

# Age of acquisition ratings for 3,000 monosyllabic words

MICHAEL J. CORTESE

*University of Nebraska, Omaha, Nebraska*

AND

MAYA M. KHANNA

*Creighton University, Omaha, Nebraska*

---

Age of acquisition (AoA) ratings made on a 1–7 scale for 3,000 monosyllabic words were obtained from 32 participants across four blocks of 750 trials (two blocks of 750 trials were completed in each of 2 days). These results, as well as those of the regression analyses and reliability and validity measures that were originally reported in Cortese and Khanna (2007), are summarized here. Here, we also report high interblock correlations across items, indicating that participants were consistent in their ratings across blocks. The norms for the 3,000 words are important for researchers interested in word processing and may be downloaded from the Psychonomic Society's Norms, Stimuli, and Data archive at [www.psychonomic.org/archive](http://www.psychonomic.org/archive).

---

Recently, much attention has been focused on the relationship between age of acquisition (AoA) and word recognition (e.g., Bonin, Barry, Méot, & Chalard, 2004; Ellis & Lambon Ralph, 2000; Monaghan & Ellis, 2002; Morrison & Ellis, 1995; Zevin & Seidenberg, 2002, 2004; for a review, see Juhasz, 2005). The present study provides age of acquisition ratings for 3,000 monosyllabic words. The primary purpose of these ratings is to provide researchers with a source of information that can be used to control, manipulate, or analyze AoA in word processing and memory studies.

Some important debates in the literature on word recognition have focused on AoA and its relationship to word frequency (e.g., Zevin & Seidenberg, 2002, 2004) and imageability (e.g., Monaghan & Ellis, 2002; Strain, Patterson, & Seidenberg, 2002). In each case, there has been uncertainty regarding the extent to which each variable relates to word recognition performance because the variables in question are correlated. In the case of word frequency, high-frequency words (e.g., *book*) are typically acquired early, and low frequency words (e.g., *boon*) are typically acquired later. Similarly, in the case of imageability, high-imageable words (e.g., *doll*) are typically acquired early, and low-imageable words (e.g., *trust*) are acquired later.

Because of the high correlations among these factors, it is difficult to design experiments that cross these variables in a factorial design. Furthermore, because the AoA ratings currently available are limited to a relatively small number of words, and the number of variables that must be controlled in standard word recognition studies is high,

it is difficult to generate enough items for there to be sufficient power to find an effect of AoA while controlling for the other factors.

One possible remedy to this problem involves collecting reaction times for large sets of words and performing regression analyses on the data (e.g., Besner & Bourassa, 1995; Spieler & Balota, 1997). However, in the multiple regression analyses of 2,428 monosyllabic words conducted by Balota and colleagues (Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004) on naming and lexical decision latencies and error rates, AoA was not assessed because of the low number of items in the corpus for which AoA ratings were available. This limitation of the Balota et al. (2004) study was one of the motivating factors for collecting AoA ratings found in the archive. In fact, we (Cortese & Khanna, 2007) reanalyzed the data originally analyzed by Balota et al. (2004). In our analyses, AoA predicted naming and lexical decision performance above and beyond the 22 predictor variables analyzed by Balota et al. (2004). This result has important implications for theories of word processing (e.g., Lambon Ralph & Ehsan, 2006), as well as theoretical conceptualizations of AoA (e.g., Brysbaert, Van Wijnendaele, & De Deyne, 2000). For example, when AoA was entered into the regression analysis of naming latencies, the formerly significant effect of imageability that was reported by Balota et al. (2004) was no longer significant.

In addition, we found a high degree of interparticipant reliability in our AoA ratings, with correlation coefficients ranging from .56 to .85. Furthermore, our AoA ratings correlated highly with other AoA norms, such as those

---

M. J. Cortese, [mcortese@mail.unomaha.edu](mailto:mcortese@mail.unomaha.edu)

---

reported by Bird, Franklin, and Howard (2001;  $r = .827$ ) and Stadthagen-Gonzalez and Davis (2006;  $r = .850$ ), despite the fact that these other norms were collected from non-American English speakers. The correlation between our norms and Morrison, Chappell, and Ellis's (1997) ratings was somewhat lower ( $r = .565$ ), but there were only 128 words in common, and Morrison et al.'s norms were collected on British English children.

