Dissociating word frequency and age of acquisition: The Klein effect revived (and reversed)

Stephen A. Dewhurst
Department of Psychology, Lancaster University

Christopher Barry
Department of Psychology, University of Essex

Please address correspondence to:
Stephen A. Dewhurst
Department of Psychology
Lancaster University
Lancaster LA1 4YF
UK
Phone 44 (0)1524 593835
Fax 44 (0)1524 593744
Email: s.a.dewhurst@lancaster.ac.uk
Dissociating word frequency and age of acquisition: The Klein effect revived (and reversed)

The Klein effect (Klein, 1964) refers to the finding that high-frequency words produce greater interference in a color-naming task than low-frequency words. The present study used the Klein effect to investigate the relationship between frequency and age of acquisition (AoA) by measuring their influence on color-naming. Two experiments showed reliable effects of frequency (though in the opposite direction to that reported by Klein) but no effects of AoA. Experiment 1 produced a dissociation between frequency and AoA when manipulated orthogonally. Experiment 2 found the same dissociation using different stimuli. In contrast, both variables reliably influenced word-naming. These findings are inconsistent with the view that frequency and AoA are two aspects of a single underlying mechanism.
Previous studies have shown that word frequency and age of acquisition (AoA) influence performance on a range of lexical tasks. A common finding is that frequency and AoA exert parallel effects, with high-frequency words processed more fluently than low-frequency words, and early-acquired words processed more fluently than late-acquired words, a pattern that has been observed in tasks such as lexical decision (Gerhand & Barry, 1999; Morrison & Ellis, 1995), word-naming (Gerhand & Barry, 1998; Morrison & Ellis, 2000), and object and picture naming (Barry, Morrison, & Ellis, 1997; Ellis & Morrison, 1998). Other studies have reported dissociations between frequency and AoA. For example, Roodenrys, Hulme, Alban, Ellis, and Brown (1994) found that frequency influenced short-term memory span but not speech rate, while AoA influenced speech rate but not memory span. Effects of AoA but not frequency have also been found in picture naming (Barry, Hirsh, Johnston, & Williams, 2001) and pronunciation durations (Gerhand & Barry, 1998).

Although the effects of both frequency and AoA are well established, there remains disagreement over whether frequency and AoA are separate variables or two aspects of a single underlying mechanism. The view that frequency and AoA reflect a single mechanism was supported by the findings of Ellis and Lambon Ralph (2000). They developed a connectionist model in which effects of both frequency and AoA occurred by influencing the weights between units on the different layers of the network. Specifically, early training produced larger weight changes than late training, and frequent training produced larger weight changes than infrequent training. According to Ellis and Lambon Ralph, this implies that any task that is affected by AoA will also be affected by
frequency, and vice versa. This view was echoed by Brysbaert and Ghyselinck (in press) who stated that an AoA effect is bound to occur whenever there is a frequency effect.

Researchers who contend that frequency and AoA are separate variables have attributed their effects to different stages of lexical processing. Models of reading typically locate the effects of frequency at the level of visual word recognition. For example, Morton’s (1969) logogen model assumes that high-frequency words require less stimulus information than low-frequency words in order to be activated. In some connectionist models of reading (e.g., Seidenberg & McClelland, 1989), frequency effects are accounted for by assuming that the weight structures of high-frequency words have been adjusted more often than those of low-frequency words, and are thus more accurate. Other models, such as Borowsky and Besner’s (1993) multistage activation model, locate frequency effects in the pathways that connect the word recognition system with representations in the semantic and phonological systems.

Several hypotheses have been proposed regarding the locus of AoA effects (see Johnston & Barry, in press, and Juhasz, 2005, for reviews). Brown and Watson (1987) suggested that early-acquired words have more complete phonological representations than late-acquired words, so that the phonology of the early-acquired words can be assembled more rapidly. More recently, Brysbaert and Ghyselinck (in press) proposed two separate AoA effects: a frequency-related effect and a frequency-independent effect. They suggested that the frequency-related AoA effect shares a common locus with the frequency effect, whereas the frequency-independent AoA effect occurs at either the semantic level or at the interface between the semantic and word production levels.
Despite disagreement regarding the precise mechanism underlying AoA effects (see Juhasz, 2005), there is evidence that AoA influences the production of phonological information. As noted above, effects of AoA (but not frequency) have been observed in speech rate (Roodenrys et al., 1994), picture naming (Barry et al., 2001), and pronunciation durations (Gerhand & Barry, 1998). In contrast, the effects of frequency appear to be located primarily at input processes such as visual word recognition. It should therefore be possible to dissociate the effects of frequency and AoA in a task that involves visual word identification but does not require any corresponding phonological output. In order to test this possibility, we chose a task originally used by Klein (1964) based on the Stroop task (Stroop, 1935). The standard finding in the Stroop task is that color-naming is slowed if the letter string represents a color name that is incongruent to the color of the text (e.g., the word “blue” in red text). Klein investigated color-naming latencies using a range of lexical stimuli and found that even noncolor words interfered with color-naming, relative to a baseline condition of naming the color of a row of asterisks.

