SECCIÓN EXPERIMENTAL

Psicológica (2009), 30, 155-179.

Identity, gender, and the role of age of acquisition in face processing

Rachel M. Richards* & Andrew W. Ellis

University of York (UK)

Two experiments examined the effects of age of acquisition (AoA) and the gender of stimulus faces on familiarity decisions (Experiment 1) and gender decisions (Experiment 2) to the same set of famous and unfamiliar faces presented as whole faces, internal features or external features. In Experiment 1, familiarity decisions were faster to whole faces than to internal or external features. Famous faces with early AoA were recognised faster than later acquired faces, though the effect was only reliable for famous male faces, and for whole faces and internal features rather than for external features. In Experiment 2, gender decisions were made more rapidly to whole faces than to internal or external features. Classification was faster to famous than to unfamiliar faces when the faces were presented as internal features or external features, but not when they were presented as whole faces. More gender classification errors were made to famous than to unfamiliar male faces, but there was no effect of familiarity on the accuracy of responses to female faces. AoA had no effect on gender classification of whole faces or external features. Classification from internal features was faster for early than for late acquired male faces, but faster for late than for early female faces. In the light of the anomalous results for female faces, responses to male faces were analysed separately. The results for the male faces form the focus of the discussion.

The influential Bruce and Young (1986) model of face processing proposed that different aspects of face perception are handled by different specialist processing subsystems. After an initial stage of perceptual encoding, information about a seen face is fed into subsystems responsible for recognising familiar faces (face recognition units), decoding facial expressions (expression analysis), analysing facial speech movements

* The first author was the recipient of a PhD studentship from the UK Biology and Biotechnology Research Council. Correspondence should be addressed to Andy Ellis, Department of Psychology, University of York, York YO10 5DD, UK. E-mail: a.ellis@psych.york.ac.uk

(facial speech analysis), and so on. The process of deciding whether a face is male or female (referred to in the literature as either sex decision or gender decision) was assigned to a component labelled 'directed visual processing', along with other decisions that can be made to both familiar and unfamiliar faces, such as deciding on the probable age of a face.

A clear prediction from the Bruce and Young (1986) model is that performance on some face perception tasks should occur independently of whether a face is familiar or unfamiliar. Tasks which involve processes independent of those responsible for recognising familiar faces should be capable of being carried out with the same speed and accuracy whether the stimulus faces are highly familiar or are being seen for the first time. Much of the initial research carried out to test this aspect of the Bruce and Young (1986) model supported the predictions. For example, Bruce (1986) reported that decisions based on the expressions on stimulus faces could be made as quickly to unfamiliar as to familiar faces, while Young, McWeeny, Hay, and Ellis (1986) found that two different images of a face could be classified as being the same person more quickly if the faces belonged to a famous rather than an unfamiliar face, but that decisions as to whether two faces were showing the same or different emotional expressions showed no speed advantage for famous compared with unfamiliar faces. Campbell, Brooks, De Haan, and Roberts (1996) found no effect of the familiarity of faces on the speed of expression decisions or the speed of extracting speech-related information from faces.

The particular concern of this paper is the independence or otherwise of the processes that judge the gender of a face from the processes that decide whether the face is familiar or unfamiliar. A. Ellis, Young, and Flude (1990) and Stevenage and Osborne (2006) found no effect of familiarity on gender decisions to whole faces, which is compatible with the proposal embodied in the Bruce and Young (1986) model that these decisions are taken by independent processes. Bruce (1986), however, reported slightly faster gender decisions to familiar than to unfamiliar faces. Upon close inspection of the items used in the experiment, the effect was shown to be carried by two male faces whose gender was relatively difficult to categorise. A follow-up study by Bruce, Ellis, Gibling, and Young (1987) employed sets of famous and unfamiliar male faces that had been rated as high masculinity or low masculinity. The male faces were interleaved with an equal number of famous and unfamiliar female faces and presented to participants for familiarity decisions or gender decisions. The rated masculinity of the male faces affected the speed of gender decisions but not the speed of familiarity decisions.

Human beings are generally accurate at judging the gender of unfamiliar faces on the basis of cues such as skin texture, facial structure, signs of facial hair, and hair style (Brown & Perrett, 1993; Burton, Bruce, & Dench, 1993; Roberts & Bruce, 1988). Bruce (1986) suggested that if gender decisions can be made easily on the basis of such superficial features, then the decision will be made without recourse to identity information. If, on the other hand, a face is hard to classify on the basis of surface features, then the decision may be informed by semantic knowledge if the face belongs to a familiar person. That may speed up the decision making process in comparison with the situation where a hard-to-classify face is unfamiliar and can only be judged on its surface properties. This line of reasoning stimulated further studies in which the ease or difficulty of making gender decisions on the basis of surface information was deliberately manipulated. One way to increase the difficulty of making gender decisions to faces is to remove the external features (hair style and face outline) leaving the eyes, nose, cheeks and mouth region. Rossion (2002) found faster gender decisions to the internal features of faces which participants had been trained to recognise than to the internal features of novel faces. Clutterbuck and Johnston (2004) replicated that effect, and found an even larger effect of familiarity on gender processing speed for the internal features of famous faces. Stevenage and Osborne (2006) found no effect of familiarity on gender decisions to whole faces regardless of whether they were presented in normal upright form or rotated by various amounts, but internal features showed an effect of familiarity, especially when rotated to 30 or 60 degrees from vertical. These findings are compatible with the notion that semantic knowledge about familiar faces (either famous or recently-learned) may be called upon to assist the process of deciding whether a face is male or female if that decision cannot easily be made on the basis of surface features, as in the case of images showing the internal features only. Under such circumstances, decisions should be made more quickly and more accurately to familiar than to unfamiliar faces (which lack semantic information).

The theoretical approach just described allows for a more strategic deployment of processing resources, with semantic information being recruited in some circumstances but not others. It does not necessarily undermine the idea that the processing mechanisms involved may be fundamentally distinct. Gender processing still could proceed within a system like Bruce and Young's (1986) 'directed visual processing system' that is independent of, and insulated from, the processes used to recognise familiar faces and activate semantic information: it is just that participants performing particular tasks would be able to draw strategically upon the

results of those independent processing strands when performing different tasks under different conditions (in this case, taking note of the results of semantic processing when making gender decisions to internal feature images more than when making gender decisions to whole face images).

