# Age of face matters: Age-group differences in ratings of young and old faces

NATALIE C. EBNER

Max Planck Institute for Human Development, Berlin, Germany

Young (n = 24) and old (n = 24) participants rated 160 faces of young and old individuals taken from the CAL/PAL Face Database (Minear & Park, 2004) with regard to attractiveness, likeability, distinctiveness, goal orientation, energy, mood, and age. Ratings are reported for each face separately. Further analyses showed that the age groups differed in their ratings of young and old faces. On average, old participants evaluated the faces as more positive (i.e., more attractive, more energetic) than did young participants. In line with research on a negative aging stereotype, old faces were judged as less positive than young faces. They were, for instance, seen as less attractive, less likeable, less distinctive, less growth-oriented, and less energetic. The findings of the present study can serve as a basis for the selection of appropriate facial stimuli in age-comparative studies of face perception, face processing, or memory for faces. All face-specific data are archived at www.psychonomic.org/archive.

Human faces constitute a unique category of objects, in that, from very early on, we are frequently exposed to them in our daily lives. The interest in faces as research stimuli is broad: Images of faces are used in various fields of research including perception, memory, emotion, and infant or adult development. Due to a lack of appropriate stimulus material, most studies so far have, however, exclusively used pictures of faces of young adults-even when addressing developmental research questions and comparing different age groups. Evidence suggests, however, that the type (e.g., age, gender, or race) of a face can influence how well the face is processed, recognized, and remembered (Bäckman, 1991; Brigham & Barkowitz, 1978; Golby, Gabrieli, Chiao, & Eberhardt, 2001; Mason, 1986). Faces that are more similar to the face of the individual studying them are remembered and discriminated better than are faces that are different. Findings of an "own-age bias" in face recognition and person identification specifically suggest that adults of different ages are more likely to identify, recognize, and remember faces and persons of their own age in comparison with faces and persons of other ages (Anastasi & Rhodes, 2006; Bäckman, 1991; Bartlett & Fulton, 1991; Lamont, Stewart-Williams, & Podd, 2005; Mason, 1986; Perfect & Harris, 2003; Wright & Stroud, 2002). This effect is generally explained by the amount of exposure to and contact with certain classes of faces an individual has, the idea being that people more frequently encounter own-age faces than other-age faces. They are therefore more familiar with, and acquire more expertise at processing, faces of their own age as compared with faces of other age groups, with positive effects on memory for own-age faces. In line with this assumption, Bartlett and Fulton (1991) found that old faces

are perceived as more familiar by old participants than by young participants, whereas young faces are perceived as more familiar by young participants than by old participants. Consequently, old participants, compared with young participants, are at a disadvantage when processing or recognizing young faces, whereas the opposite is true for young participants with regard to old faces. This ownage bias clearly challenges the findings and interpretation of current age-comparative behavioral and neuroimaging studies that have used facial stimuli but not systematically varied the age of the presented facial stimuli.

A major step to overcome the lack of appropriate agedifferential facial stimuli has been undertaken by Minear and Park (2004), who have recently created a new large set of faces representing a wide range of different age groups: the CAL/PAL Face Database. This set of faces includes over 1,000 color (and black-and-white) photographs of young, middle-aged, and old adults, ranging in age from 19 to 93 years. The database comprises neutral facial expressions for each individual face. For some of the faces, there are also happy expressions, as well as images of profiles. Just recently, the authors have added sad, angry, annoyed, grumpy or disgusted, and surprised facial expressions for some of the faces (see Minear & Park, 2004, for further information). The faces can be obtained for research purposes from agingmind.utdallas.edu/facedb/.

This new material allows researchers to systematically vary the age of the faces used as experimental stimuli when comparing different age groups. At the same time, the existence of this new material makes it possible, and necessary, to collect information about facial features (such as attractiveness or distinctiveness) that may be of relevance when using these faces (and especially faces of different ages) as

N. C. Ebner, natalie.ebner@yale.edu

research stimuli. In addition, this new material allows the examination of differences in ratings of facial attributes of young and old faces and the investigation of age-group differences in these ratings. This information will not only give researchers the possibility to experimentally control for differences in certain facial features when using faces of different ages, but will help to avoid confounds of such attributes with other variables of interest, or will make possible the use of such additional information about the faces as independent variables in order to explicitly examine their effects. In sum, this information will facilitate and help to systematize the selection of facial stimuli appropriate for use in future (age-comparative) studies.

