Impairment of Age Estimation from Faces in Alzheimer’s Disease

Evelyne Moyse*, Christine Bastinb, Eric Salmonb,c and Serge Brédarta

Department of Psychology: Cognition and Behavior, University of Liège, Belgium
bCyclotron Research Center, University of Liège, Belgium
cMemory Clinic, CHU Liège, Belgium

Handling Associate Editor: Hanna Chainay

Accepted 17 December 2014

Abstract. A prerequisite for any function in social cognition is the perception and processing of social cues. Age estimation is a skill that is used in everyday life and is fundamental in social interactions. This study evaluated whether facial age estimation is impaired in patients with mild to moderate Alzheimer’s disease (AD). The current age of faces is known to have an impact on age estimation, and therefore stimuli belonging to different age groups (young, middle-aged, and older adults’ faces) were used. As expected, an impairment of age estimation from faces was observed in mild to moderate AD patients. However, the profile of impairment depended on the age of faces and stage of the disease. Mild AD patients presented difficulties mainly in assessing the age of middle-aged adults. In moderate disease stage, these difficulties also affected the age estimation of young adult faces. Interestingly, AD patients remained relatively good at estimating the age of older adults’ faces, compared to healthy controls.

Keywords: Age discrimination, Alzheimer’s disease, face, social perception

INTRODUCTION

Social cognition involves the perception and processing of other people’s faces, emotional states, voices, and postures, as well as making judgments about others’ personalities, intentions, and likely behaviors, and planning one’s own interactions with other people. Social cognition, then, requires various cognitive processes such as sensory and perceptual functions, selective attention, executive function, working memory, language and communicative skills in a variety of situations [1]. Research on social cognition in Alzheimer’s disease (AD) has mainly focused on theory of mind and emotional facial expressions and has reported that this capacity is affected early in AD ([2–4], for a review, see [5]). However, a prerequisite for any social cognition function is the perception and processing of social cues [5]. Faces are often the main focus when people perceive others, and facial appearance is one of the most important factors influencing human social interactions [6]. Information about an unknown person, such as age, gender, and race, can be readily and quite automatically perceived from a face [7]. Age is a fundamental factor in social interactions. People’s perceived age influences how we behave with them [8]. For instance, a baby face elicits approach and protective responses [9]. Moreover, linguistic style differs according to the interlocutor’s age [10]. People tend to speak to children with “baby talk,” which consists in using diminutives, reduplications, and more concrete verbs [11]. In the same way, the speech register people use with older adults is marked by decreases in speaking rate, sentence length, the use of complex syntactic constructions, and propositional density,
and increases in the numbers of words, utterances, instructions, and repetitions. The increase in the use of semantic and discourse markers such as expansions and repetitions has been shown to benefit older listeners, especially those with cognitive impairments [12].

The ability to perceive the age of faces, an important determinant for social interactions, has rarely been explored in AD. An age estimation task is sometimes used during neuropsychological examinations to test the preservation of face perception in AD [13]. In such a task, participants are asked to classify faces from the youngest to the oldest without making a precise age estimation. Della Sala et al. found that only 20% of mildly deteriorated AD participants (6/30) performed below the cut-off score calculated from the score distribution of a normal population in the age estimation task [13]. The main objective of the present study was to evaluate precise age estimation from faces in mild to moderate AD patients in comparison with matched controls.

Faces convey numerous cues about age (for a review, see [14]). Normally, people use these different cues in a coordinated manner. However, the unavailability of one type of cue does not inevitably disrupt the accuracy of age estimation; the remaining cues are sufficient to allow fairly correct age estimation [15, 16]. According to George and Hole [16], there are at least two routes for age perception: The first one is based on surface cues, such as wrinkling, skin pigmentation, and the spatial configuration of the internal facial features, and the second is based on shape cues such as head shape, cardioidal strain level, and 3D configuration. Moreover, it seems that age estimation from faces is related to general face processing capacity. Indeed, De Renzi et al. reported a significant correlation between a facial matching test and a face age classification test [17].

