THE EFFECT OF AGE OF ACQUISITION ON SPEED AND ACCURACY OF NAMING FAMOUS FACES.

Viv Moore

University of Durham, UK

and

Tim Valentine

Goldsmiths College, University of London, UK

Short Title: Age of acquisition effects on face naming.

Address for correspondence:

Viv Moore,
Department of Psychology,
University of Durham,
Science Laboratories,
South Road,
Durham.
DH1 3LE

We would like to acknowledge the support from the Economic and Social Research Council in funding this research (Grant No. R000 234612).

Abstract

Three experiments examined whether famous faces would be affected by the age at which knowledge of the face was first acquired (AoA). Using multiple regression design, Experiment 1 showed that rated familiarity and AoA were significant predictors of the time required to name pictures of celebrities' faces and the accuracy of producing their names. Experiment 2 replicated an effect of AoA using a factorial design in which other attributes of the celebrities were matched. In both Experiments 1 and 2 several ratings had been collected from participants before naming latency data were collected. Experiment 3 investigated the accuracy and latency of naming celebrities without any prior exposure to the stimuli. An advantage for naming early acquired celebrities was observed even on the first presentation. The participants named the same celebrities in three subsequent presentations of the stimuli. The effect of AoA was not significant on the fourth presentation. The implications of these results for models of face naming and directions for future research are discussed.

Acknowledgements

We would like to thank Andy Ellis, Jonathan Grainger, Andy Young and anonymous reviewers for their most useful and detailed comments on earlier versions of this paper.

Introduction

Oldfield and Wingfield (1964) reported that the latency to name pictures of objects is a function of the frequency of the object's name in word frequency corpora (e.g. Thorndike & Lorge, 1944). Thus pictures of objects with high frequency names (e.g. house) will be named more quickly than objects with low frequency names (e.g. aardvark). Later Oldfield and Wingfield (1965) reported that the effect of word frequency was independent of word length. Humphreys, Riddoch and Quinlan (1988) have also found that word frequency affects object naming latency. Word frequency effects have been reported in many studies over a wide range of tasks. For example, in lexical decision tasks participants are faster to judge high frequency words are English words than they are to judge low frequency words are English words (e.g. Scarborough, Cortese & Scarborough, 1977) and high frequency words are read aloud more quickly than low frequency words (e.g. Monsell, Doyle & Haggard, 1989; Seidenberg, Waters & Barnes, 1984).

The emphasis on word frequency has served to validate frequency effects *per se* leaving other potentially powerful factors ignored. However, the effect of age of acquisition (AoA) has recently re-emerged as an important determinant of processing speed in picture naming and other tasks. It has been proposed that the speed of object naming is a function of the age at which the object name was acquired. The earlier a name is acquired, the quicker it can be produced on confrontation with a picture of the relevant object (e.g. Carroll & White, 1973a). Oldfield and Wingfield derived their data from only twelve participants and twenty-six stimuli, whereas Carroll and White (1973a) performed a larger study (37 participants and 94 stimuli), and used various measures of written word frequency, AoA and word length as independent variables in multiple regression analyses. These data were subsequently re-analysed using additional estimates of AoA and word frequency giving the same result (Carroll & White, 1973b). Carroll and White concluded that the age at which object names were acquired was the chief determinant of naming latency, and that some measures of word frequency only predict naming latency to the extent that they reflect AoA.

Morrison, Ellis and Quinlan (1992) re-analysed Oldfield and Wingfield's (1965) data including ratings of AoA in multiple regression analyses. This analysis revealed that AoA was the major determinant of naming speed and that word frequency played no independent role when its correlation with AoA was taken into account. Morrison *et al*. replicated this result in an experimental study of object naming and showed that there was no effect of AoA on the time taken to make semantic decisions to objects (natural vs. man-made). They concluded that the locus of AoA is at the stage of the output lexicon for speech. This conclusion is consistent with the locus of AoA proposed by Brown and Watson (1987), who suggested that the effect of AoA reflects a developmental stage in language acquisition. They proposed that early acquired words are represented in a more *complete form* in memory than later acquired words, and therefore are produced more rapidly throughout later life.

The effects of AoA are not restricted to object naming tasks. Morrison and Ellis (1995) reported that AoA but not written word frequency affected the speed of word naming and that both word frequency and AoA exerted independent effects in a lexical decision task. Their results are compatible with several other studies (e.g. Carroll & White, 1973a; Rubin, 1980; Gilhooly & Logie, 1981; Gilhooly & Logie, 1982). Turner, Valentine and Ellis (in press) also found effects of AoA and of word frequency on visual lexical decision but found an effect of only AoA on auditory lexical decision. Taken together, these results suggest that some re-appraisal of the role of word frequency on object and word naming latency may be necessary.

The studies cited above suggest that word frequency only affected picture naming to the extent that it is correlated with AoA; if the effect of AoA was taken account there was no independent effect of word frequency. However, some studies have found independent effects of word frequency and AoA on picture naming latency. Lachman (1973) reported significant effects of subjective ratings of word frequency, word codability (or name agreement) and of AoA. These findings were supported in a larger picture naming study on children aged from 4 years to 10 years old and adults (Butterfield & Butterfield, 1977). Lachman, Shaffer and Hennrikus (1974) report significant independent effects of both AoA and subjective ratings of spoken word frequency in their picture naming study.

In a recent study of picture naming based on the Snodgrass and Vanderwart (1980) pictures, Barry, Morrison and Ellis (1997) included a measure of spoken word frequency from the Celex database (Baayen, Piepenbrock, & van Rijn, 1993), the number of phonemes in the object name, subjective ratings of image agreement, visual complexity, familiarity, name agreement (with the concepts depicted in the pictures), AoA and imageability (of the written object name) and a multiplicative term between AoA and spoken word frequency in a multiple regression analysis. The results showed that spoken word frequency, the percentage of name agreement and the multiplicative term between spoken word frequency and AoA significantly predicted naming speed. The interaction between AoA and spoken word frequency was such that there was no difference in naming speed between early-acquired and late-acquired pictures with high frequency names but that pictures with early-acquired, low frequency names were named faster than were pictures with late-acquired, low frequency names.

Many models of human cognition, (e.g. connectionist models) have been designed to account for the effects of word frequency. However few, if any, can account for the effects of AoA. For example, Seidenberg and McClelland's (1989) model, which uses backward error-propagation to learn mappings between orthography and phonology, has been shown to account for a wide range of effects in normal reading. The model simulates effects of word frequency on word naming latency. However, Morrison and Ellis (1995) found that there is no effect of word frequency on word naming latency when the effect of AoA is controlled. In contradiction with this result, Gerhand and Barry (in press) report that both word frequency and AoA affect word naming latency. While the apparent effects of word frequency in these studies are somewhat contradictory, there is an unambiguous effect of AoA. It is not at all clear how Seidenberg and McClelland's model could account for an effect of AoA. Similarly, Humphreys, Lamote and

Lloyd-Jones (1995) have simulated the effect of word frequency on object naming using an interactive activation and competition architecture. It is not at all clear how an interactive activation and competition architecture could simulate the effect of AoA on object naming latency. However, it is clear that current models of human cognition face an important challenge to simulate the effects of AoA appropriately. Therefore, it can be readily appreciated how vital it is to obtain a clear understanding of the empirical effects of frequency of use and the age at which one acquires information.

The results of picture naming studies are somewhat contradictory. Some studies have found an independent effect of written word frequency, but have often failed to control for the age when a word was acquired. Other studies have found an independent effect of AoA or an interaction between spoken word frequency and AoA. These studies have used various measures of spoken and written word frequency. Most studies have analysed data using multiple regression. When evaluating the effects of these variables careful consideration needs to be given to the different measures used and choice of statistical analysis, especially in regard to the high inter-correlation between AoA and word frequency.

Age of acquisition is usually measured using subjective ratings, although objective data has validated the subjective ratings (e.g. Carroll & White, 1973a; 1973b; Morrison, Chappell & Ellis, 1997). The correlations between rated AoA and word frequency are generally high (e.g. in the range of -.59 to -.83, Carroll & White, 1973a; 1973b). AoA could be dependant upon factors such as the socio-economic background or age of a participant, for example a word like 'video' may be late acquired to many elderly people, whereas it may be a very early-acquired word for many young children). Measures of word frequency are often based on quite dated written sources that might have some significant discrepancies with modern data. Some measures are based on North American samples. However, Barry *et al.* (1997) report high correlations between two major written frequency corpora and frequencies of written and spoken language from the Celex database (a modern British sample). The Celex database correlates with Kucera and Francis (1967) for written words (r = .78) and for spoken words (r = .75) and with Hofland and Johansson (1982) for written words (r = .81) and for spoken words (r = .76).

A further problem lies in the use of multiple regression. Although this is a powerful tool, the high inter-correlations between variables can pose problems of interpretation, especially for claims of 'independent' contribution (Morris, 1981). Lorch and Myers (1990) express serious concerns over the "inappropriate use of multiple regression in cognitive psychology" and describe more appropriate methods of data treatment. They suggest that a major difficulty with analysing repeated measures involving non-orthogonal factors lies in estimating the appropriate error term for each test of interest. The error term should be estimated by separately regressing the observations from each participant on the predictor variables. The resulting equation then represents the "best description (i.e. least squares criterion)" for each participant between the dependent variable and the set of independent (or predictor) variables. The resulting regression coefficients are then employed to create a participant by predictor matrix. Each regression coefficient is then analysed with a single-group t test, to test if it differs reliably from zero.