Also, it is important to note that we (Cortese & Khanna, 2007) reported simple correlations between AoA and nine other factors for 2,342 of the items in our data set. We found significant correlations between AoA and length ( $r = .261$ ), orthographic neighborhood size ( $r = -.206$ ), subjective frequency ( $r = -.721$ ), objective frequency ( $r = -.689$ ), feedforward rime consistency ( $r = .080$ ), and imageability ( $r = -.381$ ). Feedforward onset consistency, feedback onset consistency, and feedback rime consistency were not significantly correlated with AoA (all  $r$ s  $< |.04|$ ). In addition, in a simultaneous multiple regression analysis in which AoA was predicted by these factors and initial phoneme factors, we found that over 72% of the variance associated with AoA was accounted for by these factors. All of the factors that were significantly correlated with AoA remained significant in this analysis, and feedforward onset consistency also accounted for unique variance in this analysis. Finally, we should note that for 2,046 of the items for which both AoA and frequency trajectory values (cf. Zevin & Seidenberg, 2004) were available, AoA and frequency trajectory were also significantly correlated ( $r = -.496$ ). However, in subsequent multiple regression analyses with frequency trajectory added as an additional factor, AoA still predicted naming and lexical decision performance above and beyond the standard predictor set.

Also, it is important for AoA ratings to be available for a large number of monosyllabic words, because most contemporary models of word recognition have been designed to simulate recognition for single-syllable words. Interestingly, some of the more prominent models, such as Coltheart and colleagues' dual-route cascaded (DRC) model (Coltheart, Rastle, Perry, Langdon, & Ziegler, 2001) and Seidenberg and colleagues' parallel-distributed processing model (e.g., Harm & Seidenberg, 2004; Plaut, McClelland, Seidenberg, & Patterson, 1996; Seidenberg & McClelland, 1989), do not posit a prominent role for AoA.

## METHOD

The methods used to obtain the ratings provided here have been described in Cortese and Khanna (2007). Here, we describe the main details of the norming procedure. For other methodological details, see Cortese and Khanna.

### Participants

Thirty-two College of Charleston undergraduates who were enrolled in a psychology course participated in the study for course credit.

### The Word Corpus

The word corpus consists of the 3,000 monosyllabic words that were rated for imageability in the Cortese and Fugett (2004) study. Balota and colleagues have made available naming and lexical deci-

sion latencies and error rates from younger and older adults for most of these words (Balota et al., 2004), and naming and lexical decision data are also accessible via the English Lexicon Project Web site (elexicon.wustl.edu). Finally, objective frequency estimates (see, e.g., Zeno, Ivens, Millard, & Duvvuri, 1995) and subjective frequency estimates (Balota, Piloti, & Cortese, 2001) are available for the majority of items.

### Procedure

A microcomputer was used for collecting ratings in two separate sessions that were held on separate days and were no more than 1 week apart. Each session took between 1.25 to 2.0 h to complete. Half of the words were rated in each session. With one exception, each set of 4 participants received the same set of items in a given block, but the order of blocks was counterbalanced across the set of participants, according to a Latin square design. Within a block, the order of items was random. This counterbalancing procedure was incomplete for one set of 4 participants. One of these participants received the items in a uniquely random order.

The instructions (presented in the Appendix) explained that different words are acquired at different times throughout our lives and that the purpose of the study was to provide AoA estimates for 3,000 words. On each trial, a word in lowercase letters was presented in the center of the screen, and at the bottom of the screen, a 1–7 scale appeared. The scale employed was originally developed by Gilhooly and Logie (1980). Ages 0–2 were associated with the value 1, ages 2–4 were associated with the value 2, and so on, up to and including 6. The value of 7 was given to words acquired at or after the age of 13. Each participant entered his/her rating via a numerical keypad on the right side of the keyboard. Reaction times were measured, but participants were not informed of this, and they responded at their own pace. Due to the number of trials, it was in the participants' best interest to maintain a quick pace. Thus, the recording of reaction times allowed us to eliminate ratings that were made prematurely (i.e., responses less than 500 msec).

## RESULTS

Reaction times and ratings were eliminated whenever a rating was made in less than 500 msec.<sup>1</sup> Our purpose in choosing this value was to eliminate overly fast responses without systematically eliminating responses for items that might be easier to rate. For example, a short word that was acquired early in life (e.g., *dog*) could be rated in less time than a long word that was acquired later in life (e.g., *stealth*). By applying a longer cutoff value (e.g., 750 msec), we might end up removing more responses to shorter, earlier acquired words than those to longer, later acquired words. This procedure would introduce systematic bias into the ratings. On the other hand, by applying a 500-msec cutoff criterion, some premature responses will probably end up in the ratings, but these responses would be less likely to produce any systematic bias. Speed was not emphasized in the instructions, so no upper limit for reaction times was set. From this modest screening criterion, 1.7% of the data were eliminated. The average latency across items was 2,136.9 msec ( $SD = 465.6$ ).