Human faces of different ages differ in several aspects. There are not only differences between faces of young children and grown-ups (Berry & MoArthur, 1986), but also differences between faces of young and old adults (Bruce & Young, 1998). Crudely, two types of age-related changes in features that refer to the "surface" of a face can be distinguished (Bruce & Young, 1998; Burt & Perrett, 1995): changes in the shape, which mainly occur through growth or weight gain or loss, and changes in the characteristics of the surface texture and coloration of skin and hair. More specifically, young and old faces differ in the shape of the skull and forehead (often due to receding hair with advancing age), the shape and broadness of the jaw and chin, the length of the nose and ears, the distribution of adipose tissue, the size of the eyes, and the shape of the lips. They also differ in terms of the elasticity, texture, and quality of the skin, such as with respect to skin blemishes, skin pigmentation, presence of capillary varicosities, size of pores, and, of course—being one of the most powerful conveyers of age information—prevalence of facial wrinkles. Finally, young and old faces typically differ in the color, quantity, and dispersal of hair (e.g., eyebrows, beards, head hair). It has been shown that human vision is very sensitive to these different kinds of features and age-associated changes in these features when determining the age of a face, even when the differences and changes are rather subtle and difficult to describe (Bruce & Young, 1998).

Such "surface features" of a face may strongly influence how attractive, distinctive, or old a face is seen to be. In addition to these "surface features," however, a face can also convey information about the "person behind the face," such as underlying personality characteristics or behaviors and motive structures (Berry & McArthur, 1986; Berscheid & Walster, 1974), and it seems plausible that this information differs also as a function of the age of a face, partly due to stereotypes and generalized expectations. This information from the face possibly influences how likeable or energetic we perceive a person to be, or what types of goals we would ascribe to him or her. In this sense, facial characteristics, in interplay with stereotypes and expectations, influence to some degree the impression we have of a person and his or her psychological qualities and how we judge his or her character and habits.

The perception of facial features ("surface features," as well as information a face conveys about underlying characteristics) is, however, not only influenced by the perceived face itself, but also by the perceiver's charac-

teristics, needs, goals, and abilities (cf. Gibson, 1979; McArthur & Baron, 1983). That is, although certain facial attributes may communicate information about a particular individual's traits or behavior, the meaning of this information will vary as a function of the perceiver's own traits or behavior. It can therefore be assumed that faces of different ages are perceived differently in terms of their "surface features," but also with respect to more underlying person- and behavior-related dimensions, and differently so by young and old perceivers.

One purpose of the present study is to provide facespecific ratings for a selection of young and old faces from the CAL/PAL Face Database. A second purpose is to examine whether, and in what regard, young and old faces are perceived differently, and whether young and old participants differ in their ratings of young and old faces. To obtain this information, equal numbers of young and old participants were asked to judge each of 160 young and old faces in terms of attractiveness, likeability, distinctiveness, goal orientation, energy, mood, and age.

These seven rating dimensions were selected for the following reasons: They all seem to be relevant for processing, recognizing, and remembering young and old faces and therefore need to be taken into consideration when selecting appropriate facial stimuli for age-comparative research. Attractiveness and distinctiveness ratings are typically assessed in the context of facial databases (e.g., Psychological Image Collection at Stirling, n.d.). These two dimensions, as well as the dimension of age, might be strongly influenced by facial features at the surface, as outlined above. Ratings with regard to the age of the face were primarily collected to validate the categorization of a face as being either young or old, as undertaken in the context of the CAL/PAL database as well as in the present study. In contrast to the dimensions attractiveness, distinctiveness, and age, the dimensions likeability, goal orientation, and energy refer more strongly to attributes of the person behind the face—his or her personality characteristics, behaviors, emotions, motives, and goals. It can be assumed that ratings of these dimensions require more elaborated reflection about the person and are more strongly influenced by stereotypes and general expectations. They may involve attribution processes and are possibly less spontaneous than ratings of the attractiveness, distinctiveness, or age of a face. The dimension mood, finally, constitutes an aspect that is somewhat in between facial surface attributes and attributes that refer to the person behind the face. This dimension was also assessed in order to validate the categorization of the selected faces as displaying neutral facial expressions, as undertaken by the authors of the CAL/PAL database as well as in the context of the present study.

# **METHOD**

# **Participants**

Twenty-four young (20-29 years, M = 23.9, SD = 2.6) and 24 old  $(71-85 \text{ years}, M = 76.6, SD = 3.8)^1$  participants took part in this study. The sample was stratified by gender and education. All participants were drawn from the participant pool of the Max Planck Institute for Human Development, Berlin. As usual, young participants (M = 59.5, SD = 11.4) performed better than old participants

 $[M=40.5, SD=10.4; F(1,46)=36.6, p<.01, \eta_p^2=.44]$  in the Digit-Symbol Substitution test (Wechsler, 1981). None of the participants had unusual scores on this test, indicating that all young and old participants were well functioning with regard to visual-motor processing speed.

