

Five down, *Absquatulated*: Crossword puzzle clues to how the mind works

Raymond S. Nickerson

Published online: 24 February 2011
© Psychonomic Society, Inc. 2011

Abstract Doing crossword puzzles is a popular pastime; no one knows how many people do them, but estimates go as high as 50 million or more in the United States alone. Success at crossword puzzles taxes several aspects of memory and cognition. The purpose of this article is to consider hints that crosswords provide and questions that they prompt regarding how the mind works. Implicated topics include word associations, lexical memory search, semantic priming, the sparseness of word space, list generation, the feeling of knowing and of not knowing, mental aging, and the crossword puzzle as a vehicle for studying cognition.

Keywords Cued recall · Familiarity and recollection · Semantic memory · Knowledge

I am addicted to crossword puzzles. I do not claim to be good at them, but only to enjoy them and to suffer withdrawal symptoms when deprived of them for more than a day or two. My purpose in this essay is to revisit a topic of long-standing interest (Nickerson, 1977) and to share some reflections about hints that the experience of trying to solve crossword puzzles can provide about how the mind works. My attention here is limited to English-language puzzles, but possibly the principles discussed would apply for other alphabetic languages as well.

R. S. Nickerson (✉)
Department of Psychology, Tufts University,
Medford, Massachusetts
e-mail: r.nickerson@tufts.edu

Clue types¹

A newcomer to crossword puzzles would note straight off that clues to target words are of two types at the most general level. There are semantic and thematic clues, on the one hand, and structural clues, on the other. Independently of this distinction, some clues are provided explicitly by the puzzle designer, and other clues are discovered as a consequence of finding some of the target words. Most of the discovered clues are structural, but there are exceptions.

Semantic clues

Declarative-knowledge clues Often semantic clues call upon general knowledge. Some such clues suffice to identify the target word precisely if the puzzle doer has the requisite knowledge. Examples are shown in Table 1.

It is hard to think of more effective elicitors of “feeling-of-knowing” and “tip-of-the-tongue” experiences than the declarative-knowledge-type clues that one encounters in crossword puzzles. Sometimes such a clue will elicit the target word immediately. When it does not, the crossword puzzle doer is likely to experience varying degrees of surety with respect to the feeling of knowing. In my own experience, it is often the case that I am not immediately able to call the target to mind, but I have a strong sense that I will be able to do so with the help of additional clues or, perhaps, just with the passage of time; which is to say, I am quite sure I “know” the target, even though I cannot produce it on demand. *Author of “The Ugly Duckling”*

¹ I use the word *clue* in preference to *cue* throughout mainly because it is commonly used with reference to crossword puzzles; however, it is intended to be more or less synonymous with *cue*, as used by researchers in the context of discussions of cued retrieval and cued recall.

Table 1 Declarative-knowledge semantic clues that identify their target words precisely

Clue	Target
<i>Hades river of forgetfulness</i>	LETHE
<i>Thigh bone</i>	FEMUR
<i>Felonious fire</i>	ARSON
<i>Author of “The Brothers Karamazov”</i>	DOSTOYEVSKY
<i>Capital of Tanzania</i>	DODOMA

would evoke that feeling for me. Equally compelling is the feeling of *not* knowing; given *Capital of Tanzania* as the clue, I would be reasonably certain that I did not know the target and would get it, if at all, only as a consequence of filling in intersecting words. More often, my degree of confidence as to whether additional clues or time will bring the target to mind is somewhere between these extremes.

In addition to declarative-knowledge semantic clues that identify their target words precisely, there are those that do not identify the target precisely, although they may narrow the possibilities to very few. Examples are shown in Table 2.