In view of the potential theoretical importance of AoA and the practical problems posed by the high inter-correlation between AoA and word frequency, especially in combination with multiple regression, it is highly desirable to explore thoroughly the effect of AoA in a domain in which there are *a priori* reasons to predict an effect of AoA, but in which the problem of inter-correlations between predictor variables would be less pronounced. We believe that naming photographs of famous faces provides a suitable task.

As models of face naming have developed by analogy with object naming models, it would be anticipated that AoA would affect the time taken to name a famous face. An analogy with the effect of word frequency in object naming has already been used as a basis to interpret effects of the frequency of surnames in the population on recognition and production of celebrities' names (Valentine, Brédart, Lawson & Ward, 1991; Valentine & Moore, 1995). Figure 1, shows simplified generic models of object and face naming. An object must first undergo perceptual analysis that generates an internal representation of its visual properties. If the object is familiar that percept will activate the appropriate object recognition unit. Activation will spread from the recognition unit to access semantic information about the object (e.g. knowledge of its use). Only after semantic information has been accessed can activation pass to the name retrieval stage where the appropriate phonology becomes available for the articulatory system. Such a hierarchy may be implemented within an interactive processing architecture (e.g. Humphreys *et al.*, 1995).

FIGURE 1 ABOUT HERE

The same hierarchy of representations is assumed to be required to name a familiar face. The major difference between models of object recognition and models of face recognition is the assumption that access to semantic information about people and their names is achieved via a Person Identity Node or PIN (Hay & Young, 1982). PINs play the role of token markers in memory (denoting an individual), and are assumed to be a critical difference between the processing of stimuli that take a proper name (e.g. celebrities) with those which take a common name (e.g. everyday objects). For discussion of this issue see Valentine, Brennen and Brédart (1996).

Notwithstanding, the analogy between naming an object and naming a famous face, the tasks do differ in a number of important respects. First, face naming requires production of a proper name whereas object naming requires the production of a common name. Therefore, the AoA of a celebrity refers to encountering an individual rather than a class of objects denoted by a common name. Second, the age at which one acquires familiarity with celebrities tends to be much later in life than the acquisition of many words (especially object names). One continues to acquire knowledge about new people throughout one's lifetime, as new celebrities become famous.

Therefore, there are some *a priori* reasons why the influence of AoA on face naming might be expected to be rather different from its effects in word and object naming.

Despite these differences, it has been a logical step to explore effects on face naming that may be analogous to the reported effects of word frequency on object naming. Valentine, *et al.* (1991) explored the effect of the frequency of surnames in the population, as assessed by counts of surnames in a telephone directory, on the recognition of the names of famous people. They found that the effect of surname frequency was analogous to the effect of word frequency in tasks that did not require recognition of the name (e.g. reading it aloud). However, the effect of surname frequency was analogous to the effects of facial distinctiveness in tasks that did require recognition of a celebrity's name. There was an advantage for low frequency (or distinctive) surnames in a name familiarity decision task. Valentine and Moore (1995) examined the influence of surname frequency when participants produced surnames in face naming tasks. The effect of surname frequency on recalling surnames, taught to previously unfamiliar faces, was analogous to the effect of word frequency on object naming (i.e. high frequency surnames were recalled more quickly and more accurately than low frequency surnames). However, when naming famous faces there was an advantage for participants to produce low frequency surnames. Valentine and Moore point out that these results can be explained in terms of differences between the underlying nature of surname frequency and word frequency and in terms of differences in the task demands between naming objects and naming people.

Surname frequency is not directly analogous to word frequency. For example, very many people in the UK share the name ‘Smith’, so its surname frequency will be high. The number of times the surname ‘Smith’ is encountered will be affected by the number of people who have the name and how frequently you encounter them (personally or in the media). However, if naming a famous face is assumed to require access to a representation of a full name, that in most cases is unique to an individual, a better analogy to word frequency would be the ‘familiarity’ of each celebrity. Rated familiarity may be considered as a cumulative frequency count of encounters with that individual[1]. In order to differentiate between familiarity with a surname (across individuals) and familiarity with each individual celebrity, measures of surname frequency and rated familiarity with each celebrity are included as independent variables in the first experiment[2].

EXPERIMENT 1

In Experiment 1 we used a correlational design to explore the extent to which a number of variables could predict the accuracy and latency of naming a set of famous faces. Participants were presented with digitised images of the most recent photographs available of celebrities' faces and were requested to provide three ratings for each celebrity:- familiarity with celebrity (or cumulative frequency), distinctiveness of the face and age of first encountering each celebrity. These rating tasks were followed by participants naming the same faces. Accuracy of response and naming latencies were the dependent variables.

The concept of familiarity with a word, picture or face is at first an obvious one; if a participant is not familiar with the face presented they could not name it. Familiarity is estimated by asking participants to rate their familiarity with each item (e.g. a where 1= completely unknown to 7= very familiar). In this way the measure reflects media exposure of individual celebrities and the preferences of individual participants. Ratings of familiarity are likely to reflect a subjective rating of cumulative frequency (i.e. the total degree of exposure to a celebrity) because it is an estimate of the relative frequency of encountering each celebrity. Instructions for familiarity ratings explicitly requested participants to let their ratings reflect how many times (prior to the experiment) a celebrity had been encountered in the media, etc.

As we intended the experiment to be analogous to previous work on object naming we have used a similar experimental paradigm and data treatment to that adopted by Morrison *et al.* (1992). Use of multiple regression is appropriate for an initial exploration of the number of independent variables, especially as there is no *a priori* reason to expect inter-correlations between AoA, familiarity and surname frequency. Bearing in mind the possible problems with multiple regression analyses, the intercorrelation statistics were monitored in order to test the model's strength. Our interpretation was restricted to variables with most predictive power (i.e. reaching a significant level to $p <.05$). Furthermore, we subsequently re-analysed the data factorially.

Method

Participants All participants took part in only one experiment. They were students of North East Universities, and were paid for their participation. There were two criteria for participation:- i. Participants must have spent the first 18 years of life in the UK. ii. Participants must be between 18 and 25 years old. There were 30 participants (11 male, 19 female) with a mean age of 20.2 years (s.d. = 1.6) who took part the experimental trials.

Materials and apparatus In pilot studies twenty-four 18 year old participants generated 210 celebrities' names. Twenty-four other participants (mean age = 19.2 years, s.d. = 2.0) rated the items for the AoA, and familiarity of the celebrities. None of these participants participated in the experimental trials. The most up-to-date images of 106 celebrities that received high familiarity ratings were selected for use in the experiment.

The following details apply to all materials and apparatus in all the experiments reported here. Images of 106 celebrities were created by scanning the most up-to-date, good quality photographs available or by capturing video stills. The pictures were monochromatic, 256 x 256 pixels in size (displayed at a resolution of 640 x 480 on a 14 inch screen) with 16 grey levels. The images were edited to obscure as much background and clothing as possible and pasted against a black background. Pictures were displayed individually in the centre of the PC screen. The experiment was controlled using the Micro Experimental Laboratory (MEL) software package (Schneider, 1988), which also randomised the order of stimuli for each presentation to each participant and logged responses and reaction times (with millisecond accuracy). Naming latency was measured by using a voice key connected to the port of the PC. The onset of the participants' vocal response was detected by use of a throat microphone.

Design A correlational design was used which had five independent variables: the degree of rated familiarity with celebrity (or cumulative frequency); rated distinctiveness of the celebrity's face, rated AoA for knowledge of the celebrity, log surname frequency (x + 1) and the number of phonemes in the celebrity's full name. Surname frequencies of celebrities' names were taken from a database created by a count of non-business surnames in the 1989 South Manchester telephone directory (Moore & Valentine, 1993). There were an estimated 261,105 non business surnames in the directory. Log (x + 1) of the number of occurrences of the surname per 100,000 entries was entered into the regression model. The two dependant variables were the latency to begin articulation of the correct name and the accuracy of response given.

Procedure Participants were informed the experiment had two parts; in the first part they were to make three rating responses (familiarity, distinctiveness and AoA for 106 faces) and that the second part would be explained later. They were told that this was not a memory test.

Familiarity, Distinctiveness and Age of Acquisition Rating Scales

The correlational design required the appropriate ratings to be obtained from the experimental participants who would also provide the naming latency data. The full set of stimuli were presented for each rating task. The instructions were phrased to emphasise that there was no right or wrong answers, but that personal opinion was the important factor. On each presentation participants were given as much time as required to make their decision and enter their rating. Responses were entered into the computer by moving the cursor to the chosen score and pressing the return key.

The rating tasks were always presented in a fixed order. First, familiarity ratings were made to all faces. Second, when participants were aware of the full range of facial types, the faces were rated for distinctiveness. Age of acquisition ratings were made on the third presentation.

Participants pressed the space bar when they were ready to enter their rating and the appropriate rating scale was displayed. Each subject saw the images in a different random order for each rating task.

Familiarity The instructions, presented on the PC screen, stressed that the ratings should reflect how many times, prior to the experiment, the celebrity had been encountered by the subject, on TV, films, newspapers, magazines, posters etc. Ratings were made on a 7 point scale (1 = completely unknown to 7 = very familiar). Rating scores were converted into a 6 point scale for analysis by removing stimuli rated as unknown from the analyses.