# Source and Selection of Facial Stimuli

Young and old faces used in the present study were selected from the CAL/PAL Face Database. Selection criteria were that (1) the faces matched the respective age groups (i.e., belonged to persons that were either 18 to 32 years old or 64 years and older), (2) were Caucasian, in order to be representative for the sample that rated the faces, and (3) showed a neutral facial expression. The selected 160 pictures (80 young and 80 old faces, evenly distributed across gender) were then modified by removing excessive jewelry and by reducing eye-catching hairstyles or makeup and by standardizing light effects, visible background color (colored midgray), and color (colored black) and shape (e.g., covering the shoulders) of clothes. Picture editing was done in Adobe Photoshop CS on an Apple Macintosh computer. The gray levels of the final pictures were around 640 × 480 pixels and 72 ppi.

# **Procedure and Face Rating Task**

The study consisted of one testing session. Young and old participants were tested in separate sessions. No more than 5 participants were tested simultaneously. All participants were seated in front of computers separated by partition walls. At the beginning of the session, participants received information about the testing procedure and then signed a consent form. For sample descriptive purposes, participants then responded to a short sociodemographic questionnaire and worked on the Digit-Symbol Substitution test as a paper-and-pencil cognitive-behavioral measure of perceptual-motor speed.

Participants were then asked to rate various faces along several dimensions. Specifically, after a short introduction by the experimenter and additional written instructions, as well as a practice trial provided for by the computer, the 160 faces, 1 at a time, appeared on the screen, and participants rated each face in terms of attractiveness, likeability, distinctiveness, goal orientation, energy, mood, and age. Table 1 presents the specific wordings of the questions and the response options. Participants were instructed to look at the faces carefully and then to give a spontaneous personal response. Only

Table 1
Specific Wordings and Response Options for Rating Dimensions

Wording	Response Options
"How attractive is this person?"	0 (not at all) to 4 (very much)
"How likeable is this person?"	0 (not at all) to 4 (very much)
"How distinctive is this person?"	0 (not at all) to 4 (very much)
"I think this person would like to improve or to prevent loss in functioning."	0 (growth) 1 (loss prevention)
"How energetic is this person?"	0 (not at all) to 4 (very much)
"In what mood is this person?"	0 (sad) 1 (neutral) 2 (happy)
"How old is this person, in years?"	0 (20–30 yrs.) 1 (30–40 yrs.) 2 (40–50 yrs.) 3 (50–60 yrs.) 4 (60–70 yrs.) 5 (70–80 yrs.) 6 (80–90 yrs.) 7 (90–100 yrs.)
	"How attractive is this person?"  "How likeable is this person?"  "How distinctive is this person?"  "I think this person would like to improve or to prevent loss in functioning."  "How energetic is this person?"  "In what mood is this person?"  "How old is this person,

after participants had responded to all of the dimensions for 1 face, did the next face appear on the screen.

The size of all of the pictures was approximately  $17.5 \times 23.5$  cm when displayed on a 19-in. monitor (screen size  $1,280 \times 1,024$  pixels) in the center of the screen. The background of the screen was white. The rating dimensions and the respective response options (both written in 36-point black Arial font) were presented below the image of the face. The corresponding numbers of the response options were placed in little light-gray-colored boxes. The stimulus presentation was controlled using Psy-Point software (Steinkraus & Ebner, 2006) on 1.8-GHz Apple Power Mac G5 computers. The presentation order of faces followed a randomized age- and gendermatched yoked design (i.e., 1 young man, 1 young woman, 1 old man, and 1 old woman saw an identical randomized sequence of faces, respectively) resulting in 12 presentation orders. These 12 randomized presentation orders were controlled for the mean position of each picture in the presentation sequence to avoid order effects of the pictures (i.e., each picture appeared approximately equally often at the beginning, in the middle, and at the end of the sequence across the 12 randomized presentation orders).

The session typically took 150 min, including two 5-min breaks after about every 45 min, to allow participants to maintain concentration throughout the session. At the end of the session, all participants were debriefed and reimbursed €25.

#### RESULTS

This section first presents information pertaining to the ratings for each of the pictures separately. Then, the findings about differences between ratings of young and old faces, as well as differences between young and old participants in these ratings, are reported.

