisions are made based on the sum of component psychological responses elicited by the physical dimensions. The saturated model assumes that decisions also include an interaction term that depends on the specific values of the two components in addition to their simple additive combination. The three models form a nested series with the independent model as the most constrained and the saturated model as the least constrained. They are evaluated using a nested likelihood ratio test (Fitousi, 2014; Gerardin et al., 2018; Ho et al., 2008; Rogers et al., 2016)

The additive conjoint model replaces $\psi^G(\phi_i^g, \phi_j^a)$ by $\psi_i^{G:g} + \psi_i^{G:a}$, where $\psi_i^{G:g}$ is an additive contribution of physical gender to perceived gender that is constant across the ith row of Figure 1 and $\psi_j^{G:a}$ is a comparable contribution of the physical age to perceived gender that is constant across the jth column. $\psi_i^{G:g}$ and $\psi_j^{G:a}$ are parameters of the additive model that are estimated from data. The additive model is based on the assumption that physical age and physical gender interact in determining perceived gender but that the contribution of a particular level of age $\psi_j^{G:a}$ to perceived gender is independent of the physical gender of the face.

Note that the first term $\psi_i^{G:g}$ forms the psychophysical scale mapping physical gender to perceived gender, and the second term $\psi_j^{G:a}$ represents the contribution or contamination of perceived gender owing to changes in physical age. The additive conjoint model is based on the assumption that the contamination is additive, and one of our goals is to test this assumption. Another goal is to test whether changes in physical age affect perceived gender at all, which amounts to the hypothesis that $\psi_j^{G:a} = 0$ for all j.

The additive model can therefore be rewritten as:

$$\Delta(i, j, k, l) = \psi_k^{G:g} + \psi_l^{G:a} - \psi_i^{G:g} - \psi_j^{G:a} + \varepsilon \qquad (2)$$

which can be rearranged as:

$$\Delta(i, j, k, l) = [\psi_k^{G:g} - \psi_i^{G:g}] + [\psi_l^{G:a} - \psi_j^{G:a}] + \varepsilon \qquad (3)$$

The additive model for gender is based on judgments of a comparison of the perceived gender levels of the two faces with an additive contamination from the difference in perceived age. Because there are 5 levels along each dimension, there are 2×5 levels plus 1 variance, which amount to 11 free parameters. To make the model identifiable, it is customary to fix the response at level 1 (the lowest) in each dimension to 0, $\psi_1^G = \psi_1^A = 0$, and the variance to 1. This decreases the number of free parameters to 8.

The independent model of gender is identical to the additive model except that there is no contamination of perceived gender by age. The decision in this model is then based on:

$$\Delta(i, j, k, l) = [\psi_k^{G:g} - \psi_i^{G:g}] + \varepsilon \qquad (4)$$

In this model, the perceived difference in gender depends only on the physical gender of the faces compared. Hence, the values of ψ_j^A are fixed at 0 and the total number of free parameters is 4.

Last, the saturated model includes an interaction factor that depends on the intensity levels of both gender and age. The decision variable is defined according to:

$$\Delta(i, j, k, l) = [\psi_i^{G:g} + \psi_j^{G:a} + \psi_{ij}^{G:g:a}] - [\psi_k^{G:g} + \psi_l^{G:a} + \psi_{kl}^{G:g:a}] + \varepsilon \qquad (5)$$

In this model, responses cannot be accounted for based on simple additive combination, but require the assumption of interactive terms. Hence, it is assumed that the response to each face in the grid of Figure 1 is based on a unique combination of the separate contributions of age and gender. Recall that the 25 faces in Figure 1 are composed of 5 levels of gender and 5 levels of age. One cell in this grid is fixed at 0 leading to 24 free parameters. This maximal number of free parameters gives this model its name (saturated).

There are analogous equations for additive, independent, and saturated models for judgments of facial age. The three models yield a nested set, with the independent model serving as the most constrained, the saturated model as the least constrained, and the additive model as intermediate. The goals of the experiments are to estimate the perceptual scale values and model the contributions of both dimensions from each observer's data, as well identifying the best fitting model.

Results

Data analysis was performed with the MLCM package (Knoblauch et al., 2014) in the open source software R (R Core Team, 2017). The default method uses the glm function to estimate the model coefficients, and it uses a maximum likelihood criterion. An iteratively re-weighted least squares algorithm is used to find the ML estimates. The χ^2 statistic is computed from the differences of deviance of the nested model fits (see Knoblauch & Maloney, 2012 p. 240–245). The difference in the number of parameters between a pair of models served as the test's degrees of freedom (Ho et al., 2008; Knoblauch & Maloney, 2012). I first compared the additive and independent models for judgments based on facial gender (Experiment 1) and facial age (Experiment 2). I did it separately for each observer. The degrees of freedom for this specific test were computed as the difference between the coefficient

Downloaded from abstracts.iovs.org on 04/25/2024

	Test faces (additive model × independent model)		
	DF	Deviance	*p*-value
Obs. 1	4	5.53	0.23
Obs. 2	4	3.70	0.44
Obs. 3	4	33.41	**<0.001**
Obs. 4	4	5.75	0.21
Obs. 5	4	47.95	**<0.001**
Obs. 6	4	14.89	**<0.005**
Obs. 7	4	2.40	0.66
Obs. 8	4	9.92	**<0.05**

Table 1. Experiment 1: Comparison of independent and additive observer models for judgments of gender. DF = degrees of freedom.

	Test faces (additive model × independent model)		
	DF	Deviance	*p*-value
Obs. 1	4	492.3	**<0.001**
Obs. 2	4	280.2	**<0.001**
Obs. 3	4	492.48	**<0.001**
Obs. 4	4	442.04	**<0.001**
Obs. 5	4	348.52	**<0.001**
Obs. 6	4	635.81	**<0.001**
Obs. 7	4	387.37	**<0.001**
Obs. 8	4	588.27	**<0.001**

Table 2. Experiment 2: Comparison of independent and additive observer models for judgments of age. DF = degrees of freedom.