The present study sought to shed further light on the independence or otherwise of the processes responsible on the one hand for deciding whether faces are male or female, and on the other hand whether faces are familiar or unfamiliar. Participants made speeded familiarity decisions (Experiment 1) or gender decisions (Experiment 2) to famous and unfamiliar faces presented as whole-face images, internal features only, or external features only. On the basis of the previous studies, we predicted that in Experiment 1, whole faces would be recognised as familiar more easily than internal features, which might be easier to recognise than external features (cf. Clutterbuck & Johnston, 2002; H. Ellis, Shepherd, & Davies, 1979; Young, Hay, McWeeny, Flude, & A. Ellis, 1985). We expected that in Experiment 2, gender decisions would be faster to whole faces than to internal features, and that gender decisions made to internal features would show more of an effect of familiarity than decisions made to whole faces (Clutterbuck & Johnston, 2004; Rossion, 2002; Stevenage & Osborne, 2006).

Age of acquisition, identity and gender

Experiments 1 and 2 manipulated more than just the matter of whether a face was famous or unfamiliar. The famous faces also varied on their *age of acquisition* (AoA). A large body of research indicates that, all other things being equal, early acquired words, objects and faces are processed more rapidly and more accurately than later acquired words, objects and faces in a wide variety of tasks (see Johnston & Barry, 2006, and Juhasz, 2005, for reviews). Faster recognition of early than later acquired faces has been reported in three studies. Moore and Valentine (1998) found faster naming times for early than for late acquired famous faces matched on rated familiarity and distinctiveness. Moore and Valentine (1999) mixed famous faces with unfamiliar faces and presented them to participants who were required to decide as quickly as possible whether each face they saw was famous or unfamiliar (a *familiarity decision task*). Reaction times were faster to early than to late acquired famous faces matched on rated familiarity and distinctiveness. Lewis (1999) required participants to classify the faces of actors according to which of two television series they appeared in. Correlational analyses found a significant effect of the length of time that actors had appeared in their shows that was independent of the significant effect of a measure of how often their

characters appeared in the programmes. Given that the actors concerned tended not to have been known for other roles before they began to appear on their shows, the ones who had spent more time in the series will generally have been earlier acquired.

Current theories assume that AoA exerts its influence within the processes that recognise faces, objects or words as being familiar (Johnston & Barry, 2006; Juhasz, 2005). Brysbaert and colleagues have proposed that the age (or order) of acquisition of words affects the quality of the semantic representations created for those items, and that the strength of AoA effects in different tasks will depend on the degree of semantic involvement (e.g., Brysbaert, van Wijnendaele, & de Deyne, 2000; Ghyselinck, Custers, & Brysbaert, 2004; Ghyselinck, Lewis, & Brysbaert, 2004; see also Steyvers & Tenenbaum, 2005; van Loon-Vervoorn, 1988). Frequency plays a part in determining processing speed, but the effect of AoA cannot be explained in terms of differences in frequency (or cumulative frequency) between early and late acquired items (Bonin, Barry, Méot, & Chalard, 2004; Cortese & Khanna, 2007; Ghyslinck et al., 2004). Deciding that a face is familiar is assumed to involve access to semantic representations. The familiarity decision task (*Is this face famous or unfamiliar?*) should therefore show effects of AoA whether the stimuli being judged are whole faces, internal features or external features. In contrast, the evidence reviewed above suggests that gender decisions may be made to whole faces without reference to semantic information. When gender decisions are made more difficult, for example by presenting only internal features, then semantic information may contribute to gender decisions, possibly introducing an influence of AoA.

The current study presents two experiments. Experiment 1 involved speeded familiarity decisions (famous or unfamiliar?) while Experiment 2 involved speeded gender decisions (male or female?). The same mixture of famous and unfamiliar faces was used in each experiment. The famous faces varied on AoA while being matched on distinctiveness, familiarity, frequency of encounter, and the degree to which the image was regarded as a good likeness of the famous person. The stimuli were presented in each experiment as whole faces, internal features or external features. On the basis of Moore and Valentine (1999) and reports of AoA effects in familiarity decisions to object pictures (Holmes & Ellis, 2006; Moore, Smith-Spark, & Valentine, 2004), we expected to see an effect of AoA on familiarity decisions to the famous faces. We expected to see little or no effect of AoA on gender decisions to whole faces in Experiment 2, and little or no difference in reaction times to famous and unfamiliar whole faces. We predicted that the greater difficulty of making gender decisions to internal

features in particular would lead to more involvement of semantic knowledge in the gender decision process and therefore influences of familiarity (cf. Clutterbuck & Johnston, 2004; Rossion, 2002, Stevenage & Osborne, 2006) and AoA. Because much of the theoretical interest of the results depends on the comparison of familiarity decisions with gender decisions, we will present both experiments before discussing them together.

EXPERIMENT 1

Participants in Experiment 1 made familiarity decisions to whole faces, internal features and external features. Half the faces were famous and half were unfamiliar. Half of each set were male and half were female. The famous faces were rated by a separate group of participants as either early acquired or late acquired, and were matched on distinctiveness, frequency, familiarity, and likeness. Participants saw the whole faces, internal features and external features in three separate blocks of trials, with the order being counterbalanced across participants.

METHOD

Participants. Thirty six undergraduate students from the University of York took part in this experiment (8 male, 28 female), receiving course credit or a payment of £2 for their participation. Participants were 18 – 22 years old (mean = 19.58; S.D. = 0.94) and were required to have lived in the UK for at least 18 years.

Design. The main experimental stimuli were 68 famous faces differing on AoA (early or late acquired) and gender (male or female). They were presented to each participant as a whole face, as internal features only, and as external features only. The task was familiarity decision, so the famous faces were presented with an equal number of unfamiliar faces which were half male and half female. None of the male faces had facial hair. The dependent variables were response latency and accuracy.

Materials. Sixty-eight coloured images of faces were selected from a database comprising current images of celebrity faces. The images were obtained by scanning photographs from magazines, or obtained as digital stills from the internet or courtesy of PAPhotos picture library. All the images were full-face and were selected to ensure minimal variation in pose

and expression. The images were digitally edited to obscure background and clothing, and were cropped below the neck. The images were pasted onto a white background and equated in size to approximately 350 x 350 pixels, determined by approximate inter-ocular distance.

Faces in the internal condition were created by deleting hair and contours from the whole face images, so that only the internal features (eyes, nose and mouth) remained. All internal images were equated in size and were captured in an oval template subtending 180 x 210 pixels. External images were the complement of internal images, comprising hair, ears and face-shape. Examples of the stimuli seen in each of the three conditions are shown in Fig. 1.