In the specific case of AD, several studies have demonstrated that face processing, evaluated by means of a facial matching test, is impaired in mild to moderate AD patients compared with matched controls [13, 18–21]. However, there have been different findings with regard to the relationship of disease severity and face processing task performance. Becker et al. found a significant correlation, indicating that the probability of impairment in a face processing task increased as a function of disease severity [18], but other studies showed no significant correlation between disease severity and face processing ability [13, 19, 21].

Finally, the importance of cues for age estimation differs according to the age of faces (e.g., [15, 22]). For example, George and Hole reported that a transformation that affects skin texture has a greater impact on older faces than on younger faces [16]. Similarly, Burt and Perrett showed that transformations of shape and color were more effective at increasing the perceived age of young faces than of older faces [23]. Moreover, accuracy of age estimation is influenced by the actual age of faces and of the participants themselves. For example, Sorqvist and Eriksson showed that young adults’ age estimation performance was more accurate for young stimuli (15 to 24 years) than for middle-aged (34 to 46 years) or older (56 to 65 years) stimuli [24]. Voevild et al. indicated that young and middle-aged participants were more accurate than older participants in estimating the age of faces aged 19 to 80 years [25]. Thus, it appears that there is an overall decline in estimation accuracy as the age of the faces increases and as the participants’ age increases. Moreover, some studies have also reported an interaction between the age of faces and the age of participants. In these studies [15, 26], young adults were more accurate than older adults in estimating the age of young adults’ faces; conversely, older adults were more accurate than young adults in estimating the age of older adults’ faces. This phenomenon is known as the own-age bias. The occurrence of an own-age bias has never been examined in AD patients. However, the age of faces may modulate the probable impairment in age estimation observed in AD patients. Indeed, previous research [15, 26] reported that, while healthy older adults show an overall decline in accuracy, their performance at estimating the age of older adults’ faces appears similar to or even better than that of young adults when the age of the faces is taken into account.

Given (1) the relationship between general face processing and age estimation, and (2) the occurrence of face processing deficits in AD, this study was based on the hypothesis that AD patients would show inaccurate age estimation from faces and that this impairment would correlate with impairment during a face processing task. Moreover, the study aimed to examine age estimation as a function of disease severity (mild versus moderate dementia). Finally, the impact of the actual age of faces was examined by using stimuli from different age groups.

METHODS

Participants

Forty-five patients (20 women) diagnosed with probable AD according to the National Institute on Aging and Alzheimer’s Association (NIA-AA) criteria [27] participated in this study. Patients were recruited
via memory clinics in Li`ege area. Diagnosis was based on general examination and neurological and neuropsychological assessments performed in the memory clinics. Hippocampal atrophy on structural magnetic resonance image was taken as biomarker. The disease stage classification was based on the Mini-Mental State Examination (MMSE) [28]. The patients were separated into two groups: mild AD and moderate AD. Mild AD comprised 21 patients (6 women) in the mild stage of the disease (indicated by an MMSE score above 20), and moderate AD comprised 24 patients (14 women) in the moderate stage of the disease (with an MMSE score between 10 and 20).

Forty-five healthy older controls (20 women) were recruited from seniors’ organizations and among acquaintances of the experimenters’ relatives and friends. Healthy controls had no history of neurological or psychiatric impairment. They were selected to match the AD patients based on age ($F(2, 87) = 0.58$, $p = 0.56$), years of education ($F(2, 87) = 1.59$, $p = 0.21$), and gender ($\chi^2(2) = 4.02$, $p = 0.13$). On the MMSE, the three groups scored significantly differently, $F(2, 87) = 377.58$, $p < 0.001$. The control group’s score was higher than that of AD patients, while mild AD patients scored higher than moderate AD patients. All participants (and close relatives for AD patients) gave their written informed consent to participate in the study, which was approved by the ethics committee of the Faculty of Psychology, University of Li`ege, according to the Declaration of Helsinki. Demographic characteristics of the AD and control groups are presented in Table 1.