Word associates That words are associatively linked to each other to varying degrees is a very old idea in psychology (Karwoski & Schacter, 1948; Kent & Rosanoff, 1910; Woodworth, 1938). In one form of the word association task, people are asked to respond with the first word that comes to mind when they hear or read a stimulus word. Not only is this an easy task to perform, but for many stimulus words there is a remarkably high degree of agreement among the responses that different people make. Often the most frequent response to a given word is several times as frequent as the next-most-frequent response (Woodrow & Lowell, 1916; Woodworth, 1938); a common response, especially with adults, is a word’s antonym (O’Connor, 1928). This consistency is sufficient to have motivated the development of word association norms (e.g., Jenkins & Palermo, 1964; Nelson, McEvoy, & Schreiber, 1998; Toglia & Battig, 1978).

Researchers distinguish between direct (*tiger–stripes*) and indirect (*lion–[tiger]–stripes*) associations. In principle,

Table 2 Declarative-knowledge clues that identify their target words less than precisely

Clue	Target	Could be
<i>Churns up</i>	ROILS	STIRS
<i>Contempt</i>	DISDAIN	DISLIKE
<i>Stately</i>	REGAL	ROYAL
<i>Pique</i>	IRRITATION	RESENTMENT
<i>Melody</i>	ARIA	TUNE

there is no limit to the number of steps there can be in an associative chain, and when people are asked to free associate—to emit words quickly as they come to mind—a word string emitted by a single person typically wanders over a considerable semantic range. It may be difficult or impossible to discern any relationship between two words separated by a few other words in the sequence, except via the mediating connections between the intervening links. Studies of semantic priming have found evidence of priming by associates that are one or two steps removed from direct (Balota & Lorch, 1986; McNamara, 1992b; McNamara & Altarriba, 1988). Exactly how to interpret such findings is a matter of debate (McNamara, 1992a).

Words that are directly associatively linked usually are related in an apparent way. If a participant in a word association experiment consistently gave responses to stimulus words that bore no obvious relationship to them (*vegetable–pencil*; *bread–roof*; *soft–crimson*), the experimenter would wonder what was going on. This is not to suggest that such associations could not exist—presumably any two words can become associated—but only that they would be unusual. Typically, it is possible to specify the relationship involved between associated words upon inspection of them.

A remarkable aspect of the relationships on which associations seem to be based is their considerable variety. Miller (1951/1963) summarizes the situation this way:

Some responses are related to the stimulus words by contrast (“wet–dry,” “black–white,” “man–woman”). Others are similar (“blossom–flower,” “pain–hurt,” “swift–fast”). Some are subordinate to the stimulus words (“animal–dog,” “man–father”), while others are coordinate (“apple–peach,” “dog–cat,” “man–boy”), and still others are superordinate (“spinach–vegetable,” “man–male”). There are also examples of assonance (“pack–tack,” “bread–red”), of part–whole (“petal–flower,” “day–week”), of completion (“forward–march,” “black–board”), of egocentrism (“success–I must,” “lonesome–never”), of word derivatives (“run–running,” “deep–depth”), of predication (“dog–bark,” “room–dark”). (p. 179)

Crossword puzzle designers use many, if not all, of these relationships as the basis for the semantic clues they provide.

Structural clues

Number of letters in the target word This clue is contained in the structure of the puzzle. Although usually the number of puzzle cells devoted to a given word is a reliable indication of the number of letters in the target word, that is

not invariably the case. Occasionally, a square is used for a string of letters that intersecting target words have in common. For example, a single position might be used for the letter string UAR that occurs in each of two intersecting words.

Specific letters in specific positions Specific letter clues are discovered as a puzzle is partially filled in. Finding a horizontal (or vertical) target word, for example, typically reveals a letter in a specific position of each of several vertical (or horizontal) target words. From first principles, one would expect that, on average, the larger the number of letters that serve as clues for a target word of a given length, the more effective this information will be. Thus, a set of three letters is likely to be a more effective clue for a six-letter target than is a two-letter set, on average. It is necessary to say “on average” because it is easy to think of exceptions to this rule. Generally, structural information limits the range of possibilities for filling in the remaining blanks. It may be clear that a missing letter is a vowel, for example, or that it is a consonant. And although the constraining information may come from knowledge of some of the letters of the horizontal (or vertical) target, it applies to the vertical (or horizontal) target as well (Rabbitt, 1993).