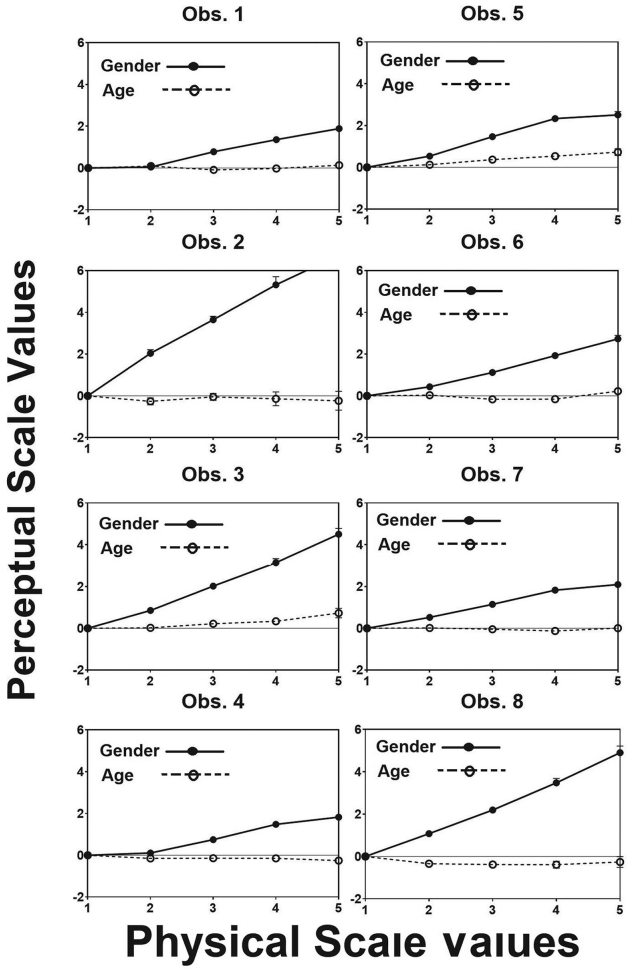

Figure 2. Experiment 1. Estimated scales for judgments based on facial gender. Additive model average estimates for the eight observers. Gender was the relevant dimension. Error bars are standard errors of mean.

estimates in each model (8 for the additive model −4 for the independent model = 4). The results of these comparisons appear on Tables 1 and 2. As can be noted, the additive model was superior to the independent model for four of eight observers in Experiment 1, and for eight of eight observers in Experiment 2. In total the independent model was rejected for 12 of 16 observers. At a Bonferroni correction level of $p < 0.00625$ these results still held for 11 out of 16 participants. These outcomes entail that for most observers judgments of gender or age were contaminated by additive contributions from the task-irrelevant dimension.

I then derived the estimated contributions of gender and age to perceived gender (Experiment 1) and to perceived age (Experiment 2). These estimates can inform us on how observers transformed the physical values of a given dimension (gender or age) onto a perceptual (psychological) scale of gender. Figure 2 plots the perceptual scale values as a function of the physical scale values in the Gender task, separately for gender and age. For all observers the perceptual scale values for gender increased monotonically as the physical scale values of gender increased. Observers judged the face as more masculine as the faces indeed became more masculine. The additive contamination of gender by age was small but significant for four observers (obs. 3, 5, 6, and 8). For two observers (3 and 5), age contributed positively to perceived gender, such that older faces were also perceived as more masculine. For two observers (6 and 8) age negatively contributed to perceived gender, with faces being judged as less masculine as they got older.

Figure 3 plots the perceptual scale values as a function of the physical scale values in the Age task, separately for gender and age. As expected, perceived age increased as physical age increased. In addition, gender contaminated age additively. For five of the eight observers (1, 2, 4, 5, and 7) this contribution was positive, with faces being judged as older as they became more masculine. For three observers (3, 6, and 8) the pattern was more complex, with negative

Downloaded from abstracts.iovs.org on 04/25/2024

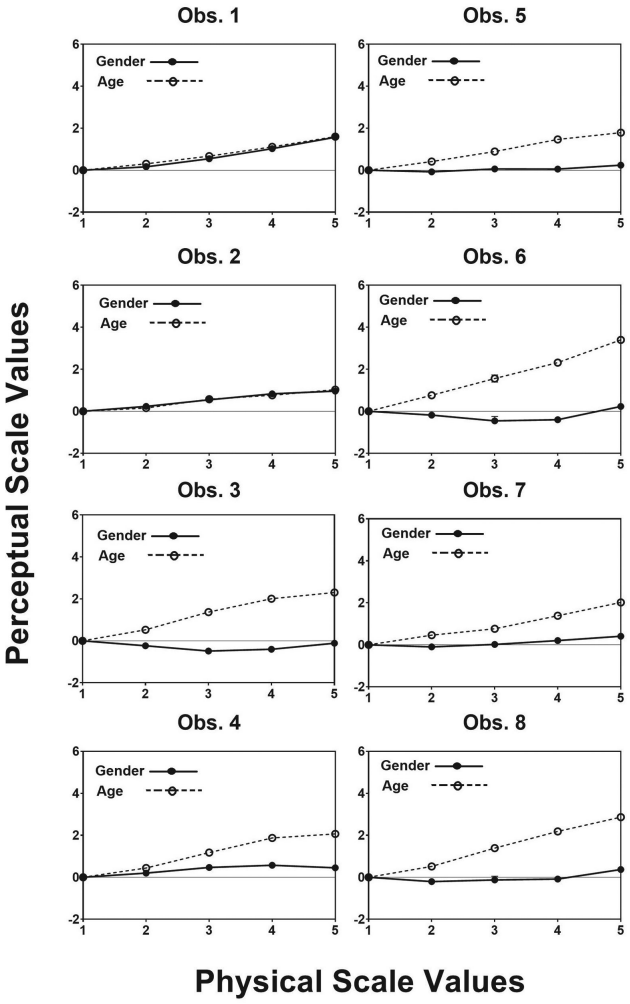

Figure 3. Experiment 2. Estimated scales for judgments based on facial age. Additive model average estimates for the eight observers. Age was the relevant dimension. Error bars are standard errors of mean.

(to level 4) and then (at level 5 = masculine) positive contribution of gender to age.

The independent observer model was rejected for all observers in the age task and for three of eight observers in the gender task. To better estimate the contributions of weighted combinations of age and gender I compared the saturated and additive models for judgments of gender (Experiment 1) and judgments of gender (Experiment 2). The χ^2 statistics in this case has 16 degrees of freedom, computed as the difference in number of free parameters of the less constrained model (the saturated model = 24) and the more constrained model (the additive model = 8). Tables 3 and 4 give the results of these comparisons. The additive model was rejected in favor of the saturated model for six out of eight observers in Experiment 1, and for six out of eight observers in Experiment 2 (*p*s'< Bonferroni-corrected level of 0.00625). Thus, the data

	Test faces (saturated model × additive model)		
	DF	Deviance	*p*-value
Obs. 1	16	14.97	0.52
Obs. 2	16	20.06	0.21
Obs. 3	16	49.93	**<0.001**
Obs. 4	16	58.86	**<0.001**
Obs. 5	16	34.05	**<0.006**
Obs. 6	16	62.64	**<0.001**
Obs. 7	16	64.95	**<0.001**
Obs. 8	16	55.38	**<0.001**

Table 3. Experiment 1: Comparison of additive and saturated models for judgments of gender. DF = degrees of freedom.

	Test faces (saturated model × additive model)		
	DF	Deviance	*p*-value
Obs. 1	16	52.16	**<0.001**
Obs. 2	16	61.43	**<0.001**
Obs. 3	16	46.23	**<0.001**
Obs. 4	16	133.49	**<0.001**
Obs. 5	16	51.36	**<0.001**
Obs. 6	16	39.82	**<0.001**
Obs. 7	16	14.72	0.54
Obs. 8	16	31.84	**<0.05**

Table 4. Comparison of additive and saturated models for judgments of age in Experiment 2. DF = degrees of freedom.

of 12 of 16 observers favored the saturated model over the additive model.