### Stimuli
Digital photographs were taken of 16 Caucasian 20- to 30-year-old adults (8 women; $M = 25.00$, $SD = 3.39$), 16 Caucasian 45- to 55-year-old adults (8 women; $M = 50.31$, $SD = 3.09$), and 16 Caucasian 70- to 88-year-old adults (8 women; $M = 78.94$, $SD = 5.98$). All the photographs were close-up, frontal-view portraits of faces showing a neutral expression, without glasses, jewelry, or facial hair. The stimuli were color photographs with a white background and were resized to 350 × 525 pixels using Adobe Photoshop.

A pilot study, carried out on 20 participants (10 females; mean age = 20.6 years, $SD = 2.8$), allowed us to select 48 photographs from the 80 stimuli on the basis of both the quality of the pictures and the emotional valence of the faces. The selected stimuli were judged to have good picture quality (i.e., they had a score of 5 or more on a scale ranging from 1 “very poor quality” to 7 “very good quality”) and were emotionally neutral (i.e., they had a score between 3 and 5 on a scale ranging from 1 “very positive” to 7 “very positive”).

### Procedure
Participants were tested individually at their homes. Two tasks were administered. The first one was an age estimation task, which included 48 stimuli presented on a computer using E-prime software. Stimuli were presented randomly. At the beginning of the experimental session, participants were informed that they were going to see male and female faces of different ages. They were instructed to evaluate the age of each face as accurately as possible by orally giving a specific number (e.g., “23 years,” not “about 20”). The experimenter instructed them to respond at their own pace. Each trial started with a fixation cross that appeared for 1000 ms, followed by the face to be evaluated, which remained on the screen until the participant responded. The experimenter noted each response before pressing the space bar to start the next trial.

After the age estimation task, unfamiliar face processing was assessed with the short form of the Benton Facial Recognition Test [29]. The test consists of 13 items where a target face is presented and participants are requested to match it to one or more faces within an array of six photographs. Three matching conditions are presented: photographs are either identical frontal-view facial portraits, taken from different angles, or taken under different lighting conditions. Maximum score is 54.

### RESULTS

#### Age estimation task
A mixed design was used, with one between-participants factor (Group: three levels: mild AD, moderate AD, and control) and one within-participants factor (Order of Groups: two levels: AD, control first, and control, AD first). There was a significant main effect of Group, $F(2, 87) = 13.36$, $p < 0.001$, with the control group showing the highest accuracy, followed by the moderate AD group, and then the mild AD group. There was also a significant main effect of Order, $F(1, 87) = 4.73$, $p = 0.03$, with the control group showing higher accuracy when the control condition was first. There was no significant interaction between Group and Order, $F(2, 87) = 1.40$, $p = 0.25$. The results suggest that the control group was more accurate in distinguishing between faces of different ages than the AD groups, and that the order of the conditions affected the performance, with the control group tending to show better performance when it was tested first.
factor (Age of faces, three levels: young, middle-aged, and older adults). Three dependent measures were calculated for each participant: 1) the directional estimation error was calculated by averaging, for each stimulus category, the absolute value of the directional estimation error for each face; and 3) the lack of precision was obtained by calculating the proportion of responses consisting in mere multiples of 5 (e.g., 20, 25, 30 . . . years old) for each stimulus category. Because several conditions of application for parametric ANOVAs were not respected (i.e., normality of distribution and homogeneity of variances), non-parametric analyses were carried out. Kruskal-Wallis tests and Mann-Whitney tests were used for the inter-group comparisons. Moreover, a Bonferroni correction was applied on the statistical significance level, with \( p < 0.0167 \) being considered as significant.

**Direction of error**

For young adults’ faces, a Kruskal-Wallis test revealed a significant main effect of Group, \( H(2) = 13.90, p = 0.001 \). Mann-Whitney tests indicated that overestimation was higher for the Moderate AD group (\( M = 4.74, SD = 7.05 \)) than for the Mild AD group (\( M = –0.29, SD = 6.18 \); \( U = 113.5, Z = –3.14, p = 0.002 \), Cohen’s \( r = 0.47 \)) and for the Control group (\( M = –0.11, SD = 3.14; U = 286.5, Z = –3.19, p = 0.001 \), Cohen’s \( r = 0.38 \)); there was no significant difference between the Mild AD and Control groups. By contrast, no significant main effect of Group was observed for middle-aged adults’ faces, \( H(2) = 0.88, p = 0.65 \), or older adults’ faces, \( H(2) = 2.09, p = 0.35 \).