What is less clear from first principles is whether, for a clue composed of a given number of letters, it makes any difference which positions within the target word these letters occupy. If one’s lexicon were organized like the standard dictionary, knowledge of the first letter of a word would be expected to be more useful than knowledge of a single letter in any other position, because this would distinguish a section of the lexicon where the wanted word was to be found from other sections where a search for it would be in vain. Clearly, mental lexicons are not organized like dictionaries; nevertheless, I strongly suspect that most crossword puzzle doers would agree that knowing the first letter of a target word is typically more helpful than knowing any other letter of the word. This intuition gets support from a well-known study by Tversky and Kahneman (1973), in which people estimated for each of the consonants K, L, N, R, and V whether it occurred more frequently in first- or third-letter position in English words. A majority of participants estimated the frequency of occurrence in first-letter position to be greater than that in third-letter position for a majority of the letters, although the reverse is true in each of these cases. Gigerenzer and Brighton (2009) argued that this subset of consonants is atypical, inasmuch as most consonants occur more often in first- than in third-letter position, which suggests that, from a broader perspective and in the absence of specific knowledge to the contrary, guessing that a consonant is more likely to occupy first-letter position than third is statistically justified.

For present purposes, the main point is that knowing one or more of the letters of a target word is useful, and how useful this knowledge is is likely to vary with the letters known and their locations within the word. Knowledge that a specific position is occupied by a specific letter limits the set of possibilities considerably, and the degree of restriction can vary depending on what the letter–position combination is. Knowledge that the first letter is J, for example, is more restricting than finding that it is D, simply because there are many more English words that begin with D than that begin with J; similarly, knowing that the word ends with Z is more restricting than knowing that it ends with E.

Let us return to the question of whether knowledge of the first letter of a target word is generally likely to be more helpful than knowledge of a letter that occupies some position other than the first. Presumably whether knowledge of the first letter is more helpful in any particular case depends, at least in part, on whether knowledge of the first letter limits the possibilities more or less than does knowledge of a letter in another position. Suppose, for example, that the target is a six-letter word, and the question is whether knowledge that the first letter is P is more helpful than knowledge that the fourth letter is K. We might expect that the answer depends, in part, on the size of the set of six-letter words that begin with P relative to the size of the set of six-letter words that have K in the fourth position. The question then becomes whether knowledge of the first letter is more helpful than knowledge of any letter not in first position when the limiting effect is the same in both cases.

I am not aware of experiments in which the effectiveness of individual letters in different positions has been studied under conditions in which the information—in the technical sense of the amount by which the uncertainty about the target is reduced by the clue(s)—has been equated for clue letters in different positions. The challenge of conducting such an experiment—controlling for artifacts—is formidable. But the crossword puzzle doer is keenly aware that knowledge of letters in specific positions in target words can vary greatly in their usefulness. Knowing that the first and last letters of a five-letter word are T and S, respectively, is helpful, but not nearly as helpful as knowing that the last two letters of a five-letter word are HT. In the latter case, what are the chances that the third letter is anything other than G? And, therefore, that the second letter is anything other than I?

Compulsive crossword puzzle doers are likely to acquire a helpful sense—not necessarily verbalizable—of bigram and trigram frequencies, as well as of other sequential statistical dependencies of English, by virtue of repeated experience with them. I suspect that they acquire, too, some useful

knowledge of word segments and their relative frequencies of occurrence, but exactly what types of segments—syllabic, phonemic, morphemic, orthographic—is a question of interest.