Distinctiveness Ratings were made on a 6, instead of a 7 point scale as 'unknown' response would be inappropriate. Participants were instructed to imagine that they had never seen each face before and so had no knowledge of individual characteristics other than those apparent in the grey scale images (e.g. height, hair colour, etc.). Participants were asked to imagine they had to go to a railway station to meet each of these people, and to "rate each face for how easy it would be to spot in a crowd". They were instructed to rate very typical faces that would be difficult to spot in a crowd as 1, and very distinctive or unusual faces that would be easy to spot in a crowd, as 6.

Age of Acquisition Ratings were made on a seven point scale with 1 being an unknown face. Participants estimated when they "first became aware...." of each celebrity. Number two related to a celebrity first acquired under 3 years of age; three, for a celebrity acquired under 6 years of age; four, a celebrity acquired under 9 years of age; five, a celebrity acquired under 12 years of age; six, a celebrity acquired under 18 years of age and seven, a celebrity acquired over 18 years of age. A key of this scale appeared while participants made their ratings. The scale was converted into a 6 point scale, as for familiarity.

The rating scales appeared on the screen when participants indicated they were ready to make their rating response. Participants were given as much time as required and advised to take a short break after completing each rating scale. It took approximately 1 to 1.5 hours to complete the rating scales, after which participants left the laboratory for a coffee break of about 20 minutes.

Face Naming Task

The naming task involved the same images as previously rated. Participants were asked to give the full name to each face as quickly and accurately as possible. If they did not recognise the celebrity to say "pass". Response latencies (with millisecond accuracy) were recorded using a voice key and a throat microphone. Participants were familiarised with this apparatus and informed of the importance to respond as quickly as possible.

Each trial began with the presentation of the fixation point of a "*" in the centre of a black screen which remained until participants indicated their readiness to continue. Participants were instructed to focus on the star and indicate their readiness to continue by tapping the desk. The experimenter initiated stimulus presentation. There was a warning tone for 250 msec. followed by a 250 msec. interval before the stimulus appeared. Participants attempted to name the face. The vocal response triggered the voice key and the image disappeared. The RT was logged by the computer. The experimenter recorded the accuracy of the response via the keyboard, rejecting incomplete or incorrect names, ‘don't know’ responses and accidental or inappropriate firing of the voice key. Only correct full names (first name and surnames) were accepted.

Ten different faces, not used in the experiment proper, were used in practice trials before all experimental tasks. None of these data were included in the analyses. Participants were encouraged to ask any questions both during and after the practice sessions. The naming task took about 30 minutes, after which participants were debriefed.

Results

Five items with high error scores: *Eric Clapton (93%), Harold Wilson (67%), Pam Ayres (60%), Kate Bush (53%) and Pat Phoenix (53%)* were removed from the analysis, reducing the error rate from 16.67% to 13%, leaving 101 items for the analyses. Out of the possible 3030 naming responses 1288 key misfirings, incorrect or 'pass' responses and tip of the tongue states were removed (43%) from the RT data. The minimum number of correct responses contributing to any mean naming latency was 14. The maximum number of responses was 30[3].

The procedure recommended by Lorch and Myers (1990) requires a full data set. Clearly with as much as 43% of the data points missing it would be inappropriate to attempt to estimate the missing data points. Naming famous faces is notoriously difficult. We note that the error rate found in this experiment is consistent with previous studies, which have produced error rates of over 40% (e.g. Cohen, 1990; Valentine, Moore, Flude, Young & Ellis, 1993; Valentine, Moore & Brédart, 1995). Therefore, we entered the mean latency and accuracy for each item into separate regression models. This was the procedure employed by Morrison, *et al.* (1992) and other authors. We include this analysis as an initial exploration of the data and to allow comparisons with previous work. We address the issues raised by Lorch and Myers (1990) by subsequently selecting matched sets of stimuli to examine the effect of AoA factorially. Furthermore, Experiment 2 reports a replication of the factorial data manipulation.

Reaction times of correct naming responses. The mean reaction times to name each of the remaining 101 celebrities was calculated. The individual ratings for familiarity, distinctiveness and AoA were collapsed across the 30 participants to give mean scores for each item. These were entered into the regression model together with the number of phonemes and log (x + 1) surname frequency as independent variables. Descriptive statistics appear in Table 1.

The mean naming latency for correct full name responses was 1383 msec. (s.d. = 206 msec. min = 1177 msec., max. = 4197 msec.). Naming latency was subjected to a reciprocal transformation ($^{100}/x$) to reduce the negative skew from -1.04 to 0.05 (Cohen & Cohen, 1983) prior to entry into the multiple regression model as the dependant variable 'naming speed'. All analyses reported were performed using one tailed Pearson's Correlation Coefficient and by simultaneous and stepwise multiple regression.

The relationships between naming speed and the five independent variables are shown in Table 1. Familiarity ($p<.01$) and AoA ($p<.01$) were significantly correlated with naming speed. Celebrities rated as very familiar or acquired early in life tended to be named faster than those rated as less familiar or later-acquired.

TABLE 1 ABOUT HERE

There were four significant intercorrelations. Celebrities rated as very familiar tended to be rated as having distinctive faces ($p<.01$), and were also rated as acquired earlier in life ($p<.01$) than celebrities rated as less familiar. Celebrities with faces rated as distinctive had earlier AoA ratings than those with typical faces ($p<.01$). Celebrities with low frequency surnames tended to have more phonemes than celebrities with high frequency surnames ($p<.05$).

When variables are intercorrelated, as with familiarity and AoA ($p<.01$) the regression model may not be robust. To ensure that this model was valid the tolerance and variance inflation factor (V.I.F.) were examined[4] The close proximity of the Tolerance (= 0.97) and VIF (=1.04) to 1 gives a good indication of the model's strength and reveals that this model is robust.

The results from the simultaneous multiple regression are shown in Table 1. Sixteen percent of the variance of naming speed was accounted for when all the independent variables were entered into the regression model, R^2=.16, S.E=.01; $F(5,94)=3.57$, $p<.005$. Two variables, familiarly and AoA, have significant standardised regression coefficients with naming speed.

A stepwise multiple regression analysis would confirm these apparent significant relationships. In the stepwise multiple regression variables are entered according to the diminishing magnitude of their simple correlation with the dependant variable.

Familiarity was first to enter the stepwise multiple regression R^2=.10, S.E.=.01; $F(1,98)$=10.54, p <.002, accounting for 10% of the variance in naming speed. Age of acquisition entered on the second step R^2 =.14, S.E.=.01; $F(2,97)$= 7.99, p <.001, accounting for a further 4% of the variance in naming speed. No other variables entered the equation.

The stimuli were originally selected on the basis of their high familiarity to facilitate accurate face naming. By removing familiarity from the regression model any overlap in the predicted variance in naming speed accounted for by familiarity, but also attributable to AoA should be revealed[5]. When familiarity was partialled out of the model, AoA significantly predicted 8% of the overall variance in naming speed. A predictive power of 2.5% had been shared between familiarity and AoA. Thus, AoA accounted for 8% of the variance in naming speed Sr^2=.08, S.E.=.01; F(1,98)=8.29, p<.005. Eight percent represents half of the variance explained by all of the independent variables when entered into a simultaneous regression model.

Accuracy of Naming Response. The mean number of correct naming responses (CR.) was calculated for each item across 30 participants (mean = 17.80, s.d. = 5.92). The descriptive data and results of correlational and regression analyses are shown in Table 2.

TABLE 2 ABOUT HERE

The data in Table 2 shows three significant correlations with accuracy of response, the higher the familiarity ratings given to celebrities, the more likely participants were to provide a correct response (p <.01). The higher the distinctiveness ratings given to celebrities, the more likely participants were to name them correctly (p <.05). More correct responses tended to be given for celebrities rated as acquired early in life, than those rated as acquired late in life (p<.01). The intercorrelations between independent variables were as described for naming speed.

The simultaneous multiple regression revealed that 33% of the variation in the accuracy data were significantly accounted for by all the independent variables, R^2 =.33, S.E.=4.76 F(5,95)=19.47, p <.0001. Only familiarity evinced a significant standardised regression coefficient (p <.01).

In the stepwise multiple regression familiarity accounted for 29% of the variance in accuracy of response R^2 =.29, S.E.=5.01; F(1,99)=40.47, p <.0001. Age of acquisition entered on the second step to significantly account for a further 3% of the variance in the accuracy data, R^2 =.32, S.E.=4.93; F(2,98)=23.01, p <.0001. No other variable significantly entered the equation. When familiarity was partialled out of the equation, AoA could significantly account for 9% of the variance, Sr^2=.09; S.E.=5.67; F(1,99)=9.99, p <.005. In calculations of unique variance AoA shared over 4% of the variance predicted by familiarity. Thus AoA significantly predicted 27% of the variance explained by all of the independent variables in the simultaneous regression model.

Discussion

Participants produced the correct full name in response to 47% of the celebrities. Multiple regression analyses showed that rated familiarity and AoA were the only independent variables which significantly predicted the variance of either of the dependant variables. Both naming speed and accuracy were significantly predicted by rated familiarity and by rated AoA. Highly familiar celebrities were named faster and more accurately than were less familiar celebrities. Celebrities of whom knowledge was acquired early in life were named faster and more accurately than celebrities first encountered later in life. As similar results were obtained in the analysis of naming speed and of accuracy there is no evidence of a trade-off between speed and accuracy.