# **Ratings of Young and Old Faces**

Online Appendixes A-G present the ratings for each of the 160 faces separately. Appendix A refers to the rating dimension attractiveness, Appendix B to the dimension likeability, Appendix C to the dimension distinctiveness, Appendix D to the dimension goal orientation, Appendix E to the dimension energy, Appendix F to the dimension mood, and Appendix G to the dimension age. To allow for easy reference, entries are listed according to the identifying labels originally assigned to the faces by Minear and Park (2004), as well as to identifying labels used in the present study. First, information referring to the total sample is presented. followed by information pertaining to the young and old subsamples, respectively. Specifically, Appendixes A, B, C, and E present information about the means and standard deviations of the ratings, whereas Appendixes D. F. and G report information about the percentages of the ratings, separately for goal orientation toward growth and loss prevention, for sad, neutral, and happy mood, or for each of the eight age categories, respectively. All face-specific data are archived at www.psychonomic.org/archive.

#### Differences in Ratings of Young and Old Faces

Having presented the picture-rating information for each of the young and old faces, results pertaining to the differences in the ratings of young and old faces, as well as differences between young and old participants in these ratings, are reported next.

Table 2 summarizes the descriptive information of the ratings of young and old faces in the total sample as well

8 4 8

as in the young and old subsamples. Table 3 presents the correlations between the seven rating dimensions. Ratings of attractiveness, likeability, distinctiveness, and energy were positively correlated: The more attractive a face was rated, the more likeable, distinctive, and energetic it was indicated to be. There were no differences between young and old participants in any of the significant correlations between the rating dimensions, and none of the correlations differed for young and old faces.

Are young and old faces rated differently in terms of attractiveness, likeability, distinctiveness, goal orientation, energy, mood, and age? And do young and old participants differ in these judgments? To investigate these questions, multilevel random coefficient modeling (MRCM; Hox, 2002; Nezlek, 2001; Raudenbush & Bryk, 2001) was used. MRCM accounts for the hierarchical structure of the data (i.e., ratings of faces nested within participants). Separate models were computed for each of the seven rating dimensions. The responses to the specific rating dimension constituted the dependent variable. At the between-person level, age of participant was used as the predictor variable, to examine whether participants' ratings of the faces varied as a function of the age group of the rater. At the withinperson level (i.e., the level of individual faces), age of face was used as the predictor variable, to examine whether participants' ratings of the faces for each of the seven dimensions separately varied as a function of the age group of the faces themselves. In addition, the cross-level interaction between age of participant and age of face was entered into the model, to examine age-related differences in the withinperson association between the age group of the face and the ratings for the respective rating dimension.

The analyses were conducted using hierarchical linear and nonlinear modeling (HLM 6; Raudenbush, Bryk, & Congdon, 2000). The MRCM results for the ratings as a function of the age of the participant and the age of the face are summarized in Table 4. The top of Table 4 presents the fixed part of the models (i.e., average effects for age of participant, average effects for age of face, and cross-level interaction effects for age of participant × age of face). The random effects (i.e., variance components) are presented at the bottom of Table 4. Different types of models (i.e., linear and logistic regression) and different methods of estimation to fit the expected models to the data (i.e., restricted maximum likelihood and restricted penalized quasilikelihood) were carried out to account for differences in the response options of the dependent variable.

As indicated by the fixed effects results, old participants rated the faces as significantly more attractive than did young participants ( $\gamma = 0.46, p < .05$ ). The size of the effect can be interpreted by considering the response options reported in Table 1 and the descriptive information reported in Table 2. Comparing the size of the regression coefficient of 0.46 with the standard deviation of the mean rating for attractiveness across all faces and all participants (SD = 0.57, reported in Table 2) shows that the difference between the two age groups was about three quarters of one standard deviation. In addition, both age groups rated old faces as about one standard deviation less attractive than young faces ( $\gamma = -0.59, p < .05$ ). This tendency was