Figures 4 presents the saturated model in the Gender task. As can be noted, for two observers (obs. 1 and 2) the lines are parallel to the *x* axis, an outcome that is in agreement with an independent observer model. The remaining six observers exhibited patterns of interaction which were consistent with the saturated observer model. The source of the interaction for these observers can be visually located at the increase in perceived gender as a function of age for female faces (levels 1 and 2 of the Gender dimension) compared with gender-neutral faces (level 3). These faces were perceived as more masculine as they became older. Four observers (4, 6, 7, and 8) also exhibited an additional trend by which male faces (level 5) were perceived as less masculine as they became older. This pattern suggests that the two extreme levels of gender: femininity and masculinity maintain their strongest degree of separation at young age (level 1 = age 20), which is then gradually decreased.

Figure 5 presents the actual pattern of Gender judgments made by the eight observers in Experiment

Downloaded from abstracts.iovs.org on 04/25/2024

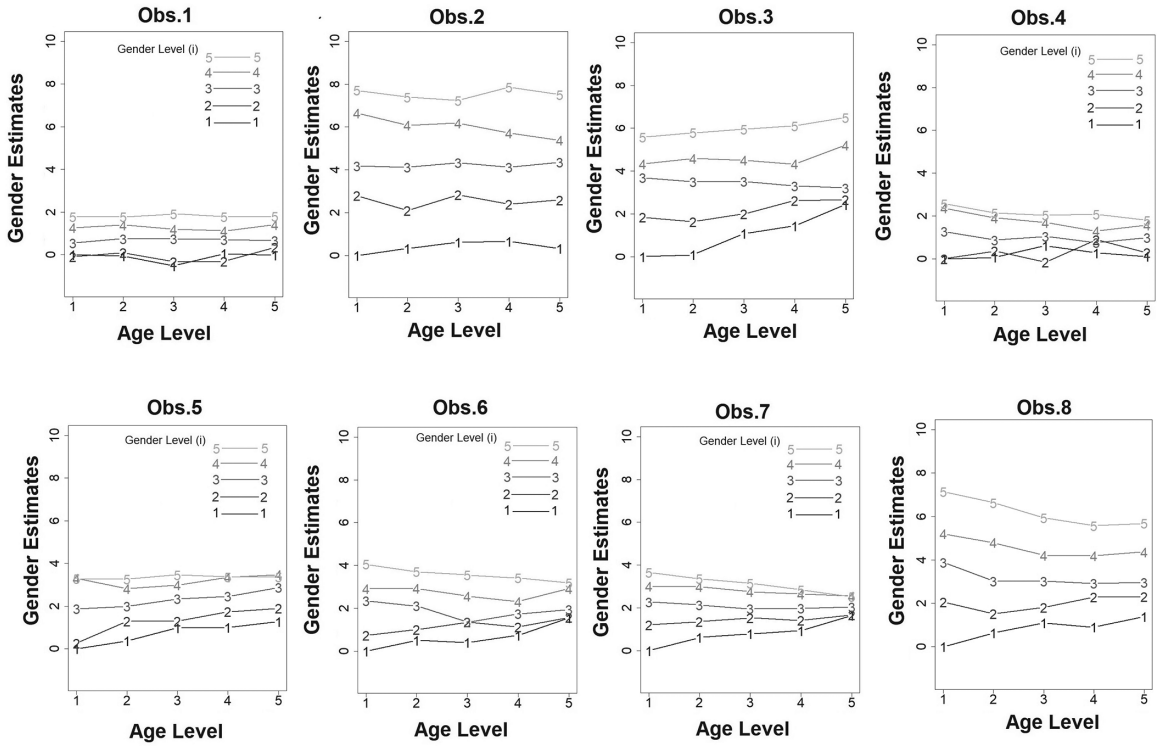

Figure 4. Experiment 1. Results of the saturated model with the estimated contributions for each combination of facial gender and age for each observer. In this experiment facial gender was the relevant dimension of judgment. Levels of facial gender are coded according to: (a) numbers 1-5, with 1 being the most feminine and 5 most masculine level, and (b) lines' color, which gradually shift from black (feminine) to gray (masculine).

1, along with the predicted judgments of Gender uncontaminated by changes in physical Age (i.e., judgments of the ideal observer). These patterns also strengthen the conclusion that observers 3 through 8 deviated from the ideal (independent) model.

Figure 6 presents the outcome of the saturated model in the Age task. As can be noted, for most observers the lines are not strictly parallel to the *x axis*, attesting to the unique contribution of gender to perceived age. The interaction is generated mostly by the very young faces (level 1 = age 20 and level 2 = 30). For these faces perceived age is increasing as the face becomes more masculine. In other words, the more feminine the face the younger it is perceived. Thus, young female faces are perceived as younger than their gender-neutral and male peers. This effect is not present at the old age levels. For most observers, relatively old faces (ages 50 and 60) were perceived as sharing similar psychological age, irrespective of the face's gender level. These results are consistent with the idea that cues for age are confounded with cues for gender mostly in young faces (20s and 30s).

Figure 7 portrays the age judgments made by the eight observers in Experiment 2 along with predicted judgments of Age uncontaminated by changes in physical Gender (i.e., judgments of the ideal observer). The patterns for all observers documented a clear deviation from an independent (i.e., ideal observer) model, and are therefore consistent with the statistical outcomes.

Taken together, these results support the idea that gender and age are for most observers perceived as dependent dimensions. Experiment 1 showed that for six of eight observers female faces were perceived as more masculine as the faces got older. Experiment 2 revealed that for eight of eight observers young faces were perceived as older as they became more masculine. These results capture a strong confound between cues for gender and cues for age. In particular, femininity is associated with young age. This confound has been documented in previous studies (Fitousi, 2020; Kloth et al., 2015; O'Neil & Webster, 2011; Schweinberger et al., 2010; Wiese et al., 2013).

General discussion

I find that facial gender and age are not perceived independently of each other. For 14 of 16 observers, judgments of gender (or age) were contaminated by age (or gender) according to a saturated observer model.

