

Imageability estimates for 3,000 disyllabic words

Jocelyn Schock · Michael J. Cortese · Maya M. Khanna

Published online: 25 October 2011
© Psychonomic Society, Inc. 2011

Abstract We provide imageability estimates for 3,000 disyllabic words (as [supplementary materials](#) that may be downloaded with the article from www.springerlink.com). Imageability is a widely studied lexical variable believed to influence semantic and memory processes (see, e.g., Paivio, 1971). In addition, imageability influences basic word recognition processes (Plaut, McClelland, Seidenberg, & Patterson, 1996). In fact, neuroimaging studies have suggested that reading high- and low-imageable words elicits distinct neural activation patterns for the two types e.g., Bedny & Thompson-Schill (Brain and Language 98:127–139, 2006; Graves, Binder, Desai, Conant, & Seidenberg NeuroImage 53:638–646, 2010). Despite the usefulness of this variable, imageability estimates have not been available for large sets of words. Furthermore, recent megastudies of word processing e.g., Balota et al. (Behavior Research Methods 39:445–459, 2007) have expanded the number of words that interested researchers can select according to other lexical characteristics (e.g., average naming latencies, lexical decision times, etc.). However, the dearth of imageability estimates (as well as those of other lexical characteristics) limits the items that researchers can include in their experiments. Thus, these imageability estimates for disyllabic words expand the number of words

available for investigations of word processing, which should be useful for researchers interested in the influences of imageability both as an input and as an outcome variable.

Keywords Imageability · Concreteness · Disyllabic words

The construct of *imageability* refers to how easily a particular word elicits a mental picture of a word's referent (Toglia & Battig, 1978). In memory and word processing studies, imageability can be manipulated (Binder, 2007; Davis, 2010; Strain & Herdman, 1999; Strain, Patterson, & Seidenberg, 1995, 2002), analyzed as a predictor variable in multiple regression analyses (Balota, Cortese, Sergent-Marshall, Spieler, & Yap, 2004; Cortese, Khanna, & Hacker, 2010), or controlled for (Cortese, 1998). However, to date, the availability of imageability information for multisyllabic words has been scant, at best. Thus, the use of imageability as a variable has been limited. For example, in their regression analyses of reading aloud and lexical decision performance for 6,115 multisyllabic words, Yap and Balota (2009) did not assess imageability because estimates were not available for the vast majority of the stimuli. The present work provides a set of imageability ratings and reactions times for 3,000 disyllabic words, to allow work on imageability to be extended to the study of multisyllabic words.

It is important to examine the influence of imageability on performance because of its theoretical implications. For example, Paivio (1971) has proposed that words are encoded on the basis of information obtained from the senses, and that this information can be classified as verbal and/or nonverbal. Imagery is thought to be an important aspect of the nonverbal code, and the images elicited by

Electronic supplementary material The online version of this article (doi:10.3758/s13428-011-0162-0) contains supplementary material, which is available to authorized users.

J. Schock · M. J. Cortese (✉)
Department of Psychology, University of Nebraska at Omaha,
Omaha, NE, USA
e-mail: mcortese@mail.unomaha.edu

M. M. Khanna
Creighton University,
Omaha, NE, USA

concepts are believed to serve as mediators in associations. In support of this dual-code hypothesis, highly imageable words are often better remembered than low-imageable words (Cortese et al. 2010; Rubin & Friendly, 1986).

In addition, imageability is theoretically important for computational models of word recognition. The theoretical framework of Seidenberg and McClelland (1989) proposes that fully interactive connections between orthographic, phonological, and semantic codes allow for the top-down influence of semantics on the computation of the orthographic and/or phonological code(s) in visual word recognition. In their computational model of reading aloud, Plaut, McClelland, Seidenberg, and Patterson (1996) demonstrated that damage to the semantic pathway can account for surface dyslexia. This supports the assumption that semantic information is involved in the visual word recognition process, and thus, semantic-level variables, such as imageability, should account for variance in visual word recognition performance. Consistent with this idea, Strain, Patterson, and Seidenberg (1995) found that the consistency effect (i.e., longer naming reaction times for words with inconsistent spelling-to-sound mappings, such as *pint*) was practically eliminated for high-imageable words (e.g., *comb* vs. *cliff*). However, this finding has been controversial, and Monaghan and Ellis (2002) as well as Ellis and Monaghan (2002) suggested that age of acquisition (AoA) is actually the variable that interacts with frequency and consistency (see also Strain & Herdman, 1999; Strain et al., 2002). In their analyses of 2,428 monosyllabic words, Balota et al. (2004) reported that imageability accounted for unique variance in reading aloud. However, in subsequent work in which both AoA and imageability were assessed simultaneously, Cortese and Khanna (2007) found that only AoA accounted for unique variance in reading-aloud reaction time performance.