Friedman tests revealed a significant main effect of Age of faces for the Control group, \( \chi^2(2) = 19.00, p < 0.001 \), the Mild AD group, \( \chi^2(2) = 9.81, p = 0.007 \), and the Moderate AD group, \( \chi^2(2) = 30.33, p < 0.001 \). Wilcoxon tests indicated the same pattern of results in the Control and Moderate AD groups: greater underestimation of older adults’ faces than of middle-aged adults’ faces (Control group: \( T = 293.5, Z = –2.98, p = 0.003 \), Cohen’s \( r = 0.44 \); Moderate AD group: \( T = 15, Z = 3.86, p < 0.001 \), Cohen’s \( r = 0.79 \), and greater underestimation of middle-aged adults’ faces than of young adults’ faces (Control group: \( T = 187, Z = 3.59, p < 0.001 \), Cohen’s \( r = 0.53 \); Moderate AD group: \( T = 47.5, Z = 2.93, p = 0.003 \), Cohen’s \( r = 0.60 \)). In the Mild AD group, the only significant difference observed was between older (\( M = –6.49, SD = 7.57 \)) and young adults’ faces (\( M = –0.20, SD = 6.18; T = 42, Z = 2.55, p = 0.011 \), Cohen’s \( r = 0.56 \)) (see Fig. 1).

**Amplitude of error**

For young adults’ faces, a Kruskal-Wallis test revealed a significant main effect of Group, \( H(2) = 12.13, p = 0.002 \). Mann-Whitney tests indicated that amplitude of error was higher for the Moderate AD group (\( M = 8.00, SD = 5.91 \)) than for the Control group (\( M = 4.59, SD = 1.29; U = 277, Z = –3.31, p < 0.001 \), Cohen’s \( r = 0.40 \)). No significant difference was observed between the Mild AD group (\( M = 5.98, SD = 4.06 \)) and the other two groups. For middle-aged adults’ faces, a significant main effect of Group was also observed, \( H(2) = 21.26, p < 0.001 \). Mann-Whitney tests showed that the Control group (\( M = 7.53, SD = 2.40 \)) was more accurate than the Mild AD (\( M = 9.75, SD = 2.59; U = 242.5, Z = –3.16, p = 0.002 \), Cohen’s \( r = 0.39 \)) and Moderate AD groups (\( M = 10.71, SD = 3.12; U = 211, Z = –4.14, p < 0.001 \), Cohen’s \( r = 0.50 \), with no significant difference between AD groups. Finally, for older adults’ faces, no significant main effect of Group was observed, \( H(2) = 5.75, p = 0.06 \).

Friedman tests revealed a significant main effect of Age of faces for the Control group, \( \chi^2(2) = 26.03, p < 0.001 \), the Mild AD group, \( \chi^2(2) = 16.65, p < 0.001 \), and the Moderate AD group, \( \chi^2(2) = 9.08, p = 0.011 \).

![Fig. 1. Signed values for the three groups according to the age of faces. Error bars represent the semi-interquartile range (SIQR).](image-url)
Wilcoxon tests indicated the same pattern of results in all groups: a lower amplitude of error for young adults' faces than for middle-aged (Control group: $T = 71.5$, $Z = 5.03$, $p < 0.001$, Cohen’s $r = 0.75$; Mild AD group: $T = 33$, $Z = 2.87$, $p = 0.004$, Cohen’s $r = 0.63$; Moderate AD group: $T = 61$, $Z = 2.54$, $p = 0.011$, Cohen’s $r = 0.52$), with no significant difference between middle-aged and older adults’ faces (see Fig. 2).