From the perspective of the present study, the most intriguing of Klein’s findings was that color-naming latencies were slower for high-frequency than for low-frequency words. Klein’s procedure therefore offers an opportunity to test the view that frequency and AoA are functionally distinct. If effects of AoA occur at output, AoA should influence word-naming but not color-naming, as color-naming does not require participants to access the phonological representations of the words themselves. If effects of frequency are located at the input stages of visual word recognition, frequency should influence both word- and color-naming.
Experiment 1

Method

Participants. Sixty undergraduate volunteers in the age range 18 to 27 took part in Experiment 1. All were native English speakers and were paid for their participation.

Stimuli and Design. Experiment 1 used the 64 words from Gerhand and Barry (1998), in which frequency and AoA were varied factorially. The set consists of 16 words in each of four categories: early-acquired, high-frequency (e.g., find, cousin); early-acquired, low-frequency (e.g., rattle, greed); late-acquired, high-frequency (e.g., rate, president); and late-acquired, low-frequency (e.g., disc, marvel). Early-acquired words had AoA values (from Gilhooly & Logie, 1980) of less than 3 (acquired at 5-6 years or younger) and late-acquired words had a value of 4 or above (acquired at 8 years or older). The mean frequencies according to CELEX (Baayen, Piepenbrock & Gulikers, 1995) were 152.31 for the high-frequency words and 3.78 for the low-frequency words. The categories were matched for concreteness, imageability, word length, and type of initial phoneme. Sixteen words were presented in each of four colors (red, blue, green, and yellow). The participants’ task was either to name the word or the color in which it was presented. The dependent measures were the word-naming (or color-naming) latencies.

Procedure. Words were presented one at a time on an Apple Macintosh computer. Half the participants were instructed to read aloud each word as soon as it appeared on the screen. The other half were instructed to name the color in which each word was presented. Item presentation was terminated when the participant named the word (or its color) into a microphone. Response times (RTs) were based on the delay between stimulus onset and the registering of a response and were recorded by the presentation
software. Any hesitations or errors were noted by the experimenter and excluded from the analysis. A different random order of presentation was used for each participant.

**Results and Discussion**

Prior to statistical analysis, the RTs from the two groups were collated into separate data files and outliers were identified. Responses of over 2s and under 200 msecs were removed, followed by responses of more than 2.5 standard deviations from the resulting means. These items represented less than 2% of the total number of responses and were removed from the individual data files. Participants’ mean RTs for each of the four categories were analysed using 2x2 (Frequency x AoA) within-subjects Analysis of Variance (ANOVA) with alpha set at .05. Separate ANOVAs were conducted on word-naming and color-naming latencies. Table 1 shows the mean word-naming and color-naming latencies for each word category.

Please insert Table 1 about here

In the word-naming condition, a significant main effect of frequency showed that participants were faster at naming high-frequency words relative to low-frequency words, $F(1, 29) = 77.21, MSE = 410.41$. A significant main effect of AoA also showed that participants were faster at naming early-acquired relative to late-acquired words, $F(1, 29) = 46.02, MSE = 354.94$. There was no significant interaction between frequency and AoA, $F < 1$. In the color-naming condition, a significant main effect of frequency indicated that participants were faster at naming the colors of high-frequency relative to low-frequency words, $F(1, 29) = 17.21, MSE = 327.89$. In contrast to the word-naming condition, the main effect of AoA in color-naming was not significant, $F < 1$. The interaction between frequency and AoA was also nonsignificant, $F < 1$. 
The color-naming data were also analysed using a by-items ANOVA in which frequency and AoA were entered as unrelated measures. The results were consistent with those of the by-subjects analysis. A significant main effect of frequency showed that high-frequency words were color-named faster than low-frequency words, $F(1, 60) = 5.64, MSE = 495.45$. Neither the main effect of AoA nor the interaction with frequency were significant, $F < 1$ in both cases).