Furthermore, it has been hypothesized that the typical lexical decision task emphasizes semantic processes because meaning differentiates words from nonwords (i.e., words have meaning but nonwords do not; see, e.g., Chumbley & Balota, 1984). In support of this idea, high-imageable words often lead to faster lexical decisions than do low-imageable words (Balota et al., 2004; Cortese & Khanna, 2007). Interestingly, James (1975) found that concreteness, a variable that is highly correlated with imageability, was related to lexical decision reaction times when pronounceable nonwords (e.g., *blark*) served as distractors, but not when nonwords that violated orthographic constraints served as distractors. This finding suggests that the relative influence of concreteness (and hence imageability) on lexical decision may depend on the level of lexical processing (e.g., semantic vs. orthographic-to-phonological mapping) required by the experimental context.

The effect of imageability on reading-aloud performance also may depend on the list context (see, e.g., Zevin & Balota, 2000).

Recently, brain-imaging studies have indicated that highly imageable words produce different patterns of neural activity than do lower imageable words during word-processing tasks (see, e.g., Graves et al., 2010). In addition, Bedny and Thompson-Schill (2006) found that activity in the left superior parietal lobule, right posterior middle temporal gyrus, left superior frontal gyrus, and left fusiform gyrus increased as imageability increased. Klaver et al. (2005) found that highly imageable words activated the P600 wave in the hippocampus more than words of low imageability did during a recognition memory task. These neuroimaging studies suggest that imageability is an influential property of words, and they encourage further study. However, these imaging studies require a large number of stimuli from each condition. Thus, in order to conduct these studies of imageability's influence on word processing, a large number of imageability estimates must be available. Graves et al. (2010) utilized six different sources when selecting their stimuli, including the estimates for 3,000 monosyllabic words collected by Cortese and Fugett (2004).

Finally, it is important to note that while the study of monosyllabic words has dominated the word recognition literature for many years, there has been a recent emphasis on the processing of multisyllabic words. Analyses of reading-aloud and lexical-decision performance for tens of thousands of words is now possible due to the completion of the English Lexicon Project (i.e., ELP; Balota et al., 2007). The ELP provides reaction time estimates in the reading-aloud and lexical-decision tasks for over 40,000 English words. While Cortese and Fugett (2004) provided imageability estimates for 3,000 monosyllabic words, imageability estimates are unavailable for most of the multisyllabic words in the ELP corpus. We note that in their analyses of the reaction time and accuracy rates of monosyllabic words, Balota et al. (2004) utilized the Cortese and Fugett estimates; however, the recent analyses on performance measures of multisyllabic words by Yap and Balota (2009) did not include imageability because estimates were not available for a large number of multisyllabic words.

The present study provides estimates for 3,000 disyllabic words that participants rated across two sessions. The procedure was very similar to that employed by Cortese and Fugett (2004). We expect that these estimates will be useful to researchers who are interested in analyzing performance from the ELP as well as those who wish to employ a large number of disyllabic words in their behavioral and neuroimaging studies of word processing.

Method

Participants

A group of 35 students enrolled in undergraduate psychology courses at the University of Nebraska Omaha ($n = 26$) and Creighton University ($n = 9$) participated for extra credit or course credit. Of the 11 participants for whom demographic information was available, 36% were male, 64 % female. They ranged from 20 to 30 years of age ($M = 23.82$, $SD = 3.28$) and had completed 14–18 years of education ($M = 16.27$, $SD = 1.11$). A total of 82% were Caucasian, 9% were Asian, and 9% other.