In drawing an analogy with the object naming literature, it is most appropriate to think of the familiarity ratings as a subjective measure of cumulative frequency. The instructions given to subjects are perfectly consistent with this interpretation. Therefore, the speed and accuracy of naming familiar faces is affected by both (cumulative) frequency and AoA. This result is not analogous to the results of Morrison *et al.*(1992) who found that AoA but not word frequency affect object naming latency. However, a more recent study by Barry *et al.* (1997) did find that both spoken word frequency and AoA affected object naming latency.

The relationship between familiarity and the ease of naming a celebrity found in Experiment 1 is not surprising. Previously, Brédart (1993) has demonstrated that rated familiarity (of celebrity's names) was positively related to naming accuracy. Given the great difficulty people experience in recalling people's names, it would be a very counter-intuitive result indeed if rated familiarity was unrelated to naming speed and accuracy.

In a previous study, Valentine and Moore (1995) demonstrated that distinctiveness and surname frequency affect both naming latency and accuracy. However, these effects have not been found in this experiment. A combination of factors may be influencing the present results. Firstly, in our previous study, participants were practised in producing *surnames* only and the test condition was *surname production* to the same faces. Here participants were required to produce the full name, therefore *surname* frequency would be expected to have less effect. In addition, participants had not previously named the faces overtly. Naming practice reduces the number of errors and naming latencies compared to those reported in this present experiment. It may be necessary to reduce the variance in naming data by practice before the effect of surname frequency could be observed. Secondly, factorial manipulations employed extreme values of surname frequency in Valentine and Moore's experiments, but surname frequency is a continuous variable in the present study.

Rated distinctiveness, familiarity and AoA were all inter-correlated. The direction of the correlations suggests that there may be some cross-talk between the ratings. Good availability of knowledge about a celebrity may produce a tendency for the celebrity to be rated as more familiar, earlier acquired and more distinctive. Although significant these inter-correlations are lower than the reported inter-correlations between AoA and word frequency in object naming studies.

One of Lorch and Myers (1990) recommendations was to regress the raw data for each participant onto the predictor variables (instead of collapsing it across items as we have in replicating Morrison *et al.'s* paradigm). We subsequently performed this analysis to find a higher proportion of the overall variance would be significantly predicted by AoA (27%) for naming speed and for the accuracy data (30%). However, as we had a high proportion of missing data replaced with a linear trend of missing points (Nourusis, 1993) we feel it would be unwise to rely on such corrupt data and prefer to report the conservative analysis only.

Morris (1981) expressed two concerns on the limitations of stepwise multiple regression as a statistical tool. Firstly, high intercorrelations between variables can make this technique inappropriate for identifying causal relationships, providing spurious relationships between factors while diluting actual relationships. To address this issue we reported the tolerance and variance inflation factor statistics which demonstrated our models were robust (Nourusis, 1993).

Morris' second concern was that high intercorrelations make this technique inappropriate for identifying the influential variables because the order of entry into the regression model could influence the attribution of that variable's importance. We believe this concern was addressed when we created the regression model, because we incorporated the necessity of partialling familiarity out of the model for the final statistic and calculating the shared and unique proportion of R^2 (Cohen & Cohen, 1983; Howell, 1992; Tabachnickk & Fidell, 1996). However, we agree with Morris's concern about the possibility of stepwise multiple regression providing spurious relationships and diluting actual relationships. The intercorrelation statistics were examined to prevent spurious interpretations from our models. To prevent the possibility of any underlying assumptions being inadvertently violated a portion of data were analysed factorially. *Post hoc* we identified stimuli forming two extreme groups on AoA ratings that were matched on all other variables. The naming and accuracy data for these stimuli (from 30 participants) were re-analysed.

FACTORIAL ANALYSIS OF EXPERIMENT 1

The multiple regression analysis partially confirms the prediction derived by analogy to Morrison *et al.'s* (1992) data on object naming. Age of acquisition was a significant predictor of response accuracy and speed. However, familiarity (or cumulative frequency) was found to be the major predictor of both naming speed and accuracy of response.

The interaction reported by Barry *et al.* (1997) was such that was no significant difference occurred for the production of early-acquired or late-acquired high frequency object names, but the difference between early-acquired and late-acquired low frequency names was significantly different. By analogy, these findings suggest that no AoA effects will be apparent for celebrities rated as highly familiar. The results from Experiment 1 showed that the major predictor of both naming speed and accuracy of response was rated familiarity. This makes intuitive sense because only celebrities who are highly familiarly to participants are likely to be named successfully. Items for factorial analysis were selected on the basis of high familiarity ratings, thus ensuring maximum availability of data points for analysis. A prediction generated by analogy to Barry *at al.'s* picture naming data would be that AoA will not affect naming speed for these items because they are rated as very familiar (or of high cumulative frequency).

Method

Participants Data from all participants in Experiment 1 were included.

Materials Two groups of 25 stimuli were selected according to the following criteria: celebrities in both groups had a mean rating greater than 5 on the familiarity scale (mean score = 5.75; s.d. = 0.24). One group consisted of celebrities with a mean AoA rating below 5.5, the other group consisted of items with mean AoA ratings above 5.5; the two groups of celebrities were statistically matched on all other variables. There was a significant difference between early and late AoA in a one tailed t-test $t(1,24)=10.20$, $p<.0001$, no significant differences occurred between the other variables. The relevant data are shown in Table 3.

TABLE 3 ABOUT HERE

Design, apparatus and procedure Age of acquisition formed one within-participants factor with two levels (early vs. late). Naming latency and accuracy of response were the dependant variables. All details of the apparatus and procedure were specified for Experiment 1. The analysis was designed to take participants as the random factor. There are two reasons why an items analysis would be inappropriate:- firstly, the stimuli were assigned to the early or late AoA groups on the basis of a median split. When the mean of naming latency across all items in an experimental condition are taken in a participants analysis this provides an adequate test of the effect of AoA. However, if the mean naming latencies are calculated for each item there will be many items of early AoA which are acquired only slightly earlier than many of the late-acquired items. Therefore, one would not expect an effect of AoA in an items analysis. Secondly, the effect of AoA is between-items, therefore an items analysis would have less statistical powerful than the within-participants analysis.

Results

The naming speed and accuracy of response for the 50 critical items were isolated from the data of Experiment 1. The mean reaction time was 1370 msec. (s.d. = 257) with 77% (s.d. = 18%) correct full name responses. A one-way analysis of variance showed a significant effect on naming speed of correct responses between celebrities rated as acquired early (mean = 1292 msec., s.d. = 236 msec.) and acquired late (mean = 1446 msec., s.d. = 257 msec.); $F(1,29)=16.76$, $p<.001$. Celebrities rated as acquired early in life were named faster than celebrities rated as acquired late in life.

There was a significant difference in the number of full names produced correctly between celebrities rated as early-acquired (mean score = 19.77; s.d. = 4.31) and late-acquired (mean score = 18.20, s.d. = 4.69) in a one-way analysis of variance $F(1,29)=19.53$, $p<.0001$. Celebrities rated as acquired early in life were named with greater accuracy than celebrities rated as acquired late in life.

Discussion

Problems inherent in multiple regression designs in cognitive psychology (e.g. Morris, 1981; Lorch & Myers, 1990) were addressed initially by investigating the collinearity statistics and then by carefully selecting two sets of stimuli matched on all variables except AoA.

The prediction of no effect of AoA derived by analogy to Barry *et al.'s* (1997) picture naming data was not confirmed in the analysis. All items were selected on the basis of the high familiarity ratings (or high cumulative frequency ratings). Selection of celebrities with high familiarity ratings also produced a higher proportion of accurate responses for analysis (early AoA = 79%; late AoA = 72%).

The significant result of the factorial analysis supports the previous interpretation that AoA affects the latency and accuracy of naming famous faces.

Having established a robust AoA effect in both regression and factorial analysis a full replication was carried out to establish that the ratings collected and the observed effect of AoA were replicable with another group of participants, from the same population. The same procedure as Experiment 1 was employed with participants rating faces prior to naming them, but only the two matched groups of stimuli identified above were included.

EXPERIMENT 2

Method

Participants There were twenty-four participants (13 male and 11 female) in this experiment (mean age = 19.46, s.d. = 1.44 years).

Materials and apparatus The materials and apparatus were as described for the previous factorial design.

Design This was a within-participants single factor design with two levels, early vs. late AoA. Naming latency and accuracy of response were the dependant variables.

Procedure The procedure was as described for Experiment 1. Participants rated 50 celebrities (25 early-acquired and 25 late-acquired) on the three attributes and following a short break they participated in the naming task. The experiment lasted approximately 25 minutes.

Results

Naming latencies were subjected to a reciprocal transformation to reduce the skew in the data. Mean naming speeds and the number of correct responses to the 50 celebrities were calculated. The mean naming latency was 1428 msec. (s.d. = 389), with 67% (s.d. = 16%) accurate responses.

There was a highly significant effect of naming speed between early (mean = 1310, s.d. = 314 msec.) and late (mean = 1545, s.d. = 426 msec.) AoA in a one way analysis of variance $F(1,23)=39.42$, $p<.0001$. Celebrities rated as acquired early in life were named faster than celebrities rated as acquired late in life.