		1	rescribance intori	ve Inior	<b>=</b> !	ation About Katings	tings of	roung and	1 Old Fa	es in the	Iotal Sa	mple ar	d in the	Young an	nS PIO P	foung and Old Faces in the Total Sample and in the Young and Old Subsamples			
								Goal Orientation	entation						Mo	Mood			
	Attract	Attractiveness	Likeability	bility	Distinct	iveness	Gro	rowth	Loss Pr	Loss Prevention	Energy	χş	Sad	P	Neutral	trai	Happy	) A	
	M	SD	W	SD	M	SD	M(%)	SD(%)	W(%)	SD (%)	M	SD	W(%)	80%	W(%)	SD (%)	W6%	800	×
Total Sample																			
All faces	1.46	0.57	1.82	0.42	1.85	0.55	54	10.2	46	10.2	2.14	0.40	1	8	3,5	13.7	36	11.7	20
Young faces	1.68	0.56	1.89	0.46	1.86	0.58	73	16.0	27	16.0	2.26	4	: =	9 5	3 8	15.7	3 %	13.7	7
Old faces	1.23	0.65	1.75	0.45	1.83	0.56	35	16.3	<b>.</b> 3	16.3	202	4	; ;	10.5		17.0	9 6	1.5	7 6
Young Participants							}		3	2	i		1		5	2.5	07	0.11	ţ
All faces	1.15	0.53	1.75	0.39	1.76	0.48	52	9.6	84	9.6	2.00	0.39	17	× ×	55	15.8	38	14.0	7 3 1
Young faces	1.45	0.52	1.81	0.40	1.72	0.51	22	17.2	78	17.2	2.13	0.42	: 2	2.	3 2	12.0	3 5	15.0	7 6
Old faces	0.85	0.58	1.68	0.42	1.80	0.48	33	15.4	19	15.4	88	0.45	: 2	11.4	48	15.6	; S	13.6	5 5
Old Participants									;		2	) ;	ì		2	2	3	0.51	į
All faces	1.76	0.43	1.89	0.45	1.93	0.61	55	10.8	45	10.8	2.27	0.37	11	7.7	36	11.4	77	0	245
Young faces	1.91	0.50	1.97	0.52	2.01	0.61	74	14.8	79	14.8	2.38	0.43	<u>~</u>	7.7	8 6	13.4	ž &	10:5	1 0
Old faces	1.61	0.48	1.82	0.47	1.86	9.0	37	17.4	63	17.4	2.16	0.39	202	10.1	23	6.11	22	9.6	44

Table 3

Correlations Between Rating Dimensions in the Total Sample

				Goal		
	Attractiveness	Likeability	Distinctiveness	Orientation	Energy	Mood
Attractiveness						
Likeability	.76*					
Distinctiveness	.43*	.47*				
Goal orientation	07	01	03			
Energy	.54*	.49*	.53*	23		
Mood	.12	.13	.34	.11	.23	
Age	.18	.12	.11	09	.26	.02

Note—Bivariate Pearson correlations were calculated using z scores. Higher scores indicate more attractiveness, more likeability, more distinctiveness, more loss-prevention orientation, more energy, happier mood, and higher age, respectively. p < 0.01 (two-tailed).

less pronounced in old compared with young participants (cross-level interaction effect between age of participant and age of face,  $\gamma = 0.30$ , p < .05).

The age groups did not differ in their ratings of likeability of the faces ( $\gamma = 0.15$ , n.s.). Young and old faces were, however, judged differently across all participants ( $\gamma = -0.14$ , p < .05): Old faces were judged as about a quarter of a standard deviation less likeable than young faces, equally so by young and old participants ( $\gamma = -0.01$ , p < .05).

With respect to distinctiveness ratings, neither of the age groups differed in their overall rating of faces ( $\gamma = 0.29$ , n.s.), nor were young and old faces rated differently across all participants ( $\gamma = 0.08$ , n.s.). The cross-level interaction effect between age of participant and age of face did, however, reach significance ( $\gamma = -0.23$ , p < .05): Old, but not young, participants rated old faces as less distinctive than young faces. This difference amounted to about half a standard deviation.

Because the response option for the dimension goal orientation of the person on the picture was dichotomous (0 = growth, 1 = loss prevention), a multilevel logistic regression was performed, and odds ratios are reported (see Hox, 2002). Across all faces, young and old participants did not differ in their categorizations (odds ratio = 0.89, n.s.). The relative chance to assign old faces to loss-prevention orientation was, however, almost seven times the chance to assign young faces to loss prevention (odds ratio = 6.88, p < .05). Respectively, the relative chance to assign young faces to growth orientation was also almost seven times the chance

to assign old faces to growth. The cross-level interaction effect between age of participant and age of face was not significant (odds ratio = 1.07, n.s.).

With regard to the dimension energy, old, compared with young, participants rated the faces as more than half a standard deviation more energetic ( $\gamma = 0.25$ , p < .05). In addition, old faces were seen as more than half a standard deviation less energetic than young faces by both age groups ( $\gamma = -0.25$ , p < .05). The cross-level interaction effect between age of participant and age of face was not significant ( $\gamma = 0.02$ , n.s.).