The results of Experiment 1 show a clear dissociation between frequency and AoA. Consistent with previous research (e.g., Gerhand & Barry, 1998), high-frequency words were named faster than low-frequency words, and early-acquired words were named faster than late-acquired words. In contrast, color-naming was influenced by frequency but not by AoA. This dissociation is consistent with the view that frequency and AoA have separate effects rather than arising from a common underlying mechanism. Although an effect of frequency was observed in color-naming, it was in the opposite direction to that originally reported by Klein (1964).

A potential problem with the results of Experiment 1 is that the AoA effect in word-naming was smaller in magnitude than the frequency effect. As both the frequency and AoA effects were approximately 20ms less for color-naming than for word naming, we cannot rule out the possibility that the null effect of AoA is simply due to the lower sensitivity of color-naming relative to word-naming. We therefore conducted Experiment 2 using the same procedure but a different set of stimuli, taken from Morrison and Ellis (1995). This consisted of two lists, one in which frequency was manipulated with AoA controlled, and one in which AoA was manipulated with frequency controlled. Morrison and Ellis found effects of AoA but not of frequency in word-naming (though more
recently Gerhand & Barry, 1998, found reliable effects of both variables using the same stimuli).

Experiment 2

Method

A new group of 60 undergraduates took part in either two word-naming sessions or two color-naming sessions, one featuring a manipulation of frequency and the other featuring a manipulation of AoA. Stimuli consisted of two lists of 48 words each. List 1 consisted of 24 high-frequency and 24 low-frequency words, matched for AoA. Mean CELEX scores were 182.13 per million for the high-frequency words and 5.63 per million for the low-frequency words. List 2 consisted of 24 early-acquired and 24 late-acquired words, matched for frequency. Early-acquired words had AoA values of 2.5 or less (acquired at 4 years or younger) and late-acquired words had a value of 5.0 or above (acquired at 9 years or older). Words were presented in red, blue, green, or yellow text. Participants took part in two tests, one with the frequency list and one with the AoA list.

Results and Discussion

Outliers (which comprised less than 2% of the total number of responses) were removed from the individual data files and the resulting mean RTs were analysed using related t-tests. Table 2 shows mean word-naming and color-naming latencies for each category. Reliable effects of both frequency, t(29) = 6.43, and AoA, t(29) = 6.70, were observed in word-naming. However, color-naming showed a reliable effect of frequency but not of AoA. For the frequency list, color-naming RTs were reliably faster to high-frequency words than to low-frequency words, t(29) = 4.20. In the AoA list, color-naming RTs to early- and late-acquired words did not differ reliably, t(29) = -1.35, p =
0.19. By-items analyses using unrelated t-tests confirmed these results. Color-naming RTs were reliably faster to high-frequency words than to low-frequency words, $t(46) = 2.21$, while the effect of AoA was not significant, $t(46) = -0.21$, $p = 0.84$.

Please insert Table 2 about here

The results of Experiment 2 confirm those of Experiment 1. A reliable effect of frequency was observed in color-naming with no effect of AoA, while effects of both frequency and AoA were observed in word-naming. In contrast to Experiment 1, the effect of AoA in word-naming was larger than the effect of frequency. The null effect of AoA in color-naming therefore cannot be due to a general reduction in the magnitude of the AoA and frequency effects relative to those observed in word-naming. The effect of frequency was again in the opposite direction to that reported by Klein (1964).

General Discussion

The results of two experiments support the view that word frequency and AoA are separate variables rather than two aspects of a single underlying mechanism. Experiment 1 found reliable effects of both frequency and AoA in word-naming, with high-frequency words named faster than low-frequency words, and early-acquired words named faster than late-acquired words. In contrast, color-naming was influenced by word frequency, with high-frequency words color-named faster than low-frequency words, but not by AoA. An effect of frequency but not of AoA in color-naming was also observed in Experiment 2 using a different set of stimuli. These findings are inconsistent with the view that frequency and AoA effects reflect a common underlying mechanism.