Downloaded from abstracts.iovs.org on 04/25/2024

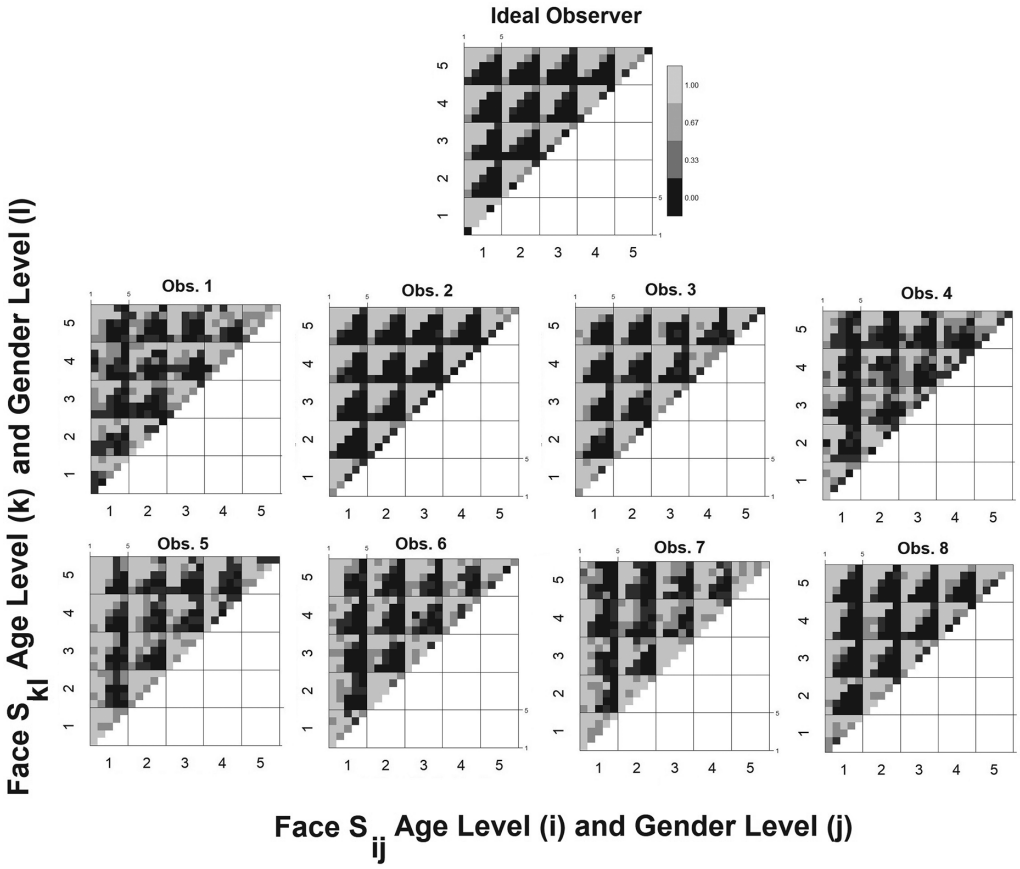

Figure 5. Experiment 1. Gender judgments made by the eight observers are shown along with predicted judgments of Gender uncontaminated by changes in physical age (i.e., judgments of the ideal observer). The gray levels of the squares in the matrices represent the proportion of time that a face S_{kl} was perceived to be more masculine than another face S_{ij} for each pair-wise comparison. Age level *i* is indicated by the large numerical labels (1,2,...,5), and gender level *j* (or *l*) is indicated by the small numerical labels (1,2,...,5).

Generally, the results suggest that female faces are perceived as more masculine as they become older; and young faces (age 20 and 30) are judged as older as they become more masculine. Why do aging and gender interact? The answer is rooted in the perceptual structure of the faces themselves. Perception of facial gender and age rely on shape and texture cues (Brown & Perrett, 1993; Bruce & Langton, 1994; Burton et al., 1993). The correlations between these phenotypic aspects can be readily demonstrated in our set of synthetic face stimuli (Figure 1), and they are likely present in real faces.[3] Facial aging is conveyed by morphological cues (Berry & McArthur, 1986; Burt & Perrett, 1995; George & Hole, 2000; O'Toole et al., 1997) such as a) an increase in the size of the jaw, b) thinning of the lips, and c) an increase in the distance between the eyebrow and the eyes. Textural cues for aging particularly affect the skin: a) skin tone becomes darker, b) it has more wrinkles, c) its luminance contrast decreases, and d) its pigmentation becomes yellower. Many of these shape and texture cues also signal masculine features (Brown & Perrett, 1993; Bruce & Langton, 1994; Burton et al., 1993; Russell, 2009). Men have bigger jaws, their lips are thinner, and their eyebrows are closer to their eyes than females; moreover, they tend to have darker skin with lower contrast (Russell, 2010; Tarr et al., 2001). The upshot is that many of the features that serve as cues for old age are also signs of masculinity. The current study shows that cues for age and for gender have the strongest interactive influence when faces are either young or feminine.

A comment is in order regarding the relations between skin lightness and gender in the current experimental setting. Despite my great efforts to eliminate the correlation between skin lightness and gender, feminine skin tone created by FaceGen were slightly lighter than masculine faces. One may argue that this undermines the current conclusions because observers could have used skin lightness as a cue for gender. However, one should note that such a confound may reflect an ecologically valid cue because a) in real

Downloaded from abstracts.iovs.org on 04/25/2024

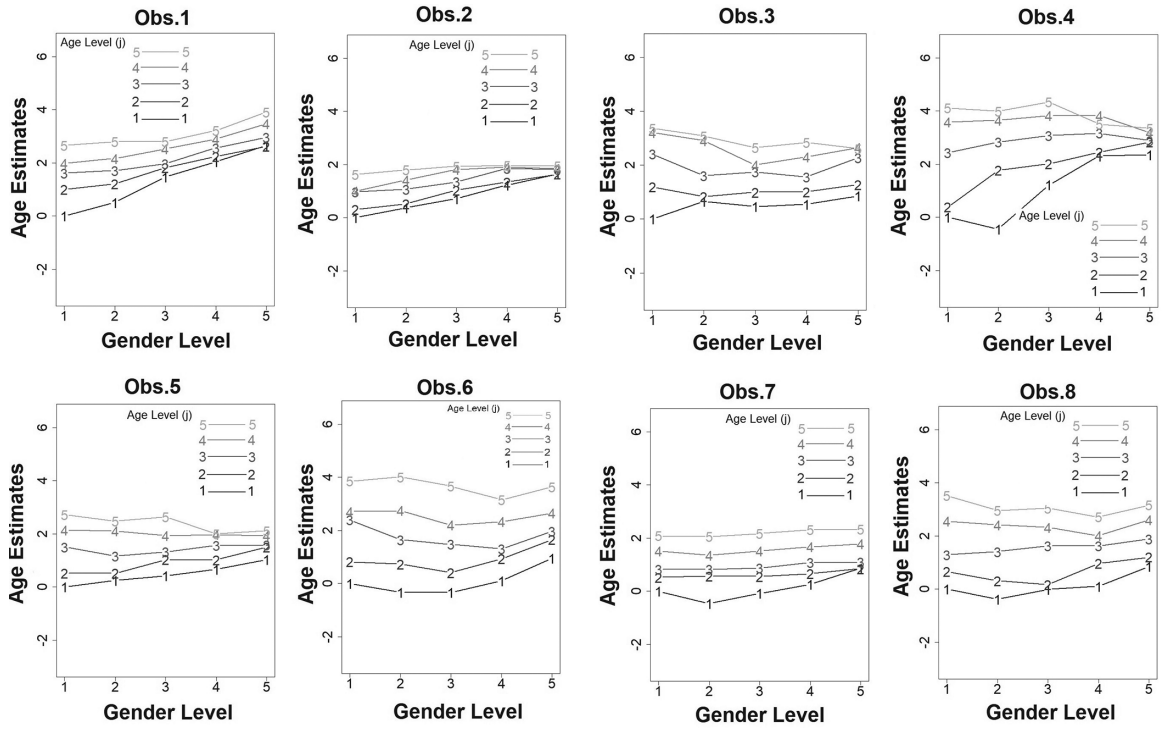

Figure 6. Experiment 2. Results of the saturated model with the estimated contributions for each combination of facial gender and age for each observer. In this experiment facial age is the relevant dimension for judgment. Levels of facial age are coded according to: (a) numbers, with 1 being the younger and 5 older level and (b) lines' color, gradually shifting from black (young) to gray (old).