In the case of mood ratings, a multilevel multinomial regression specifying two contrasts (sad vs. neutral mood; happy vs. neutral mood) was conducted, because there were three response options for the outcome variable (0 =sad, 1 = happy, 2 = neutral). Young and old participants did not differ in their ratings across all faces, neither for the contrast between sad versus neutral mood (odds ratio = 1.14, n.s.) nor for the contrast between happy versus neutral mood (odds ratio = 1.07, n.s.). When contrasting sad versus neutral mood, the relative chance to assign a sad instead of a neutral mood to old faces was more than two times the chance to assign a sad instead of a neutral mood to young faces (odds ratio = 2.39, p < .05). This tendency was the same for young and old participants (odds ratio = 0.70, n.s.). In addition, the relative chance to assign a happy instead of a neutral mood to old faces was almost one and a half times the chance to assign a happy instead of a neutral mood to young faces (odds ratio = 1.45, p < .05). The lat-

Table 4
Ratings As a Function of Age of Participant and Age of Face: Results of Multilevel Random Coefficient Modeling

				Goal		Mood		
	Attractiveness	Likeability	Distinctiveness	Orientation	Energy	S vs. N	H vs. N	Age
Fixed Effects								
Intercept	1.45*	1.81*	1,72*	0.32*	2.13*	0.17°	0.41*	0.37*
Age of participant	0.46*	0.15	0,29	0.89	0.25*	1.14	1.07	0.13
Age of face	-0.59*	-0.14°	0.08	6.88*	$-0.25^{*}$	2.39*	1.45*	3.88*
Age of participant × age of face	0.30*	-0.01	-0.23°	1.07	0.02	0.70	0.71*	0.08
Random Effects								
Intercept	0.25*	0.20*	0.31*	1.05*	0.17*	0.69*	0.61*	0.05*
Age of face	0.14*	0.09*	0.03*	2.56*	0.11*	0.31*	0.11*	0.26*
Residual	0.85	0.84	0.69	_	0.67	_	_	0.47

Note—S, sad mood, N, neutral mood, H, happy mood. For ranges of response options, see Table 1. Age of participant was coded as young participant = 0 and old participant = 1. Age of face was coded as young face = 0 and old face = 1. Unstandardized regression coefficients are reported for the fixed effects. Estimated variance components are reported for the random effects. Goal orientation: logistic regression, odds ratios reported. p < 0.05.

ter effect was less pronounced in old compared with young participants (odds ratio = 0.71, p < .05).

Finally, in terms of the rated age of the face, young and old participants did not differ across all faces ( $\gamma = 0.13$ , n.s.). As one would expect, old faces were judged older than young faces ( $\gamma = 3.88$ , p < .05), equally so by young and old participants ( $\gamma = 0.08$ , n.s.).

#### **DISCUSSION**

In the present study, equal numbers of young and old participants were asked to rate a selection of 160 faces of young and old individuals from the CAL/PAL Face Database (Minear & Park, 2004) with regard to the dimensions attractiveness, likeability, distinctiveness, goal orientation, energy, mood, and age. The central purpose of the study was to provide face-specific ratings for this selection of young and old faces. Therefore, for each face and each rating dimension, ratings across the total sample, as well as separately across young and old participants, are reported (see the online appendixes). The second purpose of the present study was to investigate whether, and in what regard, young and old faces were perceived differently, and to examine age-group differences in ratings of young and old faces. Analyses pertaining to this second purpose showed that ratings of young and old faces differed in several regards: On average, old faces were judged as less attractive (especially by young participants), less likeable, less distinctive (at least by old participants), less oriented toward growth and more oriented toward loss prevention, less energetic, more likely to be in a happy or sad compared with a neutral mood (however, less so by old participants), and as older. Interestingly, in addition to the differential judgments as a function of the age of the face, old participants evaluated the faces overall as more positive. Specifically, old participants rated the faces as more attractive and more energetic than did young participants.

The finding that old faces were rated as less attractive than young faces confirms the results of an earlier study that found that young and old raters perceive young faces as more attractive than old faces (Wernick & Manaster, 1984). Age-associated changes in shape and texture and coloration of skin and hair, which come with age (Berry & McArthur, 1986; Bruce & Young, 1998; Burt & Perrett, 1995), may strongly influence these lower attractiveness ratings of old compared with young faces. Increased facial wrinkling with advancing age might furthermore lead to the greater chance of assigning happy and sad, relative to neutral mood, to old than to young faces. One could have expected that increased prevalence of facial wrinkles might render a face more distinctive. This assumption was, however, not supported in the present study. Rather, old faces were rated as less distinctive than young faces, at least by old participants. This is in accord with the finding that old participants perceive old faces as more familiar, and in this sense also less distinctive, than do young participants (Bartlett & Fulton, 1991).

Overall, the finding of more negative evaluations of old compared with young faces by young and old participants is very much in line with evidence of a negative aging stereotype (Hummert, Garstka, O'Brien, Greenwald, & Mellot, 2002; Kite & Johnson, 1988). This research shows that old age is generally perceived as more negative than young age, by young as well as old individuals. Only with respect to attractiveness ratings do the present data suggest a self-serving bias of old participants toward their own age group, in that old participants rated old faces as less unattractive than did young participants. This finding is supported by a study by Adams and Huston (1975), who found that old participants stereotyped middle-aged stimulus photographs in more favorable ways than did young participants (the study presented no faces of old individuals).