The null effect of AoA in color-naming is also at odds with the views of Ellis and Lambon Ralph (2000) and Brysbaert and Ghyselinck (in press) regarding the ubiquity of
AoA effects. However, it is possible that effects of AoA are task-specific. In other words, a task will only show an effect of AoA if the stimuli vary in the order in which they were first encountered within that task. Although the words used in the present study varied in AoA, there were no pre-experimental differences in the order in which they were color-named. A problem with this explanation is that there were no pre-experimental differences in the frequency with which the words were color-named, yet a reliable effect of frequency was observed. However, it is possible that the frequency effect is the result of processing that involves access to the orthographic lexicon. The role of orthography in the Stroop task is indicated by the finding that the Stroop effect is reduced for pseudohomophones of color names, such as “bloo” (Dennis & Newstead, 1981; see also Besner, Stolz, & Boutelier, 1997). If frequency influences visual word recognition, an effect of frequency will be found in any task in which words are processed visually, including tasks in which the recognition of the words is incidental, such as color-naming.

What do the present findings tell us about the locus of AoA effects? Juhasz (2005) concluded that, although the precise locus of the AoA effect remains uncertain, “it is obvious that AoA influences semantic processing” (p. 707). However, the combination of our results and the account of the Stroop effect proposed by Besner et al. (1997) may present a problem for this general view. Besner et al. suggested that the Stroop effect occurs because participants fall back on the “mental set” of processing words to the semantic level, despite being instructed not to do so. If this is the case, then an effect of AoA should be found in color-naming if AoA influences semantic processing. The null effect observed in the present study suggests that AoA does not necessarily influence
semantics. The present findings are more consistent with the view that AoA effects are located at the word production level.

Ellis and Lambon Ralph (2000) argued that effects of AoA should be observed in domains other than word recognition. They cited a number of studies reporting effects of AoA using stimuli other than words, such as pictures (e.g., Vitkovitch & Tyrell, 1995), faces (Moore & Valentine, 1999), and characters from soap operas (Lewis, 1999). They further suggested that AoA effects are a natural consequence in any network whenever learning is cumulative and accompanied by a decline in plasticity. We do not dispute this. However, while it is clear that effects of AoA can be observed in nonverbal tasks, the precise mechanism underlying the effect will vary from task to task. In tasks that involve word-naming, effects of AoA are likely to be located at the word production level. This does not, however, preclude the possibility that AoA effects in other tasks can be supported by different mechanisms.

We have argued that color-naming is not influenced by AoA because it does not require participants to access the phonology of the words. However, effects of AoA have been found in lexical decision, a task that does not necessarily require access to phonology. A key difference between lexical decision and color-naming is that the former requires participants to make a response based on the word itself, whereas the latter does not. It is therefore possible that AoA effects in lexical decision are located at the interface between the semantic and word production levels, a level identified by Brysbaert and Ghyselinck (in press) as a possible locus of the frequency-independent AoA effect.

In Klein’s (1964) study, color-naming latencies were faster for low-frequency than for high-frequency words. In contrast, the present study found precisely the opposite
pattern in two experiments, using two different sets of words. Specifically, we found that color-naming latencies were faster for high-frequency words than for low-frequency words. How can these contradictory results be explained? The answer probably lies in the difference between the low-frequency words used by Klein and those used in the present study. The low-frequency words in the present study, although less common than the high-frequency words, were likely to have been known by all participants. In contrast, Klein used only four low-frequency words (eft, helot, sol, and abjure) which were particularly rare and may not have been familiar to all participants. These words would therefore have been processed as nonwords by some participants (Miozzo & Caramazza, 2003). As Klein found that nonwords were color-named faster than words, processing low-frequency words as nonwords would have reduced their mean RTs.

Why does word frequency influence performance on a task that does not require word identification? Klein (1964) discussed his findings in terms of the greater attention-catching or “attensive” power of high-frequency words. Although our results were in the opposite direction to Klein’s, the notion of attensive power could explain our results, as recent findings indicate that it is low-frequency words that attract higher levels of arousal, particularly during early stages of processing (e.g., Malmberg & Nelson, 2003). However, we believe the critical factor is the duration of processing of the words in the color-naming task. As low-frequency words are processed more slowly than high-frequency words, their interference with color-naming persists for longer.

Miozzo and Caramazza (2003) also found that low-frequency words interfered more than high-frequency words in a task in which participants were not required to name the words. They investigated the picture-word interference paradigm, in which a
picture and a word are presented simultaneously at the same fixation point and the task is to name the picture. Miozzo and Caramazza found that the interference in picture naming was greater when the pictures were paired with low-frequency words. They also discussed the possibility that the effects they attributed to frequency were due to AoA, as their high- and low-frequency words differed significantly in AoA. However, the null effects of AoA in color-naming observed in the present study suggest that the effects reported by Miozzo and Caramazza were indeed due to frequency.