Stimuli

The stimuli were 3,000 disyllabic words. Item characteristics for the 3,000-word corpus are provided in Table 1. In addition, we computed correlations among imageability, frequency, length, and AoA (based on preliminary data collected on these stimuli in our lab). These correlations are presented in Table 2. The items were mainly monomorphemic, but very common multimorphemic words (e.g., *breakfast*) were also included. We started with 23,365 disyllabic words, having as our goal to narrow the list down to 3,000 words. Our previous research had indicated that, across a number of different tasks, participants could respond to about 3,000 words in less than 4 h (or two 2-h sessions). To achieve this goal, a group of 7 undergraduate research assistants inspected the large sample of disyllabic words and indicated the words that they knew; each research assistant perused a different sample of words. This reduced the sample to 15,434 words. From this reduced sample, we eliminated many multimorphemic words, but maintained some common multimorphemic words. We included a relatively large proportion of monomorphemic words because Yap and Balota's (2009) studies on multisyllabic word processing utilized only monomorphemic words. As previously noted, Yap and Balota did not assess imageability because ratings were not available for most of these words. Thus, imageability estimates for a large sample of multisyllabic monomorphemic words would allow researchers to extend the

Table 1 Item characteristics for the 3,000-disyllabic-word corpus

	Imageability	Frequency	Length	AoA
Mean	4.58	30.9	6.2	4.87
Standard deviation	1.41	142.2	1.1	0.88
Range	1.27 to 7.00	.51 to 3,631.49	3 to 11	1.76 to 6.45

Frequency = number of occurrences per million words taken from Brysbaert and New (2009)

Table 2 Correlation matrix for imageability, frequency, length, and age of acquisition (AoA)

	Imageability	Frequency	Length	AoA
Imageability		-.112**	-.036 ⁺	-.506**
Frequency			-.101**	-.294**
Length				-.112**
AoA				

Frequency = number of occurrences per million words taken from Brysbaert and New (2009). ⁺ $p = .051$. ** $p < .01$

work of Yap and Balota utilizing a large common set of stimuli. Although we ended up with 3,000 words, more words could easily have been included. Thus, the final sample of words was a representative but not an exhaustive sample of disyllabic words. In other words, there was no a priori reason for selecting these specific 3,000 words. Instead, we wanted a set of words that exhibited variability on a range of other factors. Our long-term goal was to collect normative data for as many English words as possible on key variables, such as imageability, AoA, and so forth. For example, the stimuli ranged in frequency from 0 to 71.21 occurrences per million (Brysbaert & New, 2009), and they ranged in length from 3 to 11 letters. Extending the number of words for which normative data are available would allow for further examination of how these variables relate to performance on reading aloud, lexical decision, and so on, for a large set of stimuli, as well as extending the number of stimuli that could be selected for future research endeavors. For example, the English Lexicon Project (Balota et al., 2007) expanded the number of words from which researchers can select according to certain factors (e.g., average naming latencies, lexical decision times, etc.) to 40,481 words. However, most of these words have not been normed for imageability (as well as other lexical characteristics). Thus, certain practical constraints limit the items that researchers can include in their experiments, as well as limiting the factors that can be assessed in large-scale analyses (see, e.g., Yap & Balota, 2009). By providing imageability estimates for 3,000 disyllabic words, our study would reduce these constraints.

Procedure

The procedure of Cortese and Fugett (2004) was followed as closely as possible. A microcomputer was used to collect ratings in a laboratory. Two sessions of 1.25–2 h were conducted within one week of each other. Each session comprised two blocks of 750 words each, for a total of 1,500 words. Opportunities for breaks occurred after every 375 trials. A randomized file was generated anew for each set of 4 participants, and four blocks of 750 words were

constructed from this file. Each participant from this set of 4 received the same words in each of the four blocks. However, the four blocks of trials were counterbalanced across each set of 4 participants according to a Latin square design, such that each block occurred equally often in Block 1, Block 2, and so on. Within each block of trials, the presentation order of the stimuli was random.

On each trial, a word was presented in the center of the screen and ratings were entered using the numerical keypad on the right-hand side of the keyboard. The instructions were the same as those given by Cortese and Fugett (2004), except that the example of a highly imageable word was changed from the mascot of Morehead State University to MAVERICK (the University of Nebraska–Omaha mascot) for University of Nebraska at Omaha students, and BLUEJAY (the Creighton University mascot) for Creighton University students (see the Appendix). Two other aspects of the procedure differed from the method of Cortese and Fugett. First, we assumed that responses made in less than 500 ms were made too quickly for an accurate assessment of imageability, so responses that occurred in less than 500 ms were not accepted and were followed by the message “response too fast – slow down!” which appeared at the bottom and center of the screen. The target word then reappeared after 2,000 ms to be (re)rated. The 2,000-ms delay was intended to discourage participants from responding too quickly on future trials. Second, responses that did not represent the numbers between 1 and 7 were rejected and resulted in the message “response invalid – try again” at the bottom center of the screen for 2,000 ms. The target word then reappeared after 2,000 ms to be (re)rated. In each of these situations, the second rating (as long as it took longer than 500 ms and represented a number between 1 and 7) was recorded as the estimate for that item.