Because of the number of items rated by each individual, one might be concerned that the ratings are noisy and unreliable. To address this concern, we separated the items by the four blocks in which they were rated, and we computed interblock correlations, as well as the correlation between each block and the grand mean. The correlation matrix is presented in Table 1. After outliers were removed,

**Table 1**  
**Interblock Correlation Coefficients Across the 3,000 Items**

	Block 1	Block 2	Block 3	Block 4	Grand Mean
Block 1	1.000	.912	.909	.895	.969
Block 2		1.000	.900	.874	.960
Block 3			1.000	.890	.963
Block 4				1.000	.949
Grand mean					1.000

there were at least three ratings per item per block and no more than eight ratings per item per block. In other words, all 3,000 items went into each correlation. Even though the number of ratings per item per block was relatively small, ratings were remarkably consistent. Regarding the interblock correlations, the correlation coefficients ranged from .874 to .912. We find it interesting that, although still very high, the lowest correlation was between Block 2 and Block 4. Because there were two sessions per day, one would expect fatigue to be prominent during the second session of each day, and it did appear to be so.

## DISCUSSION

We are providing an extensive set of AoA norms in the archive, which will be of use to a variety of researchers interested in word processing. Previous work of ours (Cortese & Khanna, 2007) has shown that AoA predicts naming and lexical decision performance above and beyond the standard predictor set employed by Balota et al. (2004). In addition, we have provided useful statistics relating AoA to other factors, including length, imageability, frequency, and so on.

We end this article with a brief note about whether AoA norms can be considered true independent variables. There has been considerable discussion of this issue in the literature (Bonin et al., 2004; Cortese & Khanna, 2007; Zevin & Seidenberg, 2002, 2004). For example, Zevin and Seidenberg (2002, 2004) consider AoA to be an outcome measure. In other words, the age at which a word is acquired is affected by other variables that relate to word processing performance. Our view is that it is possible to identify these other factors that contribute to AoA, so they can be evaluated separately. If AoA affects performance above and beyond these other factors, it is difficult to argue that the time at which a word is acquired is not an independent factor related to performance. In addition, one could use similar logic to argue that word frequency and, perhaps, other variables are also outcome variables. In other words, there are sure to be reasons why some words are used more frequently than others in a language. It seems, then, that there is no obvious set of criteria we can refer to that can determine which factors should or should not be considered outcome variables.

## AUTHOR NOTE

Correspondence concerning this article should be addressed to M. J. Cortese, Department of Psychology, University of Nebraska at Omaha, 6001 Dodge St., Omaha, NE 68182 (e-mail: mcortese@mail.unomaha.edu).