Fig. 1. Illustration of the types of face stimulus employed in Experiments 1 and 2. Left to right: whole face, internal features, external features. Note that the images used in the experiments were in colour.

The whole face images were rated for AoA, distinctiveness, frequency, familiarity, and likeness by York University undergraduates, who had been brought up in the UK for at least 18 years. Approximately 20 participants rated each face. One group of participants provided the AoA and likeness ratings, a second group provided the frequency and

distinctiveness ratings, and a third group provided the familiarity ratings. Details of the ratings are as follows:

AoA. Participants were required to estimate the age at which they first became aware of a celebrity, using a 5-point scale (1 = below 5 years of age; 2 = 5-8 years; 3 = 9-12 years; 4 = 13-16 years; 5 = over 16 years).

Distinctiveness. Participants were asked to rate faces according to how typical or distinctive they appeared, using a 5-point scale (1 = very typical, 2 = fairly typical, 3 = mildly distinctive, 4 = fairly distinctive, 5 = highly distinctive). Examples of distinctive and typical faces were provided with the instructions in order to facilitate an awareness of a range of face types.

Frequency. Ratings were made on a 5-point scale pertaining to how often the celebrity was encountered at the time of testing (1 = once a year or less, 2 = a few times a year, 3 = about once a month, 4 = about once a week, 5 = almost once a day).

Familiarity. Participants were asked to rate each face with respect to how familiar it was to them. Ratings were made on a 5-point scale (1 = unfamiliar; 2 = vaguely familiar; 3 = fairly familiar; 4 = familiar; 5 = very familiar).

Likeness. A rating of how well the image resembled the person concerned permitted a quality control of the images themselves. Ratings were made on a 5-point scale (1 = very poor, 2 = poor, 3 = reasonable, 4 = good, 5 = very good).

The 68 experimental stimuli comprised 34 early acquired faces and 34 late acquired faces. Early faces were rated as learned between the ages of 5 and about 10 years of age; late faces as learned after the age of 10. At each level of AoA, there were an equal number of male and female faces. Only faces rated ‘familiar’ or ‘very familiar’ were used. The early and late sets were matched on distinctiveness, frequency, familiarity and likeness (all t’s < 1). Table 1 shows the summary data for the different experimental sets of famous faces.

The famous faces were matched to an equal number of unfamiliar faces on gender, approximate age and race. These images were prepared in the same manner as the famous faces.

Apparatus and Procedure. The familiarity decision task was executed in PsyScope (Cohen, MacWhinney, Flatt & Provost, 1993) for the AppleMac. Participants were instructed to “decide as quickly and accurately as possible, whether or not the presented face is familiar” in response to the

images presented in the centre of the screen. Each experimental trial began with a 500 ms blank screen prior to image onset. Faces were displayed until a response was made. 'Familiar' and 'unfamiliar' responses were made using the 'z' and '/' keys on the keyboard, pressed with index fingers. The labelling of keys was counter-balanced across participants. The images were displayed on a screen with a resolution of 1,280 x 1,024 pixels and the absolute dimensions of the whole face images were approximately 70 x 150 mm. Participants viewed these from a distance of approximately 30 cm.

Table 1. Mean ratings for the early and late acquired male and female faces used in Experiments 1 and 2.

		AoA	Distinctive-ness	Frequency	Familiarity	Likeness
Early male	Mean	2.85	2.92	2.53	4.40	3.97
	SD	0.41	0.92	0.46	0.41	0.34
	Range	2.14-3.41	1.65-4.80	2.00-3.65	3.68-4.95	3.00-4.62
Early female	Mean	2.76	2.96	2.72	4.39	4.03
	SD	0.48	0.57	0.75	0.41	0.39
	Range	1.74-3.43	1.80-3.80	1.70-3.95	3.73-4.95	3.30-4.62
Late male	Mean	4.06	2.97	2.62	4.42	4.00
	SD	0.49	0.74	0.52	0.24	0.37
	Range	3.57-5.00	1.60-4.13	1.85-3.70	3.95-4.82	3.85-4.71
Late female	Mean	4.07	2.91	2.68	4.40	4.04
	SD	0.47	0.59	0.53	0.31	0.32
	Range	3.48-5.00	2.05-4.35	1.43-3.40	3.95-4.82	3.25-4.52

The 136 experimental faces (68 famous and 68 unfamiliar) were presented once as whole faces, once as internal features, and once as external features. The face type conditions were blocked, and the order of the blocks was counterbalanced yielding six orders of presentation. Each block was preceded by instructions informing participants about the task requirements and the nature of the stimuli.

Forty practice trials (20 famous and 20 unfamiliar) were administered prior to each of the three blocks. After the practice session, participants were reminded to respond "as quickly and accurately as possible". The experimental trials were directly preceded by 8 lead-in trials.

After the experiment, participants performed a familiarity verification exercise. Participants were presented with the intact versions of the famous faces and asked to indicate "which faces are familiar to you (i.e. you recognize them as being famous)". They were explicitly informed that the task was not timed and accuracy was important. This allowed faces that were not known to particular participants to be removed from the analyses.

RESULTS

In order to reduce overall error rates, data was only included for participants who showed in the post-experiment, familiarity verification session that they recognised at least 75% of the celebrities, made at least 75% correct responses to famous faces in each of the three conditions (whole, internals and externals), and made at least 75% correct responses across the famous and unfamiliar faces combined. Seven participants failed to meet these criteria and were replaced.

Responses to faces which participants did not subsequently recognise in the familiarity verification session (4.5%) were removed from the reaction time analysis, along with other incorrect responses (whole faces 6.5%, internal features 24.8%, external features 30.2%). Reaction times less than 300 ms or longer than the overall mean + 3SDs were also removed from the analysis (whole faces 1.8%, internal features 1.4%, external features 1.4%). Table 2 shows the mean correct response latencies and error rates in each condition of the experiment. RTs and error rates to famous and unfamiliar faces were analysed separately.

FAMOUS FACES

Reaction times. Correct RTs to famous faces were analysed by ANOVA with face type (whole, internal or external), face gender (male or female) and AoA (early or late) as within-subjects factors. The main effect of face type was significant, $F(2,70) = 41.84$, MSe = 51191.95, $p < .001$. Post-hoc Tukey tests ($\alpha = .05$) found that whole faces (mean = 705 ms) were recognised significantly faster than internal features (mean = 872 ms) and external features (mean = 942 ms), but the difference between internal and external features was not significant (HSD = 128). The main effect of face gender was also significant, $F(1,35) = 16.55$, MSe = 9302.90, $p < .001$,

with faster overall responses to male faces (821 ms) than to female faces (858 ms).