A possible alternative explanation of the present findings is that they are due to differences in initiating the articulation of the words. For example, the late-acquired words may have initial phonemes that trigger the response key sooner than those of the early-acquired words, leading to differences in word-naming but not color-naming. However, this is unlikely given that the frequency and AoA effects produced by both word sets used in the present study have been found to be eliminated by delayed naming. Both Barry and Gerhand (1998) and Morrison and Ellis (1995) found reliable effects of AoA in immediate naming but not when naming was delayed until the presentation of a cue approximately one second after the presentation of the target. If the frequency and AoA effects were due to differences in initiating articulation then they would be present in both immediate and delayed naming conditions.

Although our aim was to dissociate word frequency and AoA, the findings are also of relevance to theoretical explanations of the Stroop effect (see MacLeod, 1991, for a review). We have already discussed the “mental set” theory proposed by Besner et al. (1997). The present findings are easily accommodated within this account: If participants revert to processing words for meaning in the Stroop task, then typical effects of
frequency will emerge. Another influential account of the Stroop effect is the parallel distributed processing model developed by Cohen, Dunbar, and McClelland (1990). In Cohen et al.’s model, separate pathways are created for word-naming and color-naming, with a single shared response mechanism. As individual units within the model can be members of more than one pathway, interference will result when two simultaneously active pathways produce conflicting activation at their intersection. This model can accommodate the effects of word frequency in color-naming provided the amount of interference from a secondary pathway (as would be activated by the phonology of the word) is influenced by the duration of the secondary task. As discussed above, the interference produced by low-frequency words persists for longer than that produced by high-frequency words, therefore color-naming will be slower for the low-frequency words.

To summarize, the main finding from the present study was that word frequency influenced color-naming but AoA did not. In contrast, both frequency and AoA exerted reliable effects in word-naming. Previous studies have found effects of AoA but not frequency in picture naming (Barry et al., 2001), auditory lexical decision (Turner et al., 1998), and pronunciation durations (Gerhand & Barry, 1998). The findings from the present study of an effect of frequency but not of AoA completes a double dissociation that cannot be accounted for by the view that frequency and AoA represent a single underlying mechanism. Instead, the results suggest that frequency and AoA have separate effects in lexical tasks, with frequency influencing visual word recognition and AoA influencing word production.
References


Table 1

Mean Word-Naming and Color-Naming Latencies (with standard errors) as a Function of Frequency and AoA for Experiment 1

<table>
<thead>
<tr>
<th></th>
<th>High Frequency</th>
<th>Low Frequency</th>
<th>WFE</th>
<th>Mean WFE</th>
<th>Mean AoA effect</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Word-naming</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early-acquired</td>
<td>508 (9)</td>
<td>538 (10)</td>
<td>30</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td>Late-acquired</td>
<td>529 (11)</td>
<td>564 (11)</td>
<td>35</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AoA effect</td>
<td>21</td>
<td>26</td>
<td></td>
<td></td>
<td>23</td>
</tr>
<tr>
<td><strong>Color-naming</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Early-acquired</td>
<td>607 (13)</td>
<td>618 (12)</td>
<td>11</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Late-acquired</td>
<td>605 (13)</td>
<td>621 (12)</td>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AoA effect</td>
<td>-2</td>
<td>3</td>
<td></td>
<td></td>
<td>0.7</td>
</tr>
</tbody>
</table>
Table 2

Mean Word-Naming and Color-Naming Latencies (with standard errors) as a Function of Frequency and AoA for Experiment 2

<table>
<thead>
<tr>
<th></th>
<th>Word-naming</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High Frequency</td>
<td>Low Frequency</td>
<td>WFE</td>
<td>Early Acquired</td>
<td>Late Acquired</td>
</tr>
<tr>
<td></td>
<td>Frequency</td>
<td>Frequency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>513 (10)</td>
<td>542 (11)</td>
<td>29</td>
<td>528 (11)</td>
<td>562 (13)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Color-naming</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High Frequency</td>
<td>Low Frequency</td>
<td>WFE</td>
<td>Early Acquired</td>
<td>Late Acquired</td>
</tr>
<tr>
<td></td>
<td>Frequency</td>
<td>Frequency</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>604 (8)</td>
<td>620 (8)</td>
<td>16</td>
<td>642 (9)</td>
<td>636 (10)</td>
</tr>
</tbody>
</table>