Analyses, results, and discussion

The estimates reported here were compiled from 32 of the 35 participants. After we collected data from 32 participants, we screened the data in the following ways. First, the overall mean for each item was computed by averaging across the 32 participants. Second, a correlation coefficient was computed between each participant’s ratings and the overall average across the 3,000 items. From the set of correlation coefficients, a mean correlation (.578) and standard deviation (.180) were established. We removed the estimates from further analysis for 3 participants for whom the correlation coefficient was less than two standard deviations below the mean. This procedure was done to eliminate the data for participants who either were not representative of the population of interest or who were not taking the task seriously. In order to complete the counter-

balancing of blocks (see the Procedure section), these estimates were replaced by estimates from 3 new participants. Of the 32 participants’ data that were retained, the correlation between each participant’s responses and the overall mean ranged between .24 and .78.

To establish criterion-related validity, we computed the correlation between the item estimates common to both our corpus and four other norm corpora. Our estimates correlated highly with Toglia and Battig (1978, $N = 535$), $r = .88$, $p < .001$; Gilhooly and Logie (1980, $N = 433$), $r = .86$, $p < .001$; Bird, Franklin, and Howard (2001, $N = 292$), $r = .82$, $p < .001$; and Stadthagen-Gonzalez and Davis (2006, $N = 543$), $r = .86$, $p < .001$, suggesting some validity and stability of imageability across these studies and across time. Previously, Cortese and Fugett (2004) had reported a correlation of .89 between their own and Toglia and Battig’s estimates of 1,153 monosyllabic items.

Although the speed of responding was not emphasized in our instructions, it was recorded. We found that the average reaction time across items was 1,825.51 ms, with a standard deviation of 479.78. This result was slightly longer and more variable than the values reported by Cortese and Fugett (2004), who reported an average reaction time of 1,423.2 ms with a standard deviation of 208.5. These differences are probably due to the length difference in the two sets of words. Furthermore, we found a negative relationship between the imageability estimates provided and the reaction times, $r = -.58$, $p < .001$. That is, participants provided estimates more quickly for high-imageable words than for low-imageable words. We note that this relationship is between those reported by Cortese and Fugett (2004, $r = -.47$) and Paivio (1968, $r = -.70$).

Combined with the estimates from our lab for monosyllabic words (Cortese & Fugett, 2004), we have now compiled imageability estimates for 6,000 words. This corpus has many potential uses. For example, we are using the estimates for disyllabic words to determine whether imageability accounts for unique variance in performance for reading aloud, lexical decision, and recognition memory. This work extends that conducted on both monosyllabic words (e.g., Balota et al., 2004; Cortese & Khanna, 2007; Cortese et al., 2010) and multisyllabic words (e.g., Yap & Balota, 2009). For monosyllabic words, imageability accounted for unique variance in lexical decision (Cortese & Khanna, 2007) and was the strongest of eight predictor variables of recognition memory performance (as measured by hits minus false alarms; Cortese et al., 2010). In addition, since current functional neuroimaging techniques require a large number of trials, these norms (as well as the combined set) facilitate stimulus selection for neuroimaging studies (e.g., Graves et al., 2010) as well as other types of studies.

In general, the use of large databases such as the ELP has been popular in recent years. Our view is that collecting

as much normative information as possible for the words in these large corpora will prove useful. For example, the Semantic Priming Project (SPP; Hutchison et al., 2011) has amassed a corpus of reading-aloud and lexical-decision reaction times and accuracy measures for 3,000 targets in related and unrelated conditions. Previous work has demonstrated that the item characteristics of the prime and target account for unique variance in priming (see, e.g., Hutchison, Balota, Cortese, & Watson, 2008). In addition, Bleasdale (1987) found significant priming in a naming task when the primes and targets were homogeneous with regard to concreteness (e.g., *chin–nose*) or abstractness (e.g., *truth–false*), but not when they were heterogeneous (e.g., *body–soul*). As imageability correlates highly with concreteness, the present set of norms allows for these types of relationships to be further explored in the SPP and/or used for stimulus selection in future studies.