The present study's finding that young and old faces were rated differentially on several dimensions and that young and old participants differed in some of their ratings of young and old faces has important implications for the selection of facial stimuli for use in age-comparative studies of face processing, face perception, or memory for faces in the future. Experiments on face recognition have, for example, shown that faces that are more deviant or distinctive in appearance are recognized more quickly and accurately than are those that are more typical (Bruce & Young, 1998). The present study found that young and old faces are perceived differently in terms of their distinctiveness, and that young and old participants differed in these ratings of distinctiveness. As a consequence, results of age-related differences in face recognition might confound the age of a face with the distinctiveness of a face. Selecting a set of young and old faces that match with regard to their distinctiveness, on the basis of the information provided for in the present study, will allow researchers to control differences in the degree of distinctiveness of young and old faces and to disentangle the age of a face from the distinctiveness of a face.

In conclusion, evidence on the own-age bias (Anastasi & Rhodes, 2006; Bäckman, 1991; Bartlett & Fulton, 1991; Lamont et al., 2005; Mason, 1986; Perfect & Harris, 2003; Wright & Stroud, 2002) clearly advises that the age of the faces used as experimental stimuli needs to be taken into consideration when investigating participants of different ages. Otherwise, old participants, compared with young participants, will be at a disadvantage when processing, recognizing, or remembering young faces, and young participants will be disadvantaged when presented with old faces. The face-specific information of young and old faces reported in this article, and the findings of differential judgments for young and old faces and of age-group differences in several of these ratings, provide a valuable resource and will assist researchers in selecting those faces that are most appropriate for their specific research endeavors. The present study's findings can serve as the basis for the selection of standardized, age-matched sets of facial stimuli. The information reported in this study can be used to avoid confounds of age-related facial attributes with, or to explicitly examine effects of such attributes on, other variables of interest. In sum, it will allow researchers to qualify their results, facilitate a more systematic study of developmental research questions using faces as experimental stimuli, and considerably contribute to our understanding of whether and why certain (young

and old) faces are perceived, processed, and remembered better than others.

## **Obtaining the Faces**

The original CAL/PAL faces, as well as the modified faces used in this study, can be freely accessed by visiting agingmind.utdallas.edu/facedb/.

#### **AUTHOR NOTE**

This research was funded and conducted at the Max Planck Institute for Human Development, Berlin, Germany. The author thanks Ulman Lindenberger, Alexandra M. Freund, Michaela Riediger, Daniel Grühn, and Christina Röcke for discussions of the study reported in this article. The author also gratefully appreciates the assistance of Dulce Erdt, Jan Steinkraus, Anja Böhnke, Ulrike Altmann, and Axinja Kalusche in data collection. Correspondence concerning this article should be addressed to N. C. Ebner, Department of Psychology, Yale University, P.O. Box 208205, New Haven, CT 06520-8205 (e-mail: natalie.ebner@yale.edu).

#### REFERENCES

- ADAMS, G., & HUSTON, T. (1975). Social perception of middle-aged persons varying in physical attractiveness. *Developmental Psychology*, 11, 657-658.
- ANASTASI, J. S., & RHODES, M. G. (2006). Evidence for an own-age bias in face recognition. *North American Journal of Psychology*, 8, 237-252.
- BÄCKMAN, L. (1991). Recognition memory across the adult life span: The role of prior knowledge. *Memory & Cognition*, 19, 63-71.
- BARTLETT, J. C., & FULTON, A. (1991). Familiarity and recognition of faces in old age. *Memory & Cognition*, 19, 229-238.
- Berry, D. S., & McArthur, L. Z. (1986). Perceiving character in faces: The impact of age-related craniofacial changes on social perception. *Psychological Bulletin*, 100, 3-18.
- Berscheid, E., & Walster, E. (1974). Physical attractiveness. In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 7, pp. 158-216). New York: Academic Press.
- BRIGHAM, J. C., & BARKOWITZ, P. (1978). Do "they all look alike"? The effect of race, sex, experience, and attitudes on the ability to recognize faces. *Journal of Applied Social Psychology*, 8, 306-318.
- BRUCE, V., & YOUNG, A. (1998). In the eye of the beholder: The science of face perception. Oxford: Oxford University Press.
- 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 B*, 259, 137-143.
- GIBSON, J. J. (1979). The ecological approach to visual perception. Boston: Houghton Mifflin.
- GOLBY, A. J., GABRIELI, J. D. E., CHIAO, J. Y., & EBERHARDT, J. L. (2001). Differential responses in the fusiform region to same-race and other-race faces. *Nature Neuroscience*, 4, 845-850.
- Hox, J. (2002). Multilevel analysis: Techniques and applications. Mahwah, NJ: Erlbaum.
- Hummert, M. L., Garstka, T. A., O'Brien, L. T., Greenwald, A. G., & Mellot, D. S. (2002). Using the implicit association test to measure age differences in implicit social cognitions. *Psychology & Aging*, 17, 482-495.
- KITE, M. E., & JOHNSON, B. T. (1988). Attitudes toward older and younger adults: A meta-analysis. Psychology & Aging, 3, 233-244.
- LAMONT, A. C., STEWART-WILLIAMS, S., & PODD, J. (2005). Face recognition and aging: Effects of target age and memory load. *Memory & Cognition*, 33, 1017-1024.
- MASON, S. E. (1986). Age and gender as factors in facial recognition and identification. Experimental Aging Research, 12, 151-154.
- McArthur, L. Z., & Baron, R. M. (1983). Toward an ecological theory of social perception. *Psychological Review*, 90, 215-238.