population female skin reflectance is 2 to 3 percentage points above that of male skin (Rahrovan et al., 2018), and b) FaceGen relies on a representative sample of real people (Inversions, 2008). Moreover, a study by Macke and Wichmann (2010) also attempted to remove textural cues for gender (including lightness), but it seems that these authors could not prevent this built-in confound. In their caption to their Figure 1 they admit that: "For some men with strong beard growth, like the gentlemen in the rightmost column, this meant that there was a slightly darker region around the mouth – at least from an introspective point of view a reasonable cue to gender" (p. 6). The upshot is that it is difficult to equate experimentally the skin lightness of feminine and masculine faces due to natural differences. Future studies may be able to circumvent this confound, but then an issue may arise as to whether such faces reflect the statistical structure of real-world faces.

Evolution, cosmetics, and facial aging

From an evolutionary stand point, the current findings make sense. Fertility in young females may be signaled by cues for femininity and cues for age. The correlation between the two types of cues lead to informational redundancy that increases the chances that information about fertility is transmitted efficiently and correctly to potential mates. This idea can also explain the success of cosmetics (Russell, 2010) and its higher prevalence among women (Etcoff et al., 2011; Russell, 2009). Sexual attractiveness and anti-aging are two main goals of the cosmetics industry, and the current study can explain why. Signals of femininity are positively correlated with attractiveness (O'Toole et al., 1998), and as we have shown here are also negatively correlated with age. This finding can explain the biological incentive for using cosmetics to highlight sexually dimorphic attributes of femininity, but also to conceal cues for old age. Both serve as signals of fertility and are expressed on the same facial cues. For example, Russell (2009) demonstrated the existence of a sex difference in facial contrast that affects the perception of gender. Females have greater luminance contrast between the eyes, lips and the surrounding skin than men. Russell (2009) showed that cosmetics consistently increase facial contrast and thus are functioning to exaggerate feminine features and consequently their attractiveness. Notably, skin contrast also differs between young and old faces and serves as a cue for age (Berry & McArthur, 1986; Burt & Perrett, 1995; George & Hole, 2000; O'Toole et al., 1997). Lower levels of contrast signal old age. Thus, cosmetics not only exaggerates sexually dimorphic attributes, but also decreases perceived age. Etcoff et al. (2011) found

Downloaded from abstracts.iovs.org on 04/25/2024

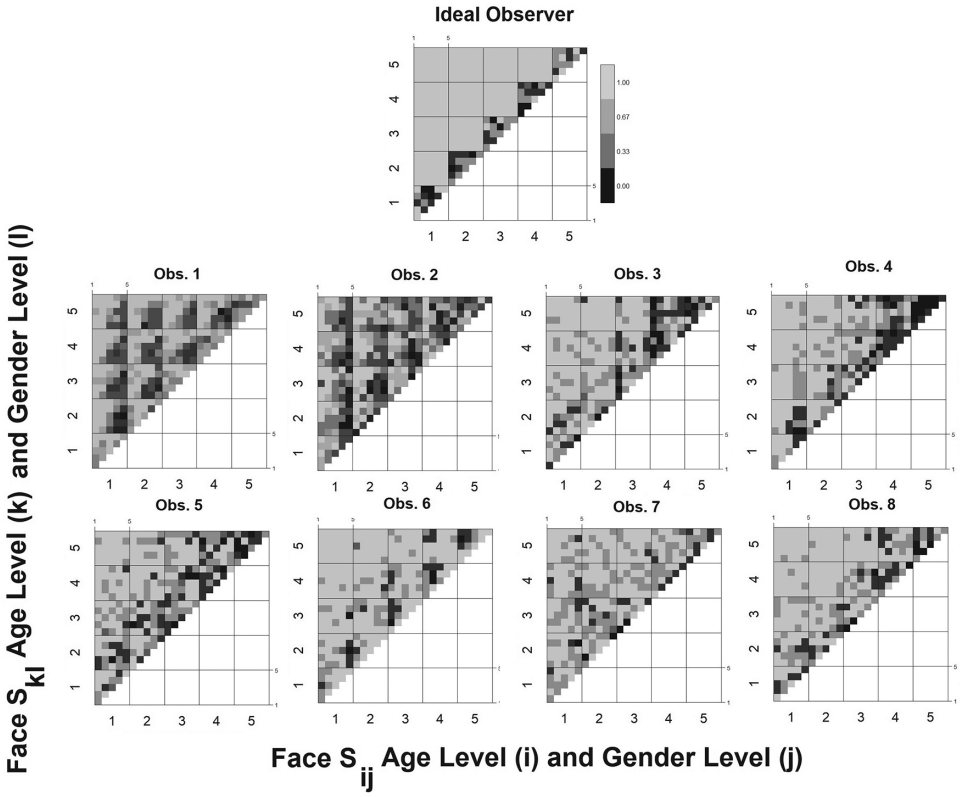

Figure 7. Experiment 2. Age judgments made by the eight observers in Experiment 2 are shown along with predicted judgments of Age uncontaminated by changes in physical Gender (i.e., judgments of the ideal observer). The gray levels of the squares in the matrices represent the proportion of time that a face S_{kl} was perceived to be older than another face S_{ij} for each pair-wise comparison. Age level *i* is indicated by the large numerical labels (1,2,...,5), and gender level *j* (or *l*) is indicated by the small numerical labels (1,2,...,5).

that the influences of cosmetics go even farther than that, exerting dramatic positive effects on judgments of competence, likability, and trustworthiness.

Nonveridical perception of facial gender and age

The present study reveals that facial age and gender are not perceived veridically, but are subjected to major influences of context. Context here refers to the contamination of each dimension by the other. In this sense, each face has a specific gender (age) level that sets a unique context for the perception of its age (gender). This finding is in accordance with the mentioned effects of cosmetics on perceived gender (Etcoff et al., 2011; Russell, 2009), and also with several recent adaptation studies that found that the appearance of both age (O'Neil & Webster, 2011) and gender (Schweinberger et al., 2010) can be altered through adaptation to a previous face. For example, a neutral-gender face seems to be male after adaptation to a female face (Schweinberger et al., 2010). Similarly, adapting to an old face causes faces of intermediate age to appear younger (O'Neil & Webster, 2011). These context effects imply that the internal representations that govern facial age and gender are dynamic and are sensitive to previous experience and correlational structures in the faces themselves. I have recently proposed a 'face file' approach to face recognition (Fitousi, 2017a, 2017b), which assumes that faces are stored as temporary episodic representations with detailed featural information about the face's gender, age, identity, and emotion. These features are bound to each other (e.g., male+young) and can be updated momentarily. Face files can be used to account for the context-dependent nature of facial age and gender (Fitousi, 2017a, 2017b).