Table 2. Mean reaction times and SDs in milliseconds and percent errors in each condition of Experiment 1 (familiarity decision).

		Early famous		Late famous		Unfamiliar	
		Male	Female	Male	Female	Male	Female
Whole faces	Mean RT	672	731	702	714	702	737
	SD	101	140	99	143	123	127
	% error	10.5	14.9	15.7	10.1	6.2	6.7
Internal features	Mean RT	829	885	869	906	856	896
	SD	146	164	165	221	171	194
	% error	29.2	30.7	33.2	29.6	8.9	10.3
External features	Mean RT	903	960	949	955	896	918
	SD	186	237	230	207	182	190
	% error	25.3	47.1	37.7	34.5	12.4	13.8

The main effect of AoA was significant, $F(1,35) = 8.48$, MSe = 4681.80, $p < .01$, with faster responses to early acquired faces (830 ms) than to late acquired faces (849 ms). The two-way interaction between AoA and face gender was significant, $F(1,35) = 9.47$, MSe = 4222.89, $p < .01$. Post-hoc comparisons using Tukey's 'honestly significant difference' test (HSD = 41) found that the overall effect of AoA approached significance for male faces (mean difference = 39 ms), but not for female faces (mean difference = -0.3 ms). No other interactions were significant.

Error rates. The overall error data were analysed using ANOVA. The main effect of face type was significant, $F(2,70) = 42.59$, MSe = 14.59, $p < .001$. Tukey tests ($\alpha = .05$; HSD = 2.2) found that whole faces (mean n = 2.2, 12.8%) were recognised significantly more accurately than either internal features (mean n = 5.2, 30.7%) or external features (mean n = 6.2, 36.2%), but the difference between internal and external features did not reach significance. The main effects of gender and AoA were not significant.

The interaction between AoA and face gender was again significant, $F(1,35) = 49.72$, MSe = 2.79, $p < .001$, as was the interaction between face type and face gender, $F(2,70) = 10.25$, MSe = 3.42, $p < .001$. These interactions were subsumed in a significant three-way interaction between face type, AoA and face gender, $F(2,70) = 13.42$, MSe = 2.09, $p < .001$. As a means of exploring this top-level interaction, error rates from each face type were submitted to independent ANOVAs with AoA and face gender as factors. In the whole face condition, the interaction between AoA and face gender was significant, $F(1,35) = 15.10$, MSe = 1.71, $p < .001$. Tukey tests ($\alpha = .05$; HSD = 0.8) revealed that error rates to whole, early male faces were lower than to whole, late male faces while the accuracy of responding to whole female faces was not affected by AoA. In the analysis of internal features, the AoA effect and its interaction with gender failed to reach significance. In the analysis of external features, the AoA by gender interaction was again significant, $F(1,35) = 61.72$, MSe = 2.63, $p < .001$. Tukey tests ($\alpha = .05$; HSD = 1.0) showed effects of AoA for both male and female external features, but the effect was in opposite directions: fewer errors were made to early than late external male faces while more errors were made to early than late external female faces.

UNFAMILIAR FACES

Reaction times. False positive responses to unfamiliar faces and RTs less than 300 ms or longer that the overall mean + 3SDs were removed from the analysis (whole faces 6.5%, internal features 9.7%, external features 13.2%). Correct RTs to unfamiliar faces were analysed with factors of face type (whole, internal, external) and face gender (male, female) as within-subjects factors. The analysis of RTs found a significant main effect of face type, $F(2,70) = 32.55$, MSe = 22389.07, $p < .001$. Tukey tests ($\alpha = .05$; HSD = 84) found significant differences between whole faces (mean = 719 ms) and both internal features (mean = 876 ms) and external features (mean = 907 ms), but the difference between internal and external features did not reach significance.

The main effect of gender was significant, $F(1,35) = 13.89$, MSe = 4020.08, $p = .001$, with male faces (818 ms) classified faster than female faces (850 ms). The interaction between face type and gender was not significant.

Error rates. Analysis of the numbers of errors to unfamiliar faces in each condition found a main effect of face type, $F(2,70) = 12.07$, MSe = 7.65, $p < .001$. Tukey tests ($\alpha = .05$; HSD = 1.6) found that whole faces (mean n = 2.2, 6.5%) were identified as unfamiliar significantly more accurately than external features (mean n = 4.5, 13.1%). Accuracy to unfamiliar internal features (mean n = 3.3, 9.6%) was not significantly different from either whole faces or external features. The main effect of gender and the interaction between face type and gender were not significant.

Summary. Familiarity decisions were faster and more accurate to whole faces than to internal or external features, which did not differ significantly. This was true for both famous and unfamiliar faces. Overall, male faces were recognised more rapidly than female faces. The factors of face gender and AoA combined in an unexpected fashion: the effect of AoA on RTs was only present for male faces. Famous female faces showed no influence of AoA on recognition speed. AoA also affected the accuracy of responding to male whole faces and male external features.

EXPERIMENT 2

Experiment 2 employed the same stimuli and the same design and procedural details as Experiment 1. The difference lay only in the task, which was changed from familiarity decision to gender decision (*Is the face male or female?*).

METHOD

Participants. Thirty six undergraduate students from the University of York took part in this experiment (8 male, 28 female), receiving course credit or a payment of £2 for their participation. Participants were 18 – 21 years old (mean = 19.67; S.D. = 1.01) and were required to have lived in the UK for at least 18 years.

Design, materials and procedure. The stimuli for Experiment 2 were the same as for Experiment 1. The design and procedure were the same except that participants were instructed to "decide as quickly and accurately as possible, whether the presented face is male or female".

RESULTS

All participants met the criteria which required them to: a) produce 75% correct responses to famous faces in the whole face condition and familiarity verification session; b) produce 75% correct responses to famous faces and unfamiliar faces pooled together, in the whole face condition and familiarity verification session; c) produce 75% correct responses to unfamiliar faces in the internal and external conditions.

The error analysis was based on a global measure which calculated errors arising for the following reasons: a) incorrect gender classifications made to known faces (a face is 'known' if it is successfully recognized in the familiarity verification session); b) a failure to recognise a face in the familiarity verification session; c) reaction times deemed outliers. As a result of incorrect gender classifications made to known faces, 74 (3.0%) whole face responses, 94 (3.8%) internal feature responses, and 96 (3.9%) external feature responses were removed. In addition, 121 (4.9%) responses from each face type condition were removed as a consequence of faces not being recognized in the familiarity verification session. As a result of eliminating outliers faster than 300ms and slower than 3 standard deviations away from the mean, 67 (2.7%) responses were removed from the whole face condition, 41 (1.7%) from the internal condition and 50 (2.0%) from the external condition. Outliers were calculated for each face type separately.