**Imprecision**

For young adults’ faces, a Kruskal-Wallis test revealed a significant main effect of Group, $H(2) = 8.28$, $p = 0.015$. Mann-Whitney tests indicated that percentage of imprecision was higher in the Moderate AD group ($M = 41.93$, $SD = 19.89$) than in the Control group ($M = 26.39$, $SD = 21.85$, $U = 292.5$, $Z = –3.11$, $p = 0.002$, Cohen’s $r = 0.37$). No significant difference was observed between the Mild AD group ($M = 37.20$, $SD = 33.39$) and the other two groups. No significant main effect of Group was observed for middle-aged adults’ faces, $H(2) = 5.17$, $p = 0.07$. Finally, for older adults’ faces, a significant main effect of Group was also observed, $H(2) = 12.84$, $p = 0.001$. Mann-Whitney tests showed that percentage of imprecision was lower in the Control group ($M = 33.05$, $SD = 27.42$) than in the Mild AD ($M = 59.82$, $SD = 32.75$, $U = 251.5$, $Z = –3.03$, $p = 0.002$, Cohen’s $r = 0.37$) and Moderate AD groups ($M = 50.78$, $SD = 27.23$, $U = 327$, $Z = –2.68$, $p = 0.007$, Cohen’s $r = 0.32$), with no significant difference between AD groups.

Friedman tests revealed a significant main effect of Age of faces for the Control group, $\chi^2(2) = 16.34$, $p < 0.001$, and the Mild AD group, $\chi^2(2) = 10.36$, $p = 0.006$. Wilcoxon tests indicated the same pattern of results in these two groups: a lower percentage of imprecision for young adults’ faces than for middle-aged (Control group: $T = 111.5$, $Z = 4.12$, $p < 0.001$, Cohen’s $r = 0.61$; Mild AD group: $T = 16$, $Z = 3.03$, $p = 0.002$, Cohen’s $r = 0.66$) and older adults’ faces (Control group: $T = 223$, $Z = 2.51$, $p = 0.01$, Cohen’s $r = 0.37$; Mild AD group: $T = 19$, $Z = 3.06$, $p = 0.002$, Cohen’s $r = 0.67$), with no significant difference between middle-aged and older adults’ faces. However, in the Moderate AD group, no significant main effect of Age of faces was observed, $\chi^2(2) = 5.45$, $p = 0.06$ (see Fig. 3).

**Benton facial recognition test**

A one-way ANOVA on the Benton test scores revealed a significant difference between the three groups, $F(2,87) = 12.72$, $p < 0.001$, $\eta^2 = 0.23$. Planned comparisons showed that performance was better in the Control group ($M = 48.29$, $SD = 3.55$) than in the Mild...
AD (M = 44.95, SD = 4.48) and Moderate AD groups (M = 42.87, SD = 5.63), with no significant difference between the two AD groups.

A significant Spearman’s correlation was observed between performance on the Benton Facial Recognition Test and MMSE score in the whole group of AD patients, indicating poorer face processing in patients with more advanced global cognitive impairment (r = 0.33, n = 45, p = 0.02).

Correlations

To assess whether AD patients’ relatively poorer age estimation performance was related to dementia severity, Spearman’s correlations were computed between each of the three measures of age estimation (calculated across the age of faces) and MMSE scores in the whole group of AD patients. None of these correlations were significant. Details of these correlations are presented in Table 2.

To assess whether age estimation performance was related to face processing performance, as shown by De Renzi et al. [17], Spearman’s correlation tests were conducted between each of the three age estimation measures and the Benton Facial Recognition Test score in the three groups. As shown in Table 3, the Moderate AD group showed a significant negative correlation between amplitude of error and the Benton Facial Recognition Test score, suggesting that larger amplitudes of error were associated with poorer face processing. However, such a moderate correlation should be viewed with caution.