Age and gender are essential for what social scientists call person 'construal' (Bodenhausen & Macrae, 1998; Fiske & Neuberg, 1990; Freeman & Ambady, 2011; Macrae & Bodenhausen, 2001), the process by which social agents construct coherent representations of themselves and others. These representations are used by observers to guide information processing and information generations towards others. According to the *dynamic interactive* model by Freeman and Ambady (2011) the initial presentation of a face launches simultaneous activation of several competing social categories (e.g., age, gender, race). Along the accrual of

Downloaded from abstracts.iovs.org on 04/25/2024

evidence, the pattern of activation gradually sharpens into clear interpretation (young female), while other alternatives are inhabited. According to this framework, a confluence of perceptual (bottom-up) and cognitive–social (top-down) factors can generate various types of interactions among social facial dimensions such as facial age and gender. The dynamic–interactive model can account for a large body of research that has documented interactive patterns in face categorization including the current findings (Cloutier et al., 2014; Freeman et al., 2012; Johnson et al., 2012). One crucial goal made explicit by the dynamic-interactive model is the need to distinguish between lower (perceptual) and higher (stereotypes, attitudes, expectations) sources of bias in face categorization (Becker et al., 2007). The former are yielded by correlated phenotypic traits in the sensory cues themselves (skin texture and cues for age), whereas the latter are generated by learned associations or social expectations that can be located in the 'head' of the observer.

The integral/separable distinction and MLCM

The application of the MLCM approach (Knoblauch et al., 2014) to psychological dimensions raises a caveat concerning a more general issue in psychology—the concept of *perceptual independence*. Garner proposed a fundamental distinction between *integral* and *separable* dimensions (Garner, 1962, 1970, 1974a, 1974b, 1976, 1991). This distinction is a pillar of modern cognitive science (for review see Algom & Fitousi, 2016). Objects made of integral dimensions, such as hue and saturation are perceived in their totality and cannot be readily decomposed into their constituent dimensions. Objects made of separable dimensions, such as shape and color, can be readily decomposed into their constituent dimensions. The integral–separable distinction cannot be decided based on the verdict of only one procedure. There is the risk that a theoretical concept (e.g., separability) would be only a restatement of the empirical result (Fitousi, 2015; Von Der Heide et al., 2018).

To avoid circular reasoning, Garner has noted the need for converging operations (Garner et al., 1956). Several methodologies have been used to support the integral–separable distinction: a) Garner's speeded classification task (Garner, 1974b), b) similarity scaling (Attneave, 1950; Melara, 1992), c) information theory (Fitousi, 2013; Garner, 1962; Garner & Morton, 1969), d) general recognition theory (GRT Ashby & Townsend, 1986; Fitousi, 2013; Townsend et al., 2012; Maddox & Ashby, 1996), and e) system factorial technology (SFT Townsend & Nozawa, 1995). Take method b) for example, in which observers are asked to rate the similarity of two objects (Hyman & Well, 1967). It has been often found that for integral objects similarity is computed according to a Euclidian distance metric, and for separable objects similarity is computed according to a city-block distance metric (Melara, 1992). It has also been shown that the outcome from the similarity procedure accords well with the Garner task results (Algom & Fitousi, 2016).

Recently, Rogers et al. (2016) and Rogers et al. (2018) have proposed that the MLCM can be used as a converging operation on the notion of integrality–separability. A case in point in their studies is the color dimensions of chroma and lightness (Munsell, 1912). In the Garnerian tradition, these dimensions are considered as classic integral dimensions: a) they produce Garner interference (Garner & Felfoldy, 1970) and b) they obey a Euclidian distance metric in similarity scaling (Burns & Shepp, 1988). If indeed the dimensions are dependent in processing, then an additive or saturated observer MLCM models should best describe the data. Rogers et al. (2016) found that the additive observer model best described the data. Lightness negatively contributed to perception of chroma for red, blue, and green hues (but not for yellow). These results are important because they demonstrate the utility of the MLCM in providing converging evidence on the notion of integrality–separability, and in identifying the internal representations that govern color dimensions. They are also highly informative in uncovering the specific pattern of dimensional interaction. One would have expected integral dimensions to be best fitted by saturated observer model rather than additive observer model. Hence, the application of multiple related methodologies to investigate questions of perceptual independence is of great practical and theoretical importance in sharpening and explicating our concepts.

The Garnerian edifice is rich in theoretical insights that can illuminate issues in MLCM, and vice versa. This can lead to a cross-fertilization of both methods. For example, an important caveat raised in the Garnerian tradition concerns the direction of interaction between a pair of dimensions. Integrality is not a symmetric concept. Dimension A can be integral with dimension B, while dimension B can be separable from dimension A (Fitousi & Algom, 2020). This notion can be readily applied to studies in MLCM. When judging dimension A and ignoring dimension B, observers can exhibit a complete independent observer model. However, when judging dimension B and ignoring dimension A, observers can exhibit an additive or saturated observer model. Moreover, Garnerian theorizing highlights the role of relative discriminability in determining the direction of asymmetry (Melara & Mounts, 1993). Often the more discriminable dimension intrudes on the less-discriminable dimension (Fitousi & Algom, 2006). It has been shown that relative discriminability can be altered by the researcher and determine the direction of interaction. Therefore, to provide a fair test of independence the dimensions should be equally discriminable (Algom et al., 1996). These factors might also be important in MLCM modeling.

Downloaded from abstracts.iovs.org on 04/25/2024

Future work should test in detail the exact relations between notions of integrality–separability in the Garner tradition and the notions of MLCM. It is not immediately clear for example, that independence in the two approaches is the same. When the dimensions of facial age and gender were subjected to the Garner test Garner (1974b) by Fitousi (2020), they were found to be separable dimensions. But the application of the MLCM to the same dimensions supported their dependency. Why age and gender can appear as separable dimensions in the Garner paradigm and as integral or interactive dimensions (Algom et al., 2017; Algom & Fitousi, 2016) in the MLCM? The solution to this caveat comes by assuming that perceptual independence is not a unitary concept, but rather a nomenclature pointing to various types of independence (Ashby & Townsend, 1986; Fitousi, 2013, 2015; Fitousi & Wenger, 2013). This idea has been originally developed by Garner and Morton (1969) and Ashby and Townsend (1986). It seems that conjoint measurement gauges different types of independence than the Garner paradigm. Future studies may be able to understand the relations between these two approaches.

Keywords: facial gender, facial age, social categorization, conjoint measurement

Acknowledgments

Supported by the ISRAEL SCIENCE FOUNDATION (grant No. 1498/21).

Data can be downloaded from https://data.mendeley.com/datasets/29z57w5989/2.

Commercial relationships: none.
Corresponding author: Daniel Fitousi.
Email: danielfi@ariel.ac.il.
Address: Ariel University, 65, Ramat HaGolan, Ariel 4077625, Israel.

Footnotes

[1]An alternative hypothesis might be that these biases reflect stereotypes or expectations that relate femininity to young age and masculinity to old age.
[2]Self-comparisons are often added to test for response bias. They do not affect the outcome of the modeling. They were not used here.
[3]Note that the software that generated the faces in the current study, FaceGen, relies on real-world statistics, and the resulting faces reflect ecological regularities with correlated texture and shape cues.