## REFERENCES

- BALOTA, D. A., CORTESE, M. J., SERGENT-MARSHALL, S. D., SPIELER, D. H., & YAP, M. J. (2004). Visual word recognition for single-syllable words. *Journal of Experimental Psychology: General*, **133**, 283-316.
- BALOTA, D. A., PILOTTI, M., & CORTESE, M. J. (2001). Subjective frequency estimates for 2,938 monosyllabic words. *Memory & Cognition*, **29**, 639-647.
- BESNER, D., & BOURASSA, D. C. (1995, June). *Localist and parallel processing models of visual word recognition: A few more words*. Paper presented at the annual meeting of the Canadian Society for Brain, Behaviour and Cognitive Science, Halifax, NS.
- BIRD, H., FRANKLIN, S., & HOWARD, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verb and function words. *Behavior Research Methods, Instruments, & Computers*, **33**, 73-79.
- BONIN, P., BARRY, C., MÉOT, A., & CHALARD, M. (2004). The influence of age of acquisition in word reading and other tasks: A never ending story? *Journal of Memory & Language*, **50**, 456-476.
- BRYSBAERT, M., VAN WIJNENDAELE, I., & DE DEYNE, S. (2000). Age-of-acquisition effects in semantic processing tasks. *Acta Psychologica*, **104**, 215-226.
- COLTHEART, M., RASTLE, K., PERRY, C., LANGDON, R., & ZIEGLER, J. (2001). DRC: A dual route cascaded model of visual word recognition and reading aloud. *Psychological Review*, **108**, 204-256.
- CORTESE, M. J., & FUGETT, A. (2004). Imageability ratings for 3,000 monosyllabic words. *Behavior Research Methods, Instruments, & Computers*, **36**, 384-387.
- CORTESE, M. J., & KHANNA, M. M. (2007). Age of acquisition predicts naming and lexical-decision performance above and beyond 22 other predictor variables: An analysis of 2,342 words. *Quarterly Journal of Experimental Psychology*, **60**, 1072-1082.
- ELLIS, A. W., & LAMBON RALPH, M. A. (2000). Age of acquisition effects in adult lexical processing reflect loss of plasticity in maturing systems: Insights from connectionist networks. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **26**, 1103-1123.
- GILHOOLY, K. J., & LOGIE, R. H. (1980). Meaning-dependent ratings of imagery, age of acquisition, familiarity, and concreteness for 387 ambiguous words. *Behavior Research Methods & Instrumentation*, **12**, 428-450.
- HARM, M. W., & SEIDENBERG, M. S. (2004). Computing the meanings of words in reading: Cooperative division of labor between visual and phonological processes. *Psychological Review*, **111**, 662-720.
- JUHASZ, B. J. (2005). Age-of-acquisition effects in word and picture identification. *Psychological Bulletin*, **131**, 684-712.
- LAMBON RALPH, M. A., & EHSAN, S. (2006). Age of acquisition effects depend on the mapping between representations and the frequency of occurrence: Empirical and computational evidence. *Visual Cognition*, **13**, 928-948.
- MONAGHAN, J., & ELLIS, A. W. (2002). What exactly interacts with spelling-sound consistency in word naming? *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 183-206.
- MORRISON, C. M., CHAPPELL, T. D., & 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. (1995). Roles of word frequency and age of acquisition in word naming and lexical decision. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **21**, 116-133.
- PLAUT, D. C., MCCLELLAND, J. L., SEIDENBERG, M. S., & PATTERSON, K. (1996). Understanding normal and impaired reading: Computational principles in quasi-regular domains. *Psychological Review*, **103**, 56-115.
- SEIDENBERG, M. S., & MCCLELLAND, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, **96**, 523-568.
- SPIELER, D. H., & BALOTA, D. A. (1997). Bringing computational models of word naming down to the item level. *Psychological Science*, **8**, 411-416.
- STADTHAGEN-GONZALEZ, H., & DAVIS, C. J. (2006). The Bristol norms for age of acquisition, imageability, and familiarity. *Behavior Research Methods*, **38**, 598-605.
- STRAIN, E., PATTERSON, K., & SEIDENBERG, M. S. (2002). Theories of word naming interact with spelling-sound consistency. *Journal*

*of Experimental Psychology: Learning, Memory, & Cognition*, **28**, 207-214.

ZENO, S. M., IVENS, S. H., MILLARD, R. T., & DUVVURI, R. (1995). *The educator's word frequency guide*. Brewster, NY: Touchstone Applied Science Associates.

ZEVIN, J. D., & SEIDENBERG, M. S. (2002). Age of acquisition effects in reading and other tasks. *Journal of Memory & Language*, **47**, 1-29.

ZEVIN, J. D., & SEIDENBERG, M. S. (2004). Age-of-acquisition effects in reading aloud: Tests of cumulative frequency and frequency trajectory. *Memory & Cognition*, **32**, 31-38.

#### NOTE

1. In our initial analyses, we (Cortese & Khanna, 2007) applied a 300-msec lower end cutoff. Therefore, the AoA values reported here deviate slightly from those that were used in Cortese and Khanna.

#### ARCHIVED MATERIALS

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

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

FILE: Cortese-BRM-2008.zip

DESCRIPTION: The compressed archive file contains three files:

Cortese(2008).txt, containing norms for 3,000 monosyllabic words.

Cortese(2008).xls, containing the above information in Excel spreadsheet format.

Cortese(2008).pdf, containing the above information in PDF format.

AUTHOR'S E-MAIL ADDRESS: [mcortese@mail.unomaha.edu](mailto:mcortese@mail.unomaha.edu).

#### APPENDIX

##### Age of Acquisition Instructions (Excerpted From Cortese & Khanna, 2007)

We acquire words throughout our lives. Some words are acquired at a very early age, some are acquired later, and others fall in between. The purpose of this study is to determine the approximate age for which 3,000 single syllable words have been acquired.

On each trial of the study, you will be presented with a word. Try to estimate the age at which you acquired that word according to the scale shown at the bottom of the screen. Note that words acquired at an early age should be given a low rating, and words acquired at an older age should be given a higher rating. For any word that you do not know, you should assign a value of 7. The scale will appear below each word for easy reference.

	1	2	3	4	5	6	7
Age	0-2	2-4	4-6	6-8	8-10	10-12	13+

When making your ratings, try to be as accurate as possible, but do not spend too much time on any one word. If you have any questions, ask the experimenter now. Otherwise,

PRESS <ENTER> TO BEGIN THE STUDY

(Manuscript received September 6, 2007;  
revision accepted for publication February 12, 2008.)