There was a significant effect of accuracy between early (mean = 17.67, s.d. = 3.73) and late (mean = 15.96, s.d. = 4.14) AoA, in a one way analysis of variance $F(1,23)=8.44$, $p<.008$. Participants were able to correctly name more celebrities rated as acquired early in life than celebrities rated as acquired late in life.

Item Reliability Measures Data from participants' rating scores were analysed by paired t tests to confirm the validity of the measure used for the experimental groups. A priori (ratings from Experiment 1) and *post hoc* rating (from participants from this Experiment) confirmed the validity of the ratings obtained in Experiment 1 (see Table 3).

Discussion

By replicating the advantage for naming celebrities rated as acquired early in life using an experimental design free of the problems inherent in multiple regression designs, we have established that rated AoA does have a consistently robust effect on both participants' speed and accuracy of producing celebrities' names. The advantage of early AoA was found for highly familiar celebrities. This result differs from the results Barry *et al's* (1997) study of picture naming, inasmuch as they found the effect of AoA was restricted to pictures with low spoken frequency names.

It could be argued that data from Experiments 1 and 2 are contaminated by prior exposure to the stimuli, because three ratings were obtained from participants prior to collecting the face naming data. The rating tasks were performed prior to naming because naming faces is an enormously difficult task to perform. Priming from prior exposure usually enhances subsequent recognition and therefore may have been expected to increase the accuracy of subsequent face naming. Nevertheless, the error rate in Experiment 1 was very high (approximately 40% of trials). It may be the case that collecting ratings was not effective in enhancing face naming accuracy because the rating tasks did not require production of the celebrities' names.

Experiment 3 had two main aims. First, to investigate whether an effect of AoA would be observed in naming data collected from the participants' first encounter with the stimuli in the experiment. Second, to investigate the possibility of using practice in naming the faces to reduce the error rate and variance of naming latency. Highly practised subjects may produce more homogeneous naming data but the practice could reduce or eliminate the effect of AoA. The purpose of exploring the effect of practice was to provide data to inform design of future experiments. Experiment 3 also serves the purpose of a further replication of the effect of AoA on naming familiar faces.

The aims of Experiment 3 meant that any effect of AoA on the first and last encounter with the stimuli were of particular interest. Participants named the same faces repeatedly over four blocks. To equate practice in name production across items, participants were instructed not to guess the name of any celebrity they could not produce, but to say "pass". When this occurred the name was supplied for them to repeat aloud, thus participants were practised equally on face recognition and name production for all items. Any effect due to priming should be apparent from Block 2 (the second presentation), and be maximally apparent in Block 4.

EXPERIMENT 3

Method

Participants There were 24 participants (9 male, 15 female) in this experiment, they had a mean age of 19.3 years (s.d. = 1.3).

Materials, Apparatus and *Stimuli* The materials and apparatus used were the same as described for Experiment 2.

Design This was a within-participants two factor design with two levels of AoA (early vs. late) and four levels of Block (1 to 4). Naming latency and accuracy of response were the dependent variables.

Procedure Participants were informed that the experiment was in two parts and that the second part would be explained later, but that it was not a memory test. In the naming task participants were asked to give the full name to each face as quickly and accurately as possible. If they could not name a celebrity they were instructed *"do not guess, but say pass and the name will be provided for you. It is important that you repeat such names aloud"*. Participants were prompted to repeat names aloud on each such occurrence.

The "*" fixation point began each trial until participants indicated readiness to continue by tapping the desk. Stimulus presentation was initiated by the experimenter. Each trial commenced with a 250 msec. tone and followed by 250 msec. interval before the stimulus appeared. Participants attempted to name the face. The vocal response triggered the voice key terminating the display of the image and logging the naming latency by the computer. The experimenter accepted only correct full names for data analysis by entering a code on the keyboard. There were four blocks in this experiment with a short break between each block. Participants were provided with the correct name if they were unable to produce it after 10 seconds and reminded to repeat it aloud. There were 10 practice trials using ten celebrities not included in the experimental trials. Data from the practice trials were excluded from the analyses. The task took about 15 min, after which participants performed rating tasks as described for Experiments 1 and 2. The rating tasks took approximately 25 minutes to perform. Finally, participants were debriefed.

Results

Mean naming latencies were subjected to a reciprocal transformation to reduce a skew in the data. Mean naming speeds of correct responses and accuracy of response to the 50 celebrities were calculated, the mean data are shown in Figure 2 and 3 respectively.

FIGURES 2 and 3 ABOUT HERE

Analysis of Naming Speed From the possible maximum (4,800) data points a total of 1295 (27%) errors, key misfiring or tip of the tongue states were removed from the analyses. The mean naming latencies of correctly named faces are shown as a function of experimental block and AoA in Figure 3 (mean = 1578, 1276, 1160 and 1108 msec. for blocks 1 to 4 respectively). These data were subjected to a two way repeated measures analysis of variance.

There was a highly significant main effect of block, $F(1,69)=56.05$; $p<.0001$. Participants named the celebrities progressively faster as the stimuli were repeated across blocks. Multiple comparisons showed that naming speed increased significantly between blocks 1 and 2 $F(1,47)=57.79$, $p<.0001$ (mean latency decreased by 302 msec.). Naming speed also increased significantly between blocks 2 and 3 $F(1,47)=18.50$, $p<.0001$ (mean latency decreased by 116 msec.) and between blocks 3 and 4, $F(1,47)=4.20$; $p<.05$. (mean latency decreased by 52 msec.). There was a highly significant main effect of AoA $F(1,23) = 13.99$; $p<.001$. Participants were faster to correctly name celebrities rated as acquired early in life than late-acquired celebrities (mean = 1233 msec. vs. 1328 msec., respectively). There was no interaction between block and AoA $F<1$.

Analysis of Accuracy The mean number of correct responses for each block are shown (in Figure 3) as a function of experimental block and AoA (mean = 51%, 73%, 81% and 86% for blocks 1 to 4 respectively). There was a highly significant main effect of block, $F(1,69)=116.37$; $p<.0001$. Participants produced more correct celebrity' names with repetition. Multiple comparisons showed that participants' accuracy of response improved significantly across blocks 1 and 2 (difference = 5.56) $F(1,47)=152.47$; $p<.0000$; between blocks 2 and 3 (difference = 2.00) $F(1,47)=33.93$; $p<.0000$; and between blocks 3 & 4 (difference = 1.21) $F(1,47)=18.94$; $p<.0001$). There was a significant main effect of AoA $F(1,23) = 6.99$; $p<.02$. Celebrities rated as acquired early in life were named with greater accuracy than celebrities rated as acquired late in life (18.73 vs. 17.76, respectively). There was no interaction between block and AoA $F< 1$.

Analysis of Naming Speed and Accuracy by Block The stated aims of the experiment required separate analysis of the effect of the first and last presentation. For completeness all blocks were analysed separately.

Block 1 Forty-nine percent of trials were excluded from the analysis of naming latency (errors, key misfiring or tip of the tongue states: 45.2% of early AoA items and 52.3% of late AoA items). The mean naming latency was 1281 msec. (s.d. = 312 msec.). In a one way analysis of variance there was a significant effect of AoA on naming speed (early mean = 1526 msec., s.d. = 175 msec.; late mean = 1630 msec., s.d. = 186 msec.) $F(1,23)=4.69$, $p <.04$. There were 51% (s.d. = 16%) accurate responses given. A one way analysis of variance revealed a significant effect of AoA on the accuracy data (early mean latency = 13.71, s.d. = 4.95; late mean latency = 11.83, s.d. = 4.36), $F(1,23)=7.50$, $p<.01$. On the first presentation, participants were significantly faster and more accurate to name celebrities rated as acquired early in life than to name celebrities rated as acquired late in life.

Block 2 Twenty-seven percent of trials were excluded from the analysis of naming latency (errors, key misfiring or tip of the tongue states: 25.7% of early AoA items and 27.7% of late AoA items). The mean naming latency was 1276 msec. (s.d. = 237 msec.). There was a highly significant effect of AoA on naming speed in a one way analysis of variance (early mean latency = 1219 msec., s.d. = 214 msec.; late mean latency = 1334 msec., s.d. = 249 msec.) $F(1,23) = 8.55$, $p < .008$. There were 73% (s.d. = 14%)accurate responses given. There was no effect of AoA in the accuracy data (early mean = 18.58, s.d. = 3.50; late mean = 18.08, s.d. = 3.49; $p > .1$). On the second presentation participants were significantly faster to name celebrities rated as acquired early in life than celebrities rated as acquired late in life.

Block 3 Nineteen percent of trials were excluded from the analysis of naming latency (errors, key misfiring or tip of the tongue states: 16.5% of early AoA items and 20.8% of late AoA items). The mean naming latency was 1160 msec. (s.d. = 244 msec.). There was a significant effect of AoA on naming speed in a one way analysis of variance (early mean = 1116 msec., s.d. = 221 msec.; late mean = 1205 msec., s.d. = 263 msec.), $F(1,23) = 4.17$, $p < .05$. There were 81% (s.d. = 13%) accurate responses given. There was no effect of AoA on the accuracy data (early mean = 20.88, s.d. = 2.79; late mean = 19.79, s.d. = 3.74), $p > .1$). On the third presentation participants were significantly faster to name celebrities rated as acquired early in life than celebrities rated as acquired late in life.