Researchers have sometimes used a partial-word task to study aspects of verbal memory. People are shown fragments of words, much like those encountered in partially filled-in crossword puzzles, and their task is to attempt to identify the entire words of which the fragments are shown. The task has been used to study the effects of priming on lexical access. It is easy to find instances in which the same fragment can be extracted from two or more different words: NGL, for example, occurs in the same location in GANGLIER, RINGLETS, TINGLING, and WINGLESS, among other eight-letter words. If one is primed with a strong associate of one of the words that this fragment could represent, such priming is likely to make that word more accessible—more likely to be produced as the target word given this fragment—than alternative possibilities (Tulving, Schacter, & Stark, 1982).

The partial-word task has also been considered appropriate for investigating insight on the grounds that, typically, solution words are thought of suddenly, if they are thought of at all (Metcalf & Wiebe, 1987). However, Farvolden (1991; see also Bowers, Farvolden, & Mermigis, 1995) obtained evidence that the process of target-word identification is less sudden and all-or-none than it may appear. He used four-letter fragments of seven-letter low-frequency words, and the participants' task was to give, for each fragment, either a solution word or any word that occurred to them when trying to come up with the solution word. Farvolden analyzed the incorrect responses to the items that were eventually solved correctly when the four-letter fragment was supplemented with an additional letter. He found that the incorrect responses to these fragments were associated more closely with the correct solution words than with control words, and concluded from this finding that there was enough semantic information in the fragments to activate relevant semantic information, even when there was not enough to give access to the correct solution word, and that, more generally, even the solving of insight-type verbal problems may proceed in a graded fashion.

Goldblum and Frost (1988) investigated the effectiveness of several types of three-letter clues in an experiment addressed to the question of whether there are units in the lexicon larger than the individual letter but smaller than the complete word. If there are no such units, they argued, “any three letters of a word should be just as good a retrieval clue as any other three letters situated in similar positions within the word” (p. 160). They gave the following example of four groupings of three letters that might be expected, on the assumption of no units in the lexicon

larger than a single letter, to be equivalently good retrieval clues. (The target word is given in the [Appendix](#)).

1. _ _ _ DIC _ _ _ (syllable)
2. _ _ _ ICT _ _ _ (pronounceable nonsyllable)
3. _ _ NDI _ _ _ (unpronounceable cluster)
4. _ _ N I T _ _ _ (nonadjacent letters)

If the lexicon does contain units larger than an individual letter, these clues would probably not be equally effective, and in particular, if the lexicon contains syllables but not other letter clusters, the first clue should be superior to the others. Conversely, if the clues proved to be equally effective, this could be taken as evidence that there are no (nonword) lexical units larger than the single letter.

In an experimental comparison of the effectiveness of the four kinds of clues distinguished here, Goldblum and Frost (1988) found syllabic fragments to be superior to all the other types of letter combinations, and any cluster of adjacent letters to be better than the same number of nonadjacent letters. In a second experiment, these investigators found syllabic clues to be superior to comparable morphemic-unit clues (e.g., _ _ NOT _ _ _ vs. _ _ _ TON _ _ _ as clues for MONOTONOUS). They concluded that phonological units not only play a role in word retrieval but that they are more effective than all other clues. A weakness in their study was that the syllabic clues were invariably the stressed syllables of the target words, so the phonological–morphological distinction was confounded with pronunciation stress.

Goldblum and Frost (1988) considered their results to be consistent with the assumption that word recognition is mediated, at least sometimes, by syllable recognition. Letter recognizers are connected directly to word recognizers, but also to syllable recognizers that are, in turn, connected to word recognizers and can therefore facilitate the word recognition process. The general idea of a hierarchy of pattern recognizers, with outputs of low-level feature recognizers serving as inputs to higher-order pattern recognizers, has been developed into specific models of word recognition, notable among them Pandemonium by Selfridge and Neisser (1960) and the interactive activation network model of McClelland and Rumelhart (1981).

Thematic clues

Often a puzzle has a theme that is reflected in several of its target words