In summary, using procedures similar to those of Cortese and Fugett (2004), we have collected imageability estimates for 3,000 disyllabic words. Based on the correlation between our ratings and Toglia and Battig's (1978) ratings, these estimates appear to be valid. We expect these estimates to be of use to researchers interested in the relationships between imageability and other variables.

Appendix: The instructions used for the imageability rating task

Instructions

Words differ in their capacity to arouse mental images of things or events. Some words arouse a sensory experience, such as a mental picture or sound, very quickly and easily, whereas other words may do so only with difficulty (i.e., after a long delay) or not at all.

The purpose of this experiment is to rate a list of 3,000 words as to the ease or difficulty with which they arouse mental images. Any word that in your estimation arouses a mental image (i.e., a mental picture or sound, or other sensory experience) very quickly and easily should be given a high imagery rating (at the upper end of the numerical scale). Any word that arouses a mental image with difficulty or not at all should be given a low imagery rating (at the lower end of the numerical scale). For example, think of the word MAVERICK. MAVERICK would probably arouse an image (e.g., of the University of Nebraska at Omaha mascot) relatively easily, and so would be rated as high. RELEVANT would probably do so with difficulty and be rated as low imagery.

Because words tend to make you think of other words as associates, it is important that your ratings not be based on

this, and that you judge only the ease with which you get a mental image of an object or event in response to each word.

Your imagery ratings will be made on a 1 to 7 scale. A value of 1 will indicate a low imagery rating, and a value of 7 will indicate a high imagery rating. Values of 2 to 6 will indicate intermediate ratings.

1	2	3	4	5	6	7
Low			Medium			High

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 experiment.

Note: For Creighton University participants, MAVERICK was replaced with BLUEJAY, and “University of Nebraska at Omaha” was replaced with “Creighton University.”

References

- Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., & Yap, M. J. (2004). Visual word recognition of single-syllable words. *Journal of Experimental Psychology: General*, *133*, 283–316. doi:10.1037/0096-3445.133.2.283
- Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., et al. (2007). The english lexicon project. *Behavior Research Methods*, *39*, 445–459. doi:10.3758/BF03193014
- Bedny, M., & Thompson-Schill, S. L. (2006). Neuroanatomically separable effects of imageability and grammatical class during single-word comprehension. *Brain and Language*, *98*, 127–139. doi:10.1016/j.bandl.2006.04.008
- Binder, J. R. (2007). Effects of word imageability on semantic access: Neuroimaging studies. In J. Hart Jr. & M. A. Kraut (Eds.), *Neural basis of semantic memory* (pp. 149–181). New York: Cambridge University Press. doi:10.1017/CB09780511544965.007
- Bird, H., Franklin, S., & Howard, D. (2001). Age of acquisition and imageability ratings for a large set of words, including verbs and function words. *Behavior Research Methods, Instruments, & Computers*, *33*, 73–79. doi:10.3758/BF03195349
- Bleasdale, F. A. (1987). Concreteness-dependent associative priming: Separate lexical organization for concrete and abstract words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*, 582–594. doi:10.1037/0278-7393.13.4.582
- Brysbaert, M., & New, B. (2009). Moving beyond Kučera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods*, *41*, 977–990. doi:10.3758/BRM.41.4.977
- Chumbley, J. I., & Balota, D. A. (1984). A word's meaning affects the decision in lexical decision. *Memory & Cognition*, *12*, 590–606.
- Cortese, M. J. (1998). Revisiting serial position effects in reading. *Journal of Memory and Language*, *39*, 652–665. doi:10.1006/jmla.1998.2603
- Cortese, M. J., & Fugett, A. (2004). Imageability ratings for 3,000 monosyllabic words. *Behavior Research Methods, Instruments, & Computers*, *36*, 384–387. doi:10.3758/BF03195585