Only correct gender classifications, and those made to faces successfully recognized in the familiarity verification session, were submitted to the analysis of RTs. Table 3 shows the mean correct response latencies in each condition of the experiment and the associated error rates. The logic of the analysis was first to discover whether gender decisions showed an influence of familiarity, and whether that influence was different for whole faces, internal features and external features. The initial analysis compared RTs and error rates to famous and unfamiliar male and female faces presented as wholes, internals or externals. Early and late famous faces were combined in this analysis. Responses to the famous faces were then analysed with AoA, face type and face gender as factors, looking to see if there were differential effects of AoA on the internal features compared with whole faces and internal features.

FAMOUS vs UNFAMILIAR FACES

Reaction times. Gender classification RTs (correct responses only) were analysed with face type (whole, internal, external), familiarity (famous, unfamiliar) and face gender (male, female) as within-subjects factors. AoA cannot be used in this analysis because AoA applies only to the famous faces.

The main effect of face type was significant, $F(2,70) = 36.20$, MSe = 7178.12, $p < .001$. Tukey tests ($\alpha = .05$; HSD = 48) found that whole faces (mean = 509 ms) were classified significantly faster than external features (mean = 577 ms) or internal features (mean = 588 ms), but the difference between internal and external features was not significant. The main effect of familiarity was also significant, $F(1,35) = 24.87$, MSe = 956.01, $p < .001$, with famous faces (551 ms) being classified for gender faster than unfamiliar faces (566 ms). The main effect of face gender was not significant.

Table 3. Mean reaction times and SDs in milliseconds and percent errors in each condition of Experiment 2 (gender decision).

		Early famous		Late famous		Unfamiliar	
		Male	Female	Male	Female	Male	Female
Whole faces	Mean RT	517	499	518	502	502	518
	SD	78	76	70	74	67	66
	% error	13.1	8.8	13.9	7.0	7.8	13.2
Internal features	Mean RT	578	582	600	557	591	602
	SD	78	88	88	85	94	105
	% error	11.4	9.5	14.7	6.2	7.1	16.8
External features	Mean RT	565	558	574	559	575	606
	SD	64	87	80	72	78	93
	% error	13.9	8.2	14.9	6.7	7.1	20.5

Familiarity entered into significant two-way interactions with face type, $F(2,70) = 8.07$, MSe = 759.75, $p = .001$, and face gender, $F(1,35) = 38.89$, MSe = 873.62, $p < .001$. No other interactions were significant. A simple main effects analysis of the interaction between familiarity and face type showed that RTs to whole faces were not affected overall by familiarity, but famous faces were classified faster than unfamiliar faces in both the internal features condition, $F(1,35) = 11.05$, MSe = 486.07, $p < .01$, and the external features condition, $F(1,35) = 24.42$, MSe = 513.24, $p < .001$.

Analysis of the interaction between familiarity and face gender using Tukey tests ($\alpha = .05$; HSD = 19) showed that across the three types of face stimulus, gender classification responses to female faces were significantly (33 ms) faster to famous than to unfamiliar faces whereas classification of male faces was not affected by familiarity, showing a non-significant trend in the opposite direction.

Error rates. The main effect of face type on gender classification errors was significant, $F(2,70) = 4.89$, MSe = 2.77, $p < .05$, with increasing numbers of errors to whole faces, internal features and external features though none of the pairwise comparisons emerged as significant on Tukey tests ($\alpha = .05$; HSD = 0.9). The main effect of face gender was significant, $F(1,35) = 4.77$, MSe = 8.33, $p < .05$, with fewer errors to male faces than to female faces. The main effect of familiarity was not significant.

Each of the three possible two-way interactions was significant (face type x familiarity, $F(2,70) = 5.64$, MSe = 1.83, $p < .01$; face type x gender, $F(2,70) = 4.09$, MSe = 2.92, $p < .05$; familiarity x gender, $F(1,35) = 53.09$, MSe = 13.92, $p < .001$). Those interactions were subsumed into a significant three-way interaction involving face type, familiarity and gender, $F(2,70) = 18.57$, MSe = 1.25, $p < .05$. Data from each face type were analysed independently in separate within-subject ANOVAs to explore the effect of familiarity on gender categorisation accuracy. The interaction between familiarity and gender was significant in the whole face condition, $F(1,35) = 20.04$, MSe = 6.22, $p < .001$, the internal features condition, $F(1,35) = 41.51$, MSe = 5.54, $p < .001$, and the external features condition, $F_1(1, 35) = 92.63$, MSe = 4.65, $p < .001$. In each case, Tukey tests indicated that famous male faces attracted more gender classification errors than unfamiliar male faces, whereas famous female faces were classified more accurately than unfamiliar female faces ($\alpha = .05$; whole faces, HSD = 1.6; internal features, HSD = 1.5; external features, HSD = 1.4).

FAMOUS FACES

Reaction times. Gender classification RTs to famous faces were analysed separately with face type (whole, internal or external), AoA (early or late) and face gender (male or female) as within-subjects factors. The main effect of face type was significant, $F(2,70) = 30.61$, MSe = 6422.60, $p < .001$. Tukey tests ($\alpha = .05$; HSD = 45) found that whole faces (mean = 509 ms) were classified significantly faster than external features (mean = 564 ms) or internal features (mean = 579 ms), but the difference between internal and external features was again not significant. The main effect of gender was also significant, $F(1, 35) = 11.23$, MSe = 2,515.16, $p < .01$, with faster gender classification of famous female faces (543 ms) than famous male faces (559 ms).

The main effect of AoA was not significant, but there was a significant interaction between AoA and face gender, $F(1,35) = 6.41$, MSe = 1251.20, $p < .05$, and a significant three-way interaction between face type, AoA and gender, $F(2,70) = 4.89$, MSe = 1228.03, $p < .05$. Data from each face type condition were analysed separately. The whole face and external features conditions showed no effect of AoA on gender classification speed, and the interaction between AoA and face gender was not significant for those stimulus types. In the internal features condition there was a significant interaction between AoA and face gender, $F(1,35) = 18.89$, MSe = 1031.29, $p < .001$. Tukey tests ($\alpha = .05$; HSD = 20) found significantly faster gender decisions to early than late male internal features but significantly slower decisions to early than late female internal features.