Block 4 Fourteen percent of trials were excluded from the analysis of naming latency (errors, key misfiring or tip of the tongue states: 13.0% of early AoA items and 14.7% of late AoA items). The mean naming latency was 1107 msec. (s.d. = 222 msec.). The difference in naming speed between early and late AoA was not significant (early mean latency = 1073 msec., s.d. = 208 msec.; late mean latency = 1143 msec., s.d. = 234 msec., $p < .08$). There were 86% (s.d. = 10%) correct responses given. There was no significant difference between early and late AoA in the accuracy data (early mean = 21.75, s.d. = 2.35; late mean = 21.33, s.d. = 2.46, $p > .1$).

The data in Figures 3 and 4 reveal a consistent pattern of results. Participants named celebrities rated as acquired early in life faster than they named celebrities rated as acquired later in life. This effect was significant in blocks 1 - 3, but just short of statistical significance in block 4. Participants named early acquired celebrities more accurately than they named late acquired celebrities in block 1. The effect of AoA on accuracy was not significant in blocks 2 - 4.

Item Reliability Measures Data from participants' ratings were analysed by paired *t* tests to confirm the validity of the measure used to select the two sets of stimuli. A *priori* ratings (from Experiments 1 and 2 combined) and *post hoc* ratings (from participants from this Experiment) confirmed the validity of the selection of stimulus sets (see Table 3).

Discussion

Participants were consistently faster and to name celebrities rated as acquired early in life than they were to name celebrities rated as acquired later in life, even on the first presentation. This result supports our interpretation of Experiment 1 and 2 as showing robust effects of AoA. There is no support for any suggestion of a possible artefact arising from repeated exposure to the stimuli. In fact, repetition of spoken name production appeared to reduce or remove the effects of AoA. The effect on naming latency was not statistically significant in block 4. The effect of AoA on accuracy of naming response observed in block 1 was not present in blocks 2 - 4. In all cases the differences in accuracy and latency, even when not significant, were in the direction of faster and more accurate naming of early acquired celebrities.

Practice in naming celebrities clearly increased naming accuracy and reduced naming latency. However, in view of our finding that the effect of AoA in block 4 was not significant, it would be unwise to use practice in *naming* faces to enhance accuracy and reduce the variability of naming latency in any future experiments that aim to investigate the effects of AoA. Experiments 1 and 2 showed robust effects of AoA in naming celebrities on the fourth presentation of the stimuli. However, participants had rated the faces rather than named them on the previous exposures. The effect of mere exposure to the celebrity faces was very different to practice in naming them (with feedback). The accuracy of naming in block 4 of Experiment 3 was considerably higher than in Experiment 2 which used the same stimulus set (86% vs. 67% respectively). A reliable effect of AoA on naming accuracy and latency was found in Experiment

2 but neither dependant variable showed a significant effect of AoA in block 4 of Experiment 3.

General Discussion

As models of face naming have developed by analogy to models of object naming, the finding that AoA affects speed of object naming raised the question of whether AoA affects the speed of naming celebrities' faces. In this paper we have demonstrated a robust advantage in both speed and accuracy of name production for early-acquired celebrities' names in both multiple regression and factorial designs. Participants evinced a facilitation in speed and accuracy when producing celebrities' names rated as acquired early in life, both with and without prior presentation of the face. As far as we are aware, studies of object naming have not reported an effect of AoA on naming accuracy (e.g. Morrison, *et al.*, 1992; Morrison & Ellis, 1995; Barry *et al.,* 1997).

Our results establish that the effect of AoA on name production can be found in a task that requires production of proper names, acquired much later in life than the acquisition of object names.

Age of acquisition effects have been demonstrated in object naming tasks (e.g. Morrison, *et al*. 1992) and have been reported to act independently of word frequency in lexical decision tasks (e.g. Morrison & Ellis, 1995; Turner, *et al*. in press). Barry *et al.* (1997) found that spoken word frequency, name agreement and the interaction of spoken word frequency with AoA are the predominant predictors of object naming speed. The interaction was such that AoA did not affect the latency to produce high frequency names. In this paper we have established an effect of AoA in the domain of naming people. We have established that AoA affects the accuracy and speed of name production and that this effect is found for faces rated as highly familiar (or of high cumulative frequency). It is important to note that even when familiarity was partialled out of the regression model in Experiment 1, AoA was the only variable to predict either speed or accuracy of name production.

The effects AoA now established for tasks involving processing of words, objects and faces may present serious problems current models of cognition. Connectionist models which use backward error propagation to learn distributed representations, can readily model the effects of cumulative frequency or familiarity. However, these networks suffer from interference of early learned material by subsequently acquired material. Therefore, it is not clear how such an architecture could model an effect of AoA. Interactive activation models do not generally include a learning mechanism, however Burton (1994) has developed an algorithm which enables interactive activation models to learn localist representations of new stimuli. It can be readily appreciated how this algorithm can model the effects of cumulative frequency (or familiarity) by increasing the weight of connections between nodes. However, it is remains to be seen whether this class of model can simulate the effect of AoA.

Kohonen (1984; 1990) proposed a model based on 'self-organising maps'. This type of network is capable of learning to distinguish between different patterns of input by unsupervised learning. Similar patterns cluster at units in the same area, whereas dissimilar patterns are topographically distant. When Morrison (1993) attempted to simulate AoA effects by introducing a specific order of different patterns, there was a suggestion that early-acquired patterns remained distributed over a greater area, with later-acquired patterns sandwiched between them. Therefore, early encountered patterns play a prominent role in organising the representation of inputs. However it remains to be seen whether a self-organising network could provide an adequate model of AoA and cumulative frequency.

It has been suggested the effect of AoA reflects the development of mental lexicons. Brown & Watson (1987) proposed that the locus of the effect of AoA for object naming is at the stage of name retrieval. Yamazaki, Ellis, Morrison and Lanbon Ralph (in press) have demonstrated that the speed of reading Japanese Kanji characters showed effects of both the age at which the word entered the spoken vocabulary and the age at which Japanese children learn the characters used to write the words. These effects suggest that AoA affects the quality of lexical representations in both the visual input lexicon and the speech output lexicon, requiring at least two loci of AoA and Morrison and Ellis's (1995) AoA effects for reading speed and lexical decision to English words is consistent with this interpretation. However, Gerhand and Barry (in press) report additive effects of AoA and frequency in a word reading paradigm, where these variables were manipulated orthogonally. They interpret the frequency effects as an input variable and the effects of AoA as an output variable.

The data reported here have been restricted to experiments on face naming. It is not possible to speculate on the locus of the effect of AoA on naming people on the basis of these data alone. However, we have recently reported investigations of the effect of AoA on a number of tasks involving processing of people's names and faces with the aim of identifying the locus or loci of AoA in person recognition (Moore & Valentine, 1997).

Significant effects of AoA were found when participants read aloud celebrities' printed names and made familiarity decisions to famous faces and to printed names. However, effects of AoA were not found when participants made semantic decisions, based on occupation, to famous faces or to printed names. We suggest that there must be more than a single locus of AoA in person processing to account for these data. The finding that familiarity decisions to names and faces are affected by AoA suggest that either face recognition units and the representations used to recognise people's names provide loci of AoA. An alternative is that person identity nodes provide a locus for these effects. The finding that face naming and reading names aloud show effects of AoA would be consistent with a locus of AoA at the stage of lexical access for production of people's names. On the basis of these data on person recognition and the data on word and object processing it seems very unlikely that there is a single locus of AoA. It seems more likely that AoA is a fundamental effect of developing mental representations in a number of stimulus domains.

REFERENCES

Baayen, H., Piepenbrock, R., & van Rijn, H. (1993). *The CELEX lexical database (CD-ROM).* Philadelphia: PA: University of Pennsylvania Linguistic Data Consortium.

Barry, C., Morrison, C. M., & Ellis, A. W. (1997). Naming the Snodgrass and Vanderwart pictures: Effects of age of acquisition, frequency and name agreement. *Quarterly Journal of Experimental Psychology,* **50A**, 560-585.

Brédart, S. (1993). Retrieval failures in face naming. *Memory*, **1**, 351-366.

Brown, G. D. A., & Watson, F. L. (1987). First in, first out: Word learning age and spoken word frequency as predictors of word familiarity and word naming latency. *Memory and Cognition*, **15**, 208-216.

Burton, A. M. (1994). Learning new faces in an interactive activation and competition model. *Visual Cognition*, **1**, 313-348.

Butterfield, G. B. & Butterfield, E. C. (1977). Lexical codability and age. *Journal of Verbal Learning and Verbal Behaviour,* **16**, 113-118.

Carroll, J. B. & White, M. N. (1973a). Word frequency and age of acquisition as determiners of picture-naming latency. *Quarterly Journal of Experimental Psychology,* **25**, 85-95.

Carroll, J. B. & White, M. N. (1973b). Age-of-Acquisition norms for 220 picturable nouns. *Journal of Verbal Learning and Behaviour,* **12**, 563-576.

Cohen, G. (1990). Why is it difficult to put names to faces? *British Journal of Psychology,* **81**, 287-297.

Cohen, J. & Cohen, P. (1983). *Applied Multiple Regression / Correlation Analysis for Behavioural Sciences.* (2nd edition), Hillisdale NJ: L.E.A.

Gerhand, S. & Barry, C. (In press). Word frequency effects in oral reading are not merely age of acquisition effects in disguise. *Journal of Experimental Psychology, Learning Memory and Cognition*.