- 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. *The Quarterly Journal of Experimental Psychology*, *60*, 1072–1082. doi:10.1080/17470210701315467
- Cortese, M. J., Khanna, M. M., & Hacker, S. (2010). Recognition memory for 2,578 monosyllabic words. *Memory*, *18*, 595–609. doi:10.1080/09658211.2010.493892
- Davis, C. J. (2010). The spatial coding model of visual word identification. *Psychological Review*, *117*, 713–758. doi:10.1037/a0019738
- Ellis, A. W., & Monaghan, J. (2002). Reply to strain, Patterson, and Seidenberg (2002). *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 215–220. doi:10.1037/0278-7393.28.1.215
- Gilhooly, K. J., & Logie, R. H. (1980). Age-of-acquisition, imagery, concreteness, familiarity, and ambiguity measures for 1,944 words. *Behavior Research Methods & Instrumentation*, *12*, 395–427. doi:10.3758/BF03201693
- Graves, W. W., Binder, J. R., Desai, R. H., Conant, L. L., & Seidenberg, M. S. (2010). Neural correlates of implicit and explicit combinatorial semantic processing. *Neuro Image*, *53*, 638–646. doi:10.1016/j.neuroimage.2010.06.055
- Hutchison, K. A., Balota, D. A., Cortese, M. J., & Watson, J. M. (2008). Predicting semantic priming at the item-level. *Quarterly Journal of Experimental Psychology*, *61*, 1036–1066.
- Hutchison, K. A., Balota, D. A., Neely, J. H., Cortese, M. J., Cohen-Shikora, E., Yap, M., & Tse, C.-S. (2011, November). *The semantic priming project: A megastudy*. Paper presented at the 51st Annual Meeting of the Psychonomic Society, Seattle, WA.
- James, C. T. (1975). The role of semantic information in lexical decisions. *Journal of Experimental Psychology: Human Perception and Performance*, *1*, 130–136. doi:10.1037/0096-1523.1.2.130
- Klaver, P., Fell, J., Dietl, T., Schur, S., Schaller, C., Elger, C. E., et al. (2005). Word imageability affects the hippocampus in recognition memory. *Hippocampus*, *15*, 704–712. doi:10.1002/hip.20081
- Monaghan, J., & Ellis, A. (2002). What exactly interacts with spelling-sound consistency in word naming? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 183–206. doi:10.1037/0278-7393.28.1.183
- Paivio, A. (1968). A factor-analytic study of word attributes and verbal learning. *Journal of Verbal Learning and Verbal Behavior*, *7*, 41–49. doi:10.1016/S0022-5371(68)80161-6
- Paivio, A. (1971). *Imagery and verbal processes*. New York: Holt, Rinehart & Winston.
- Plaut, D., McClelland, J., Seidenberg, M., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, *103*, 56–115. doi:10.1037/0033-295X.103.1.56
- Rubin, D. C., & Friendly, M. (1986). Predicting which words get recalled: Measures of free recall, availability, goodness, emotionality, and pronunciability for 925 nouns. *Memory and Cognition*, *14*, 79–94. doi:10.3758/BF03209231
- Seidenberg, M., & McClelland, J. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, *96*, 523–568. doi:10.1037/0033-295X.96.4.523
- Stadthagen-Gonzalez, H., & Davis, C. J. (2006). The Bristol norms for age of acquisition, imageability, and familiarity. *Behavior Research Methods*, *38*, 598–605. doi:10.3758/BF03193891
- Strain, E., & Herdman, C. M. (1999). Imageability effects in word naming: An individual differences analysis. *Canadian Journal of Experimental Psychology*, *53*, 347–359. doi:10.1037/h0087322
- Strain, E., Patterson, K., & Seidenberg, M. S. (1995). Semantic effects in single-word naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1140–1154. doi:10.1037/0278-7393.21.5.1140
- Strain, E., Patterson, K., & Seidenberg, M. S. (2002). Theories of word naming interact with spelling-sound consistency. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 207–214. doi:10.1037/0278-7393.28.1.207
- Toglia, M. P., & Battig, W. F. (1978). *Handbook of semantic word norms*. Hillsdale, NJ: Erlbaum.
- Yap, M. J., & Balota, D. A. (2009). Visual word recognition of multisyllabic words. *Journal of Memory and Language*, *60*, 502–529. doi:10.1016/j.jml.2009.02.001
- Zevin, J. D., & Balota, D. A. (2000). Priming and attentional control of lexical and sublexical pathways during naming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 121–135. doi:10.1037/0278-7393.26.1.121