As with familiarity decisions, faster responses to early than to late acquired faces seemed to be true for male but not female faces. Gender classification RTs to famous male faces only were analysed separately with face type and AoA as factors. The main effect of face type was significant, $F(2,70)$ 23.42, MSe = 4188.31, $p < .001$. Tukey tests ($\alpha = .05$; HSD = 36.5) found that whole faces were classified faster than external features or internal features, but the difference between internal and external features was not significant. The main effect of AoA was significant, $F(1, 35) = 6.36$, MSe = 958.77, $p < .05$, with faster gender decision RTs to early male faces (overall mean = 553 ms) than to late male faces (overall mean = 564 ms). The interaction between AoA and face type was not significant, though the trend was for the difference in favour of early acquired faces to be largest for internal features, smaller for external features, and smallest for whole faces.

Error rates. There was no significant effect on accuracy of face type or AoA. The main effect of face gender was significant, $F(1,35) = 17.34$, MSe = 6.29, $p < .001$, with female faces (mean n = 1.3, 3.9 %) attracting fewer errors overall than male faces (mean n = 2.3, 6.8 %). The two-way interaction between AoA and gender was significant, $F(1,35) = 4.05$, MSe = 2.88, $p = .05$. Tukey tests ($\alpha = .05$; HSD = 1.1) showed that late acquired females were classified for gender more accurately than late acquired males but there was no significant difference between classification accuracy for early acquired male and female faces. No other interactions were significant.

Summary. Gender decisions were faster and more accurate to whole faces than to internal or external features, which did not differ. Gender classification RTs to whole faces showed no effect of familiarity, but the internal and external features of famous faces were classified faster than the internal and external features of unfamiliar faces. Female faces showed no consistent effect of AoA on gender classification, but an effect was present for RTs to the male faces, where responses were significantly faster to early than to late acquired faces.

DISCUSSION

Whole male faces were responded to more quickly and more accurately than either internal or external features. This was true for both famous and unfamiliar faces, and was seen in both the familiarity and gender decision tasks. H. Ellis et al. (1979) found that the recognition of famous faces was most accurate from whole face images, less accurate from internal features, and least accurate from external features. Our accuracy and RT results confirm the advantage for whole faces, and the trends in our data were in the direction of better familiarity decision performance on internal features than external features, but the differences between the latter two conditions failed to reach significance. Our results imply that there may not be any great difference between internal and external features in the quality of the cues they provide for distinguishing familiar from unfamiliar faces. Gender decisions were also faster and more accurate for whole faces than for internal or external features. Our results therefore replicate the finding of Stevenage and Osborne (2006) that whole faces provide better cues for gender classification than internal features only.

An unexpected feature of the results of Experiments 1 and 2 was the difference in the responses to male and female faces[1,2]. Familiarity

[1] Our experiments involved separate blocks of trials in which participants responded to whole faces, internal features, or external features. Order of these blocks was counterbalanced across participants. A reviewer suggested that we should include Order as a factor to adjust the degrees of freedom and eliminate possible variation due to different presentation orders. We ran those analyses on the RTs to famous faces (where most of the interest lies) and found that the results were essentially unchanged: a) In Experiment 1 (familiarity decision), the main effect of Order on RTs to famous faces was not significant. There was a significant interaction of Order with Face type, caused by the fact that RTs to whole faces were unaffected by whether those stimuli occurred in the first, second or third block, while RTs to internal and external features improved from first to second to third blocks, with particular speeding of responses to part-faces after participants had been exposed to a block of whole faces. This is presumably as a result of repetition priming from whole to part faces. The interactions between Order, Face gender and AoA were not significant, and the pattern of results for those variables was unchanged. b) The main effect of Order was significant in the analysis of errors to famous faces in Experiment 1, and the interaction of Order with Face type was again significant. Error rates declined from first to second presentations, particularly for internal features which were identified correctly following a block of whole faces than following a block of external features. Otherwise, the pattern of results we reported previously was unchanged. c) In Experiment 2 (gender decision), the main effect of Order on RTs to famous faces was not significant. The Order x Face type interaction was significant: as with familiarity decision, gender decision RTs to whole faces were largely unaffected by order while RTs to internal features decreased from first to second presentation, particularly if the first presentation involved whole faces. Order did not interact with any of the other factors, and the pattern of results reported previously was unchanged. d) The main effect of Order on errors in Experiment 2 was also not significant, and there were no interactions involving Order. The pattern of results reported previously was unchanged. In summary, the order in which the different types of face stimuli were presented had an effect on RTs to the different types of face stimuli, with RTs to internal and external features being particularly prone to priming from whole faces. But including Order as a factor did not change the effects of the gender of the stimulus face, or its age of acquisition.

[2] Both of our experiments had 8 male and 28 female participants. This reflects the proportions of male and female students of Psychology in England. A reviewer asked if participant gender had an effect on the results. Between-groups analyses with 8 participants in one group and 28 in the other are far from ideal, but we ran those analyses to see what happened: a) In Experiment 1 (familiarity decision), the main effect of Participant gender on RTs was not significant. There was a significant interaction between Participant gender and Face gender, caused by the fact that the female participants tended to recognise the famous female faces faster than the famous male faces, while the male participants showed the opposite tendency. Participant gender did not interact with Face type (whole vs internal features vs external features) or with age of acquisition, and its inclusion did not affect the effects of those factors. b) The main effect of Participant gender on familiarity decision errors was also not significant. There was a significant interaction between Participant gender and Face gender, caused by the fact that the female participants made equal

decisions to both famous and unfamiliar female faces were significantly slower than decisions to male faces. This occurred despite the fact that the two sets were matched on age of acquisition, distinctiveness, frequency, familiarity and likeness (see Table 1). AoA affected the speed of responding to famous male faces, but not famous female faces, in both the familiarity decision task (Experiment 1) and the gender decision task (Experiment 2). The accuracy of responding to male faces in the familiarity decision task also showed an effect of AoA, with fewer recognition errors to early than to late acquired faces. No studies of AoA effects in face recognition have separated responses to male and female faces, so we have no way of knowing whether our observation of AoA effects for male but not female faces is aberrant or not. We will leave the question of whether AoA and other effects are genuinely different for male and female faces to future work and will concentrate this Discussion on the results obtained with male faces[3].