Gilhooly, K. J. (1984). Word age of acquisition and residence time in lexical memory as factors in word naming. *Current Psychological Research and Reviews,* **3**, 24-31

Gilhooly, K. J. & Gilhooly, M. L. M. (1979). Age of acquisition effects in lexical and episodic memory tasks. *Memory and Cognition,* **7**, 214-223.

Gilhooly, K. J. & Logie, R. H. (1981). Word age-of-acquisition and visual recognition thresholds. *Current Psychological Research*, **1**, 215-226.

Gilhooly, K. J. & Logie, R. H. (1982). Word age-of-acquisition and lexical decision making. *Acta Psychologia,* **50**, 21-34.

Hay, D. C., & Young, A. W. (1982). The human face. In A. W. Ellis (ed.), *Normality and Pathology in Cognitive Functions*. London: Academic Press.

Hofland, K., & Johansson, S. (1982). *Word frequencies in British and American English.* Bergen. The Norwegian Computing Centre for the Humanities.

Howell, D. C. (1992). *Statistical methods for Psychology. 3rd Edition*. Duxbury Press, Belmont, California: An imprint of Wadsworth Pub. Co.

Humphreys, G. W., Lamote, C. & Lloyd-Jones, T. J. (1995). An interactive activation approach to object processing: Effects of structural similarity name frequency, and task in normality and pathology. *Memory,* **3**, 535-586.

Humphreys, G., Riddoch, M. J., & Quinlan, P. T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology*, **5**, 67-103.

Kohonen, T. (1984). *Self organising and associative memory.* Berlin: Springer Verlag.

Kohonen, T. (1990). *Self organising map.* The Proceedings of the IEEE, **78**, 1464-1480.

Kucera, H. & Francis, W. N. (1967). *Computational analysis of present-day American English.* Providence, RI: Brown University Press.

Lachman, R. (1973). Uncertainty effects on time to access the internal lexicon. *Journal of Experimental Psychology*, **99**, 199-208.

Lachman, R., Shaffer, J. P. & Hennrikus, D. (1974). Language and cognition: Effects of stimulus codability, name-word frequency, and age of acquisition on lexical reaction time. *Journal of Verbal Learning and Verbal Behaviour*, **13**, 613-625.

Legrand, J. (1993). (Ed.) *Chronicle of the 20th Century.* Hampshire. UK: JOL International Publications.

Lorch, Jr. R. F. & Myers, J. L. (1990). Regression analyses of repeated measures data in cognitive research. *Journal of Experimental Psychology: Learning Memory and Cognition*, **16**, 149-157.

Monsell, S., Doyle, M. C., & Haggard, P.N. (1989). Effects of frequency on visual word recognition tasks: Where are they? *Journal of Experimental Psychology: General,* **118,** 43-71.

Moore, V. & Valentine, T. (1993)**.** *Names Count: An assessment of surname frequencies in North West England.* Unpublished software database.

Moore, V. & Valentine, T. (1997). Proposed loci for age of acquisition effects found in processing celebrity faces and names. Dublin, July 1997: Paper presented to the Fifth European Congress of Psychology.

Morris, P. E. (1981). Age of Acquisition, imagery, recall, and the limitations of multiple-regression analysis. *Memory and Cognition*, **9**, 277-282.

Morrison, C. M. (1993). *Loci and roles of word age of acquisition and word frequency in lexical processing.* University of York, York, UK: Unpublished D.Phil. thesis.

Morrison, C. M., Chapell, T. & Ellis, A. W.(1997). Age of acquisition norms for a large set of object names and their relation to adult estimates and other variables. *Quarterly Journal of Experimental Psychology,* **50A**, 528-559.

Morrison, C. M., Ellis, A. W. & Quinlan, P. T. (1992). Age of acquisition, not word frequency, affects object naming not object recognition, *Memory and Cognition,* **20**, 705-714.

Morrison, C. M. & Ellis, A. W. (1995). The roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory and Cognition,* **21**, 116-133.

Nourusis, M. J. (1993). *SPSS for Windows*^{TM}*: Base Systems User's Guide: Release 6.0.* Illinois 60611, Chicago: SPSS Inc.

Oldfield, R. C. (1966). Things, words and the brain. *Quarterly Journal of Experimental Psychology,* **18**, 340-353.

Oldfield, R. C. & Wingfield, A. (1964). The time it takes to name and object. *Nature,* **202**, 1031-1032.

Oldfield, R. C. & Wingfield, A. (1965). Response latencies in naming objects. *Quarterly Journal of Experimental Psychology,* **17**, 273-281.

Rubin, D. C. (1980). 51 properties of 125 words: A unit analysis of verbal behaviour. *Journal of Verbal Learning and Verbal Behaviour.* **19**, 736-755.

Scarborough, D. L., Cortese, C., & Scarborough, H. S. (1977). Frequency and repetition effects in lexical memory. *Journal of Experimental Psychology,* **3**, 1-17

Schneider, W. (1988). Micro Experimental Laboratory. An integrated system for IBM-PC compatibles. *Behaviour Research Methods. Instruments and computers.* **20**, 206-217.

Seidenberg, M. S., Waters, G. S., & Barnes, M. A. (1984). When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behaviour*, **23**, 282-404.

Seidenberg, M. S., & McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, **96**, 523-68.

Snodgrass, J. G. & Vanderwart, M. (1980). A standard set of 260 pictures: Norms for name agreement, image agreement, familiarity and visual complexity. *Journal of Experimental Psychology: Human Learning and Memory,* **6**, 174-215.

Tabachnickk, B. G. & Fidell, L. S. (1996). *Using Multivariate Statistics: Third Edition.* New York: Harper Collins College Publishers.

Thorndike, E. L. & Lorge, L., I. (1944). *The Teacher's Wordbook of 30,000 Words.* New York: Teachers College Press.

Turner, J. E., Valentine, T. & Ellis, A. W. (in press). Contrasting effects of age of acquisition and word frequency on auditory and visual lexical decision. *Memory and Cognition.*

Valentine, T. (1991). A unified account of the effects of distinctiveness, inversion and race of face in face recognition. *Quarterly Journal of Experimental Psychology*, **43A**, 161-204.

Valentine, T., Brédart, S., Lawson, R. & Ward, G. (1991). What's in a name? Access to information from people's names. *European Journal of Cognitive Psychology.* **3**, 147-176.

Valentine, T., Brennen, T. & Brédart, S. (1996). *The cognitive psychology of proper names. On the importance of being Ernest.* London: Routledge.

Valentine, T., Moore, V., Flude, B., Young, A., & Ellis, A. (1993). Repetition priming and proper names. *Memory,* **1**, 329-349.

Valentine, T. & Moore, V. (1995). Naming Faces: The effects of facial distinctiveness and surname frequency. *Quarterly Journal of Experimental Psychology*, **48A**, 879-894

Valentine, T., Moore, V. & Brédart, S. (1995). Priming production of people's names. *Quarterly Journal of Experimental Psychology*, **48A**, 513-535.

Yamazaki, M. Ellis, A. W., Morrison, C. M., Lambon Ralph, M. A. (1997). Two age of acquisition effects in the reading of Japanese Kanji. *British Journal of Psychology,* **88**, 407-421.

Footnotes

[1] Carroll and White (1973a) used this term to denote a combination of AoA and word frequency combination, similar to Gilhooly's (1984) "residence time in memory" (deducting AoA ratings from participant's age). We employ this term as a measure of participant's individual estimates of how well they knew celebrities, which would include frequency of encounter accumulated over the life span.

[2] Measures of frequency employed here refer to a count of surname frequencies taken from the South Manchester telephone directory. This measure have evinced high correlations of samples drawn from other telephone directories. For example, between South Manchester and Durham, r = .87; between North Manchester and South Manchester, r = .94; between South Manchester and Exeter r = .91 (Moore & Valentine, 1993).

[3] An anonymous reviewer suggested scoring 'errors' for completeness of response. This may be an appropriate method of reducing errors for object naming data. For example, a picture of a plant pot may be named as a pot, tub, plant holder etc. Such alternatives would be acceptable identifiers of the picture. However, face naming is different. The response "Prime Minister" to Tony Blair's face (the British PM) does uniquely identify the person. However, this response is a semantic decision by occupation and not face naming. In terms of the face recognition model naming a face requires extra processing and would therefore affect speed of response. A further problem is that production of a low frequency surname (e.g. Thatcher) in response to a picture of Margaret Thatcher could be considered to uniquely identify her, whereas the response of "Smith" (a high frequency surname) to a picture of John Smith would not uniquely identify a single person, as other politicians share the same surname (e.g. Cyril Smith, Chris Smith etc.)

[4] Tolerance refers to the degree to which one predictor can itself be predicted by the other predictors in the model. The tolerance of a variable i is defined as $1 - R_i^2$ where R_i is the multiple correlation coefficient when the ith independent variable is predicted from the other

independent variables. The variance inflation factor is defined as the reciprocal of the tolerance, that is for the *i*th regression coefficient, VIF $i = 1/(1 - R_i^2)$. This quantity is called the V.I.F, since the term is involved in the calculation of the variance of the *i*th regression coefficient. As the VIF increases, so does the variance of the regression coefficient (Nourusis, 1993)

5 Two stepwise regressions partial out the other independent variable's (I.V.) contribution:- I.V. A = greater proportion of the variance; I.V. B = the smaller contributor.

$$\frac{\text{I.V. B}}{\text{I.V. B} + \text{I.V. A}} = \text{shared variance}$$

(see Tabachnickk, & Fidell, 1996; pp 146-157).