- MINEAR, M., & PARK, D. C. (2004). A lifespan database of adult facial stimuli. Behavior Research Methods, Instruments, & Computers, 36, 630-633.
- NEZLEK, J. B. (2001). Multilevel random coefficient analyses of eventand interval-contingent data in social and personality psychology research. Personality & Social Psychology Bulletin, 27, 771-785.
- Perfect, T. J., & Harris, L. J. (2003). Adult age differences in unconscious transference: Source confusion or identity blending? *Memory & Cognition*, 31, 570-580.
- PSYCHOLOGICAL IMAGE COLLECTION AT STIRLING (PICS) (n.d.). University of Stirling Psychology Department, pics.psych.stir.ac.uk/.
- RAUDENBUSH, S. W., & BRYK, A. S. (2001). Hierarchical linear models: Applications and data analysis methods (2nd ed.). Thousand Oaks, CA: Sage.
- RAUDENBUSH, S. [W.], BRYK, A. [S.], & CONGDON, R. (2000). HLM 6: Hierarchical linear and nonlinear modeling [Computer software]. Chicago: Scientific Software International.
- STEINKRAUS, J., & EBNER, N. C. (2006). A manual to Psy-Point. Unpublished manual, Max Planck Institute for Human Development, Berlin.
- WECHSLER, D. (1981). Manual for the Wechsler Adult Intelligence Scale-Revised (WAIS-R). New York: Psychological Corporation.
- WERNICK, M., & MANASTER, G. J. (1984). Age and the perception of age and attractiveness. *Gerontologist*, 24, 408-414.
- WRIGHT, D. B., & STROUD, J. N. (2002). Age differences in lineup identification accuracy: People are better with their own age. Law & Human Behavior, 26, 641-654.

#### NOTE

1. Due to time reasons, one old man did not finish the computer task.

#### **ARCHIVED MATERIALS**

The following materials associated with this article may be accessed through the Psychonomic Society's Norms, Stimuli, and Data archive, www.psychonomic.org/archive.

To access these files, search the archive for this article using the journal name (*Behavior Research Methods*), the author's name (Ebner), and the publication year (2008).

FILE: Ebner-BRM-2008.zip.

DESCRIPTION: The compressed archive file contains fourteen files:

EbnerAPP-A\_Attractiveness.txt, containing attractiveness files.

EbnerAPP-A\_Attractiveness.xls, containing the above information in Excel spreadsheet format.

EbnerAPP-B\_Sympathy.txt, containing sympathy files.

EbnerAPP-B\_Sympathy.xls, containing the above information in Excel spreadsheet format.

EbnerAPP-C\_Distinctiveness.txt, containing distinctiveness files.

EbnerAPP-C\_Distinctiveness.xls, containing the above information in Excel spreadsheet format.

EbnerAPP-D\_Goal Orientation.txt, containing goal orientation files. EbnerAPP-D\_Goal Orientation.xls, containing the above information in Excel spreadsheet format.

EbnerAPP-E\_Energy.txt, containing energy files.

EbnerAPP-E\_Energy.xls, containing the above information in Excel spreadsheet format.

EbnerAPP-F\_Mood.txt, containing mood files.

EbnerAPP-F\_Mood.xls, containing the above information in Excel spreadsheet format.

EbnerAPP-G\_Age.txt, containing age files.

EbnerAPP-G\_Age.xls, containing the above information in Excel spreadsheet format.

AUTHOR'S E-MAIL ADDRESS: natalie.ebner@yale.edu.

(Manuscript received May 18, 2007; revision accepted for publication July 18, 2007.)