References

Abbatecola, C., Gerardin, P., Beneyton, K., Kennedy, H., & Knoblauch, K. (2021). The role of unimodal feedback pathways in gender perception during activation of voice and face areas. *Frontiers in Systems Neuroscience, 15,* 46.

Algom, D., Dekel, A., & Pansky, A. (1996). The perception of number from the separability of the stimulus: The stroop effect revisited. *Memory & Cognition, 24*(5), 557–572.

Algom, D., & Fitousi, D. (2016). Half a century of research on garner interference and the separability–integrality distinction. *Psychological Bulletin, 142*(12), 1352–1383.

Algom, D., Fitousi, D., & Eidels, A. (2017). Bridge-building: Sft interrogation of major cognitive phenomena. In *Systems Factorial Technology* (pp. 115–136). New York: Elsevier.

Allport, G. W., Clark, K., & Pettigrew, T. (1954). The nature of prejudice. Addison-wesley Reading, MA.

Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review, 93*(2), 154–179.

Attneave, F. (1950). Dimensions of similarity. *American Journal of Psychology, 63*(4), 516–556.

Barrett, S. E., & O'Toole, A. J. (2009). Face adaptation to gender: Does adaptation transfer across age categories? *Visual Cognition, 17*(5), 700–715.

Beale, J. M., & Keil, F. C. (1995). Categorical effects in the perception of faces. *Cognition, 57*(3), 217–239.

Becker, D. V., Kenrick, D. T., Neuberg, S. L., Blackwell, K., & Smith, D. M. (2007). The confounded nature of angry men and happy women. *Journal of Personality and Social Psychology, 92*(2), 179.

Berry, D. S., & McArthur, L. Z. (1986). Perceiving character in faces: the impact of age-related craniofacial changes on social perception. *Psychological Bulletin, 100*(1), 3.

Blanz, V., & Vetter, T. (1999). A morphable model for the synthesis of 3d faces. In *Proceedings of the 26th annual conference on computer graphics and interactive techniques* (pp. 187–194), https://doi.org/10.1145/311535.311556.

Bodenhausen, G. V., & Macrae, C. N. (1998). Stereotype activation and inhibition. *Advances in Social Cognition, 11,* 1–52.

Brown, E., & Perrett, D. I. (1993). What gives a face its gender? *Perception, 22*(7), 829–840.

Bruce, V., & Langton, S. (1994). The use of pigmentation and shading information in recognising the sex and identities of faces. *Perception, 23*(7), 803–822.

Burns, B., & Shepp, B. E. (1988). Dimensional interactions and the structure of psychological space: The representation of hue, saturation, and brightness. *Perception & Psychophysics, 43*(5), 494–507.

Downloaded from abstracts.iovs.org on 04/25/2024

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*(1355), 137–143.

Burton, A. M., Bruce, V., & Dench, N. (1993). What's the difference between men and women? Evidence from facial measurement. *Perception, 22*(2), 153–176.

Cloutier, J., Freeman, J. B., & Ambady, N. (2014). Investigating the early stages of person perception: The asymmetry of social categorization by sex vs. age. *PloS One, 9*(1), e84677.

Eidinger, E., Enbar, R., & Hassner, T. (2014). Age and gender estimation of unfiltered faces. *IEEE Transactions on Information Forensics and Security, 9*(12), 2170–2179.

Enlow, D. H., & Moyers, R. E. (1982). *Handbook of facial growth*. WB Saunders Company.

Etcoff, N. L., Stock, S., Haley, L. E., Vickery, S. A., & House, D. M. (2011). Cosmetics as a feature of the extended human phenotype: Modulation of the perception of biologically important facial signals. *PloS One, 6*(10), e25656.

Fiske, S. T., & Neuberg, S. L. (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,* 1–74.

Fitousi, D. (2013). Mutual information, perceptual independence, and holistic face perception. *Attention, Perception, & Psychophysics, 75*(5), 983–1000.

Fitousi, D. (2014). On the internal representation of numerical magnitude and physical size. *Experimental Psychology, 61,* 149–163.

Fitousi, D. (2015). Composite faces are not processed holistically: Evidence from the garner and redundant target paradigms. *Attention, Perception, & Psychophysics, 77*(6), 2037–2060.

Fitousi, D. (2017a). Binding sex, age, and race in unfamiliar faces: The formation of "face files." *Journal of Experimental Social Psychology, 71,* 1–15.

Fitousi, D. (2017b). What's in a "face file"? Feature binding with facial identity, emotion, and gaze direction. *Psychological Research, 81*(4), 777–794.

Fitousi, D. (2020). Evaluating the independence of age, sex, and race in judgment of faces. *Cognition, 202,* 104333.

Fitousi, D., & Algom, D. (2006). Size congruity effects with two-digit numbers: Expanding the number line? *Memory & Cognition, 34*(2), 445–457.

Fitousi, D., & Algom, D. (2020). A model for two-digit number processing based on a joint garner and system factorial technology analysis. *Journal of Experimental Psychology: General, 149*(4), 676–700.

Fitousi, D., & Wenger, M. J. (2013). Variants of independence in the perception of facial identity and expression. *Journal of Experimental Psychology: Human Perception and Performance, 39*(1), 133.

Freeman, J., & Ambady, N. (2011). A dynamic interactive theory of person construal. *Psychological Review, 118*(2), 247.

Freeman, J., Johnson, K., Adams, R., Jr, & Ambady, N. (2012). The social-sensory interface: Category interactions in person perception. *Frontiers in Integrative Neuroscience, 6,* 81.

Garner, W. R. (1962). *Uncertainty and structure as psychological concepts.* New York, NY: Wiley.

Garner, W. R. (1970). The stimulus in information processing. *American Psychologist, 25*, 350–358.

Garner, W. R. (1974a). Attention: The processing of multiple sources of information. In E.C. Carterette, & M. P. Friedman (Eds.), *Handbook of Perception* (Vol. *2*, p. 23–59). New York: Academic Press.

Garner, W. R. (1974b). *The processing of information and structure.* Potomac, MD: Erlbaum.

Garner, W. R. (1976). Interaction of stimulus dimensions in concept and choice processes. *Cognitive Psychology, 8*(1), 98–123.

Garner, W. R. (1991). Afterword. In G. R. Lockhead, & J. R. Pomerantz (Eds.), *The perception of structure: Essays in honor of Wendell R. Garner* (p. 327–332)). Washington, DC: American Psychological Association.

Garner, W. R., & Felfoldy, G. L. (1970). Integrality of stimulus dimensions in various types of information processing. *Cognitive Psychology, 1*(3), 225–241.

Garner, W. R., Hake, H. W., & Eriksen, C. W. (1956). Operationism and the concept of perception. *Psychological Review, 63*(3), 149.

Garner, W. R., & Morton, J. (1969). Perceptual independence: Definitions, models, and experimental paradigms. *Psychological Bulletin, 72*(4), 233.

George, P. A., & Hole, G. J. (2000). The role of spatial and surface cues in the age-processing of unfamiliar faces. *Visual Cognition, 7*(4), 485–509.