Figure Caption.

Figure 1: A comparison of generic models of object and face recognition.

∗ note. For simplicity, most links are represented as unidirectional and the processes of lexicalization are represented by a single output lexicon.

Figure 2: Experiment 3: Mean naming speeds as a function of block.

Figure 3: Experiment 3: Mean accuracy data, as a function of block.

Variables	1 N.Sp.	2 Fam.	3 Dist.	4 AoA	5 Frequ.	6 Phoneme	*B*	β	Sr.^{2} (unique)
1. (DV) N.Sp	1.00	.31**	.09	-.28**	.04	-.10	.	.	
2. Familiarity		1.00	.27**	-.26**	-.07	.12	.004**	.30	.10**
3. Distinctiveness			1.00	-.37**	.10	-.03	-.000	-.08	
4. Age of Acquisition				1.00	-.07	.10	-.003*	-.23	.08**
5. Log Frequency					1.00	-.18*	.000	.04	
6. Phoneme Length						1.00	-.000	-.10	
						Intercept = .08			
Means	1383.46	5.23	3.55	5.34	1.22	9.58			
Sd.	205.96	0.82	0.94	0.82	0.92	1.90			

R^2 = .16[a]
Adjusted R^2 = .11
R = .40**
S.E. = .01

N.Sp. = Naming Speed *= p<.05 **= p<.01
B = Unstandardize Regression Coefficients; β = Standardised Regression Coefficients; Sr.^{2} = Semi Partial Correlation
a. Unique variability = 13%; Shared variability = 2.5%.

TABLE 1: Regression on naming speed data for analysis of Experiment 1

Variables	1 C.R.	2 Fam.	3 Dist.	4 AoA	5 Frequ.	6 Phoneme	*B*	β	$Sr.^2$ (unique)
1. (DV) CR	1.00	.54**	.23*	-.30**	-.11	-.00	.	.	
2. Familiarity		1.00	.27**	-.26**	-.07	.12	3.51**	.49	.29**
3. Distinctiveness			1.00	-.37**	.10	-.03	.28	.05	
4. Age of Acquisition				1.00	-.07	.10	-1.18	-1.78	.09**
5. Log Frequency					1.00	-.18*	-.066	-1.21	
6. Phoneme Length						1.00	-.19	-.70	
						Intercept = 7.42			
Means	17.80								
Sd.	5.92								
									R^2 = .33[a]
								Adjusted	R^2 = .30
									R = .58**
									S.E. = 4.96

C.R. = Number of Correct Responses; *= $p<.05$ **= $p<.01$

B = Unstandardize Regression Coefficients; β = Standardised Regression Coefficients; $Sr.^2$ = Semi Partial Correlation

a. Unique variability = .25%; Shared variability = 4.2%.

TABLE 2: Regression on accuracy data from analysis of Experiment 1

	Early Age of Acquisition		Late Age of Acquisition		
EXPERIMENT 1	**Mean**	**Range**	**Mean**	**Range**	**Difference**
Familiarity	5.85	2.35	5.65	1.82	<n.s.>
	(.63)		(.41)		
Distinctiveness	3.81	3.52	3.35	2.99	<n.s.>
	(1)		(.86)		
Age of Acquisition	4.51	2.19	6.05	1.25	t(1,24)=10.20;
	(.61)		(.34)		p<.0001
EXPERIMENT 2					
Familiarity	5.26	3.35	5.23	2.77	<n.s.>
	(.89)		(.59)		
Distinctiveness	3.84	3.62	3.40	3.38	<n.s.>
	(.99)		(.85)		
Age of Acquisition	4.53	1.89	5.72	1.20	t(1,24)=9.08;
	(.49)		(.37)		p<.0001
EXPERIMENTS 1 & 2					
Familiarity	5.89	2.30	5.67	1.79	<n.s.>
	(.60)		(.48)		
Distinctiveness	3.75	3.49	3.55	3.04	<n.s.>
	(.97)		(.86)		
Age of Acquisition	4.48	2.73	6.15	1.52	t(1,24)=9.07;
	(.57)		(.34)		p<.001
EXPERIMENT 3					
Familiarity	5.17	3.33	5.39	2.77	<n.s.>
	(.80)		(.52)		
Distinctiveness	3.78	3.54	3.36	3.38	<n.s.>
	(.92)		(.85)		
Age of Acquisition	4.50	2.34	5.80	1.82	t(1,24)=9.88;
	(.44)		(.37)		p<.0001

TABLE 3: Mean Rating Scores from Experiments 1 - 3 as a function of age of acquisition. Standard deviations are shown in parentheses. Range = the difference between the lowest and highest rating scores. Difference = the outcome of paired t tests between ratings for early and late age of acquisition items.

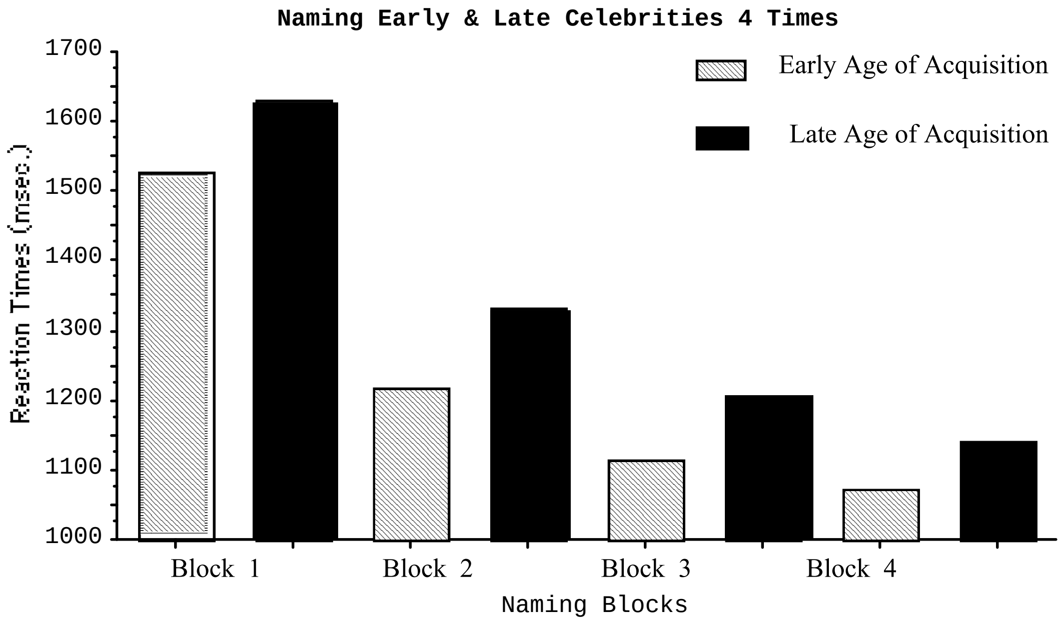

FIGURE 2: Mean reaction times for naming celebrity faces in Experiment 3

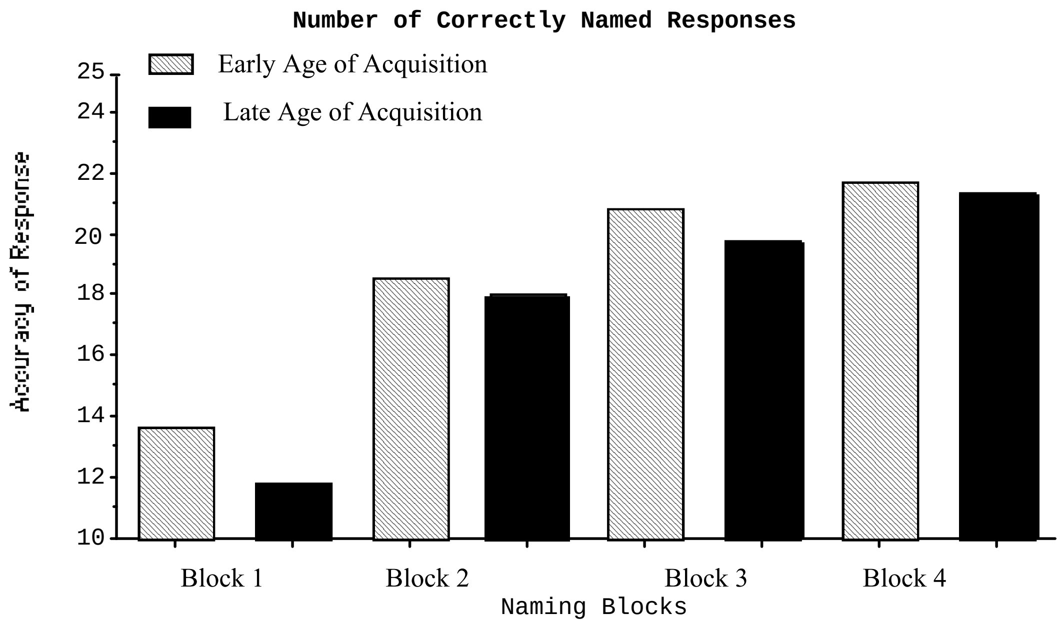

FIGURE 3: Mean accuracy scores for naming celebrity faces in Experiment 3