Richards and A. Ellis (2008) analysed familiarity and gender decisions to famous male faces presented as whole faces or internal features. The faces varied on both AoA and rated masculinity. As in the present study, AoA modulated the speed of responding in both tasks. AoA did not interact significantly with face type (whole *vs* internal features), but it did interact with the rated masculinity of the famous male faces: for both

---

numbers of errors on the male and female faces while male participants made more errors on the female than the male faces. Participant gender did not interact with face type or age of acquisition, and its inclusion did not affect the effects of those factors. c) In Experiment 2 (gender decision), the main effect of Participant gender on RTs was significant, with female participants responding more quickly than male participants. The interaction between Participant gender and Face gender was significant, with female participants classifying female faces for gender faster than male faces while male participants showed no difference between female and male stimuli. Participant gender did not interact with face type or age of acquisition, and its inclusion did not affect the effects of those factors. d) The main effect of Participant gender on gender decision errors was not significant, but the interaction between Participant gender and Face gender was again significant, with female participants tending to make more gender classification errors to male than to femal faces while male participants showed a trend in the opposite direction. Error rates were generally very low in gender classification. In summary, female and male participants tended to show opposite patterns of results to female and male faces. That pattern warrants further investigation. But the groups were of dramatically different sizes, and the inclusion of Participant gender did not change the effects of face type or age of acquisition, which is what this paper is about.

[3] Separate ANOVAs on the RTs and errors to famous male faces in Experiment 1 (familiarity decision) found significant effects of AoA on RTs ($p<.001$) and errors ($p<.001$). Separate ANOVAs on the results for famous male faces in Experiment 2 (gender decision) found a significant effect of AoA on RTs ($p<.05$) but not errors.

familiarity and gender decisions, AoA had no effect on RTs to high masculinity faces, but RTs were slower to late than to early acquired low masculinity faces. Richards and A. Ellis (2008) suggested that the fact that AoA influences both familiarity and gender decisions rather contradicts the idea embodied in the Bruce and Young (1986) model that gender classification occurs independently of the processes that judge the familiarity of a face. Similar effects of AoA and masculinity on familiarty and gender decisions argue instead in favour of models in which the processes responsible for those two forms of judgment are more closely related. Such models include the cognitive model of H. Ellis (1986) and the neurological theory of Haxby, Hoffman and Gobbini (2002) in which relatively static, invariant features of faces, including their gender, age and familiarity, are processed by a ventral route (along the fusiform gyrus) while ever-changing features such as expressions and speech movements are processed along a route involving the superior temporal gyrus. Richards and A. Ellis (2008) proposed that deciding whether a face is familiar or unfamiliar, or whether it is male or female, may both depend on the activation of semantic information. Early acquired faces may be faster than late acquired faces to access their semantic representations because those representations are more detailed (Ghyselinck et al., 2004a,b; Steyvers & Tenenbaum, 2005), or because the processing of mapping from one form of representation (a face) to another form of representation (stored knowledge about a familiar person) may be easier for early than for late items (cf. A. Ellis & Lambon Ralph, 2000). If the decision as to whether a stimulus is familiar or not is based upon the amount of semantic activation it generates in a given time (cf. Plaut, 1997), then we would expect early acquired faces to be classified as familiar faster than late acquired faces (Moore & Valentine, 1999; Richards & A. Ellis, 2008). An impact of AoA on gender decisions could be explained in a similar manner if we propose that the decision as to whether a face is male or female is not based on analyzing the visual features, but is based instead on whether the face activates semantic nodes representing the knowledge that a person is male or female. The possibility that gender decisions could be based on semantic as well as visual information has long been recognised (Bruce, 1986; Bruce et al., 1987): the idea being mooted here is that semantics may form the *normal* basis for gender decisions. Familiar faces will be associated with semantic knowledge, including the fact that they are male or female, which could be thought of as being encoded within shared semantic information units of the sort proposed by Burton, Bruce, and Johnston (1990) and Burton, Bruce, and Hancock (1990). Early acquired familiar faces should activate that specific piece of gender information faster than late acquired faces, giving

rise to the impact of AoA on gender decision speed noted in the present Experiment 2. Unfamiliar faces can also be classified for gender using this mechanisms. Unfamiliar faces will activate the face recognition units of familiar faces they resemble (face neighbours). Those face recognition units will, in turn, activate the semantic representations of the face neighbours, including the gender nodes. Unfamiliar male faces will mostly activate the face recognition units of familiar males, and hence will activate the 'male' semantic node more than the female. Unfamiliar female faces will mostly activate the face recognition units of familiar females, and hence will activate the 'female' semantic node more than the male. Some thoughts on why the impact of AoA on gender and familiarity decisions to male faces might be further modulated by masculinity can be found in Richards and Ellis (2008).

We began this study believing that familiarity and gender decisions are based on very different and unconnected processes. We expected AoA to affect familiarity decisions but not gender decisions. We found something quite different - at least for males - with AoA affecting both types of decision. The results lead us to propose that familiarity and gender decisions are much more closely related that we had originally believed, and that both may have a common basis in semantic representations.

RESUMEN

Identidad, género y el rol de la edad de adquisición en el procesamiento facial. Dos experimentos examinaron los efectos de la edad de adquisición (AoA) y el género de rostros utilizados como estímulos en tareas de decisión de familiaridad (Experimento 1) y decisión de género (Experimento 2) con el mismo set de rostros famosos y rostros desconocidos presentados completos y sólo con sus características internas o externas. En el experimento 1, las decisiones de familiaridad fueron más rápidas para los rostros completos que para las características internas o externas de los rostros. Los rostros famosos con edad de adquisición temprana fueron reconocidos más rápidamente que aquellos con edad de adquisición tardía, aunque el efecto fue sólo significativo para rostros masculinos famosos, y para rostros completos y características internas que para características externas. En el experimento 2, las decisiones de género se llevaron a cabo más rápidamente para rostros completos que para características internas y externas. La clasificación fue más rápida para rostros famosos que para rostros desconocidos cuando éstos fueron presentados con sus características internas o externas, pero no cuando fueron presentados completos. Se registraron más errores de clasificación para los rostros masculinos famosos que para los desconocidos, pero no hubo efecto de familiaridad en la precisión de las respuestas para los rostros femeninos. La AoA no mostró efecto en la clasificación de género para rostros completos o características

externas. La clasificación de características internas fue más rápida para los rostros aprendidos a temprana edad que para aquellos aprendidos a una edad tardía, pero más rápida para rostros femeninos aprendidos a una edad tardía que para aquellos aprendidos a temprana edad. Debido a que los resultados para los rostros femeninos fueron anómalos, las respuestas correspondientes a los rostros masculinos fueron analizadas por separado. Los resultados de los rostros masculinos forman el centro de la discusión.