Gerardin, P., Dojat, M., Knoblauch, K., & Devinck, F. (2018). Effects of background and contour luminance on the hue and brightness of the watercolor effect. *Vision Research, 144,* 9–19.

Ho, Y.-X., Landy, M. S., & Maloney, L. T. (2008). Conjoint measurement of gloss and surface texture. *Psychological Science, 19*(2), 196–204.

Downloaded from abstracts.iovs.org on 04/25/2024

Hyman, R., & Well, A. (1967). Judgments of similarity and spatial models. *Perception & Psychophysics, 2*(6), 233–248.

Inversions, S. (2008). *Facegen Modeller (Version 3.3)[Computer Software]*. Toronto, ON: Singular Inversions.

Johnson, K. L., Freeman, J., & Pauker, K. (2012). Race is gendered: How covarying phenotypes and stereotypes bias sex categorization. *Journal of Personality and Social Psychology, 102*(1), 116.

Kloth, N., Damm, M., Schweinberger, S. R., & Wiese, H. (2015). Aging affects sex categorization of male and female faces in opposite ways. *Acta Psychologica, 158,* 78–86.

Knoblauch, K., Maloney, L., & Aguilar, G. (2014). Mlcm: Maximum likelihood conjoint measurement. *R package version 0.4, 1*.

Knoblauch, K., & Maloney, L. T. (2012). *Modeling psychophysical data in R* (Vol. *32*). New York: Springer Science & Business Media.

Krantz, D., Luce, D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement* (Vol. I). Mineola, New York: Dover Publications, Inc.

Levi, G., & Hassner, T. (2015). Age and gender classification using convolutional neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (pp. 34–42).

Lisi, M., & Gorea, A. (2016). Time constancy in human perception. *Journal of Vision, 16*(14), 3–3.

Luce, R. D., & Tukey, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology, 1*(1), 1–27.

Macke, J. H., & Wichmann, F. A. (2010). Estimating predictive stimulus features from psychophysical data: The decision image technique applied to human faces. *Journal of Vision, 10*(5), 22–22.

Macrae, C. N., & Bodenhausen, G. V. (2001). Social cognition: Categorical person perception. *British Journal of Psychology, 92*(1), 239–255.

Maddox, W. T., & Ashby, F. G. (1996). Perceptual separability, decisional separability, and the identification–speeded classification relationship. *Journal of Experimental Psychology: Human Perception and Performance, 22*(4), 795.

Maloney, L. T., & Knoblauch, K. (2020). Measuring and modeling visual appearance. *Annual Review of Vision Science, 6,* 519–537.

Melara, R. D. (1992). The concept of perceptual similarity: From psychophysics to cognitive psychology. In D. Algom (Ed.), *Psychophysical approaches to cognition* (pp. 303–388). Amsterdam, the Netherlands: Elsevier.

Melara, R. D., & Mounts, J. R. (1993). Selective attention to stroop dimensions: Effects of baseline discriminability, response mode, and practice. *Memory & Cognition, 21*(5), 627–645.

Munsell, A. H. (1912). A pigment color system and notation. *American Journal of Psychology, 23*(2), 236–244.

Nichiporuk, N., Knoblauch, K., Abbatecola, C., & Shevell, S. (2018). Does observer's ethnicity affect perceived face lightness? A study of the face-lightness distortion effect for African American and Caucasian observers. *Journal of Vision, 18*(10), 1099–1099.

O'Neil, S. F., & Webster, M. A. (2011). Adaptation and the perception of facial age. *Visual Cognition, 19*(4), 534–550.

Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences, 105*(32), 11087–11092.

O'Toole, A. J., Vetter, T., Volz, H., & Salter, E. M. (1997). Three-dimensional caricatures of human heads: Distinctiveness and the perception of facial age. *Perception, 26*(6), 719–732.

O'Toole, A. J., Deffenbacher, K. A., Valentin, D., McKee, K., Huff, D., & Abdi, H. (1998). The perception of face gender: The role of stimulus structure in recognition and classification. *Memory & Cognition, 26*(1), 146–160.

Qi, L., Chantler, M. J., Siebert, J. P., & Dong, J. (2015). The joint effect of mesoscale and microscale roughness on perceived gloss. *Vision Research, 115,* 209–217.

Quinn, K. A., & Macrae, C. N. (2005). Categorizing others: the dynamics of person construal. *Journal of Personality and Social Psychology, 88*(3), 467.

Rahrovan, S., Fanian, F., Mehryan, P., Humbert, P., & Firooz, A. (2018). Male versus female skin: What dermatologists and cosmeticians should know. *International Journal of Women's Dermatology, 4*(3), 122–130.

R Core Team. (2017). *R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing*. Vienna, Austria: URL https://www.R-project.org.

Rogers, M., Franklin, A., & Knoblauch, K. (2018). A novel method to investigate how dimensions interact to inform perceptual salience in infancy. *Infancy, 23*(6), 833–856.

Rogers, M., Knoblauch, K., & Franklin, A. (2016). Maximum likelihood conjoint measurement of lightness and chroma. *JOSA A, 33*(3), A184–A193.

Russell, R. (2009). A sex difference in facial contrast and its exaggeration by cosmetics. *Perception, 38*(8), 1211–1219.

Downloaded from abstracts.iovs.org on 04/25/2024

Russell, R. (2010). Why cosmetics work. In: R. B. Adams, N. Ambady, K. Nakayama, & S. Shimojo (Eds.), *The Science of Social Vision* (pp. 186–204). Oxford University Press.

Schweinberger, S. R., Zäske, R., Walther, C., Golle, J., Kovács, G., & Wiese, H. (2010). Young without plastic surgery: Perceptual adaptation to the age of female and male faces. *Vision Research, 50*(23), 2570–2576.

Tarr, M. J., Kersten, D., Cheng, Y., & Rossion, B. (2001). It's Pat! Sexing faces using only red and green. *Journal of Vision, 1*(3), 337–337.

Townsend, J. T., Houpt, J. W., & Silbert, N. H. (2012). General recognition theory extended to include response times: Predictions for a class of parallel systems. *Journal of Mathematical Psychology, 56*(6), 476–494.

Townsend, J. T., & Nozawa, G. (1995). Spatio-temporal properties of elementary perception: An investigation of parallel, serial, and coactive theories. *Journal of Mathematical Psychology, 39*(4), 321–359.

Von Der Heide, R. J., Wenger, M. J., Bittner, J. L., & Fitousi, D. (2018). Converging operations and the role of perceptual and decisional influences on the perception of faces: Neural and behavioral evidence. *Brain and Cognition, 122,* 59–75.

Wiese, H., Komes, J., & Schweinberger, S. R. (2013). Ageing faces in ageing minds: A review on the own-age bias in face recognition. *Visual Cognition, 21*(9–10), 1337–1363.

Young, A. W., & Burton, A. M. (2017). Recognizing faces. *Current Directions in Psychological Science, 26*(3), 212–217.

Downloaded from abstracts.iovs.org on 04/25/2024