DISCUSSION

This study examined the accuracy of age estimation from faces in mild to moderate AD patients in comparison with matched controls. As expected, an impairment of age estimation from faces was observed in AD patients. However, the profile of impairment depended on the age of the faces and the stage of the disease, although disease severity was not linearly related to age estimation performance. Mild AD patients mainly presented difficulties in estimating the age of middle-aged faces, characterized by high amplitudes of error. In moderate disease stage, these difficulties also affected the age estimation of young faces, with high amplitudes of error and overestimation of age. For the estimation of the age of older adults, AD patients and their matched controls differed only with respect to their response strategies: controls’ responses were relatively precise, while AD patients’ responses were expressed in multiples of 5. Addis and Tippett reported that AD patients have a general bias toward vague and abstract responses [30]. Thus, the AD patients’ estimations in terms of multiples of 5 seem to reflect this general response bias rather than age estimation problems.

Previous research showed that age cues differ as a function of the age of faces (e.g., [15, 22]). For example, skin texture seems to be more important in estimating the age of older faces [16] whereas shape (cardioidal strain) seems more important for young faces [15]. So the performance of AD patients may be due to a selective deficit in the perception of these cues, and this deficit may become stronger as the disease progresses. To test this hypothesis, future research should use face transformation methods (inversion, removing some internal features, negation) to determine whether AD patients, like healthy adults, are able to use surface information when estimating the age of older faces and shape information when estimating the age of young faces.

Moreover, AD patients showed impaired face processing. This result adds further evidence to a body of data reporting a face processing deficit in AD using the Benton Facial Recognition Test [13, 18–21]. However, this test may not be really appropriate for examining the preservation of face processing capacity. Indeed, Duchaine and Weidenfeld [31] showed that, although impaired performance on the Benton Facial Recognition Test indicates impaired face processing, it is far from certain that a normal score necessarily reflects
intact face processing. This test may lack sensitivity. Moreover, it measures a general face processing capacity and does not assess the involvement of configural and featural facial processing. Therefore, future research should be conducted using more sensitive tasks that appropriately distinguish between the different processes involved in face processing. Given that age perception requires the processing of both local and global facial features ([16], for a review, see [32]), the relationship between performance on a task that taps into either configural or featural facial processing and age estimation in AD should also be examined.

Previous research reported that, although older adults performed worse on age estimation of faces than young adults, when the age of stimuli was taken into account, they were better than young adults at estimating the age of older adults’ faces [15, 26]. Based on these observations of an own-age bias in age estimation in older adults, we hypothesized that the probable impairment in age estimation observed in AD patients would be modulated by the age of the faces. Interestingly, our results support this hypothesis: in age estimation of older adults’ faces, no significant difference in performance was observed between the AD groups and their matched controls, as measured with absolute value and directional estimation error. Thus, although a real own-age bias in AD groups, with better performance for own-age faces than other-age faces, was not found, the AD patients remained relatively able to estimate the age of own-age faces. Given that this study is the first to examine the impact of age of faces on age estimation in AD, additional research is necessary to confirm this result. Moreover, it might also be interesting to examine general face processing capacity in AD by using tests with stimuli belonging to different ages. For instance, the Benton Facial Recognition Test includes only young adult stimuli and thus may accentuate the face processing deficit in AD.

Research on social cognition in AD has mainly investigated theory of mind and recognition of emotional facial expressions, suggesting that both are affected early in AD (for a review, see [5]). Yet more basic capacities, such as estimating the interlocutor’s age, that are important determinants of social interactions, have not been previously investigated. Our study is the first to provide evidence that the perception of age of faces is impaired in the early stages of AD. These AD-related difficulties perceiving and processing social cues such as faces, and specifically the age of faces, could be an additional obstacle for adjustment during social interactions. Interestingly, our findings also revealed that there is no global impairment of age estimation in AD, but rather that the ability to estimate the age of individuals is modulated by the actual age of the stimuli faces. Generally, this study is important in that it points to a deficit in a capacity that represents a prerequisite for appropriate social interaction. Future work should determine which specific impaired processes are responsible for inaccurate age estimation and the extent to which this impairment contributes to altered social cognition in AD.

ACKNOWLEDGMENTS

C. Bastin was supported by grants from the Interuniversity Pole of Attraction P711 and from SAO/FRA Belgium and by the Belgian National Fund of Scientific Research (F.R.S.-FNRS).

Authors’ disclosures available online (http://j-alz.com/manuscript-disclosures/14-2253r1).

REFERENCES