REFERENCES

Bonin, P., Barry, C., Méot, A., & Chalard, M. (2004). The influence of age of acquisition in word reading and other tasks: A never ending story? *Journal of Memory and Language*, *50*, 456-476.

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

Bruce, V. (1986). Influences of familiarity on the processing of faces. *Perception, 15,* 387-397.

Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology, 77*, 305-327.

Bruce, V., Ellis, H., Gibling, F., & Young, A. (1987). Parallel processing of the sex and familiarity of faces. *Canadian Journal of Psychology, 41*, 510-520.

Brysbaert, M., Van Wijnendaele, I., & De Deyne, S. (2000). Age of acquisition effects in semantic tasks. *Acta Psychologica*, *104*, 215-226.

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

Burton, A. M., Bruce, V., & Johnston, R. A. (1990). Understanding face recognition with an interactive activation model. *British Journal of Psychology*, *81*, 361-380.

Burton, A. M., Bruce, V., & Hancock, P. J. B. (1999). From pixels to people: A model of familiar face recognition. *Cognitive Psychology*, *23*, 1-31.

Campbell, R., Brooks, B., De Haan, E., & Roberts, T. (1996). Dissociating face processing skills: decisions about lip-reading speech, expression, and identity. *Quarterly Journal of Experimental Psychology, 49A,* 295-314.

Clutterbuck, R., & Johnston, R. A. (2002). Exploring levels of face familiarity using an indirect face-matching measure. *Perception, 31*, 985-994.

Clutterbuck, R., & Johnston, R. A. (2004). Demonstrating the acquired familiarity of faces by using a gender-decision task. *Perception, 33*, 159-168.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: An interactive graphic system for designing and controlling experiments in the psychology laboratory using Macintosh computers. *Behaviour Research Methods, Instruments, and Computers, 25*, 257-271.

Cortese, M. J., & Khanna, M. M. (2007). Age of acquisition predicts naming and lexical-decision performance above and beyond 22 other predictor variables: An analysis of 2,342 words. *Quarterly Journal of Experimental Psychology, 60,* 1072-1082.

Ellis, A. W., & Lambon Ralph, M. A. (2000). Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: Insights from connectionist networks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1103-1123.

Ellis, A. W., Young, A. W., & Flude, B. M. (1990). Repetition priming and face processing: Priming occurs within the system that responds to the identity of a face. *Quarterly Journal of Experimental Psychology, 42A*, 495-512.

Ellis, H. D., Shepherd, J. W., & Davies, G. M. (1979). Identification of familiar and unfamiliar faces from internal and external features: some implications for theories of face recognition. *Perception, 8,* 431-439.

Ganel, T., & Goshen-Gottstein, Y. (2002). Perceptual integrity of sex and identity of faces: Further evidence for the single-route hypothesis. *Journal of Experimental Psychology, 28*, 854-867.

Ghyselinck, M., Custers, R., & Brysbaert, M. (2004). The effect of age of acquisition in visual word procession: Further evidence for the semantic hypothesis. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *30*, 550-554.

Ghyselinck, M., Lewis, M. B., & Brysbaert, M. (2004). Age of acquisition and the cumulative-frequency hypothesis: A review of the literature and a new multi-task investigation. *Acta Psychologica*, *115*, 43-67.

Goshen-Gottstein, Y., & Ganel, T. (2000). Repetition priming for familiar and unfamiliar faces in a sex-judgement task: Evidence for a common route for the processing of sex and identity. *Journal of Experimental Psychology: Learning, Memory, & Cogntion, 26*, 1198-1214. ß

Haxby, J. V., Hoffman, E. A., Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4*, 223-233.

Holmes, S. J., & Ellis, A. W. (2006). Age of acquisition and typicality effects in three object processing tasks. *Visual Cognition, 13,* 884-910.

Johnston, R. A., & Barry, C. (2006). Age of acquisition and lexical processing. *Visual Cognition, 13,* 789-845.

Juhasz, B. J. (2005). Age of acquisition effects in word and picture identification. *Psychological Bulletin, 131,* 684-712.

Lewis, M. B. (1999). Age of acquisition in face categorisation: is there an instance-based account? *Cognition*, *71*, B23-B39.

Moore, V., & Valentine, T. (1998). The effect of age of acquisition on speed and accuracy of naming famous faces. *Quarterly Journal of Experimental Psychology*, *51A*, 485-513.

Moore, V., & Valentine, T. (1999). The effects of age of acquisition in processing famous faces and names: Exploring the locus and proposing a mechanism. In *Proceedings of the twenty-first annual meeting of the Cognitive Science Society*, *Vancouver, 1999* (pp. 416-421). Mahwah, NJ: Erlbaum.

Moore, V., Smith-Spark, J. H., & Valentine, T. (2004). The effects of age of acquisition in object recognition. *European Journal of Cognitive Psychology*, *16*, 417-439.

Plaut, D. C. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language and Cognitive Processes*, *12*, 765-805.

Richards, R. M. & Ellis, A. W. (2008). Mechanisms of identity and gender decisions to faces: Who rocked in 1986? *Perception, 37,* 1700-1719.

Roberts, T., & Bruce, V. (1988). Feature saliency in judging the sex and familiarity of faces. *Perception*, *17*, 475-481.

Rossion, B. (2002). Is sex categorization from faces really parallel to face recognition? *Visual Cognition, 9*, 1003-1020.

Stevenage, S. V., & Osborne, C. D. (2006). Making heads turn: The effect of familiarity and stimulus rotation on a gender-classification task. *Perception, 35*, 1485-1494.

Steyvers, M., & Tenenbaum, J. B. (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognitive Science*, *29*, 41-78.

Van Loon-Vervoorn, W. A., Van Ham, P., & Van der Koppen, M. (1988). The importance of age of acquisition for imageability in word processing. In M. Dennis, J. Engelkamp, & J. T. E. Richardson (Eds), *Cognitive and neuropsychological approaches to mental imagery* (pp. 99-107). Dordrecht: Martinus Nijhoff.

Young, A. W., Hay, D. C., McWeeny, K. H., Flude, B. M., & Ellis, A. W. (1985). Matching familiar and unfamiliar faces on internal and external features. *Perception*, *14*, 737-746.

Young, A. W., McWeeny, K. H., Hay, D. C., & Ellis, A. W. (1986). Matching familiar and unfamiliar faces on identity and expression. Psychological Research, 48, 63-68.

(Manuscript received: 20 November 2007; accepted: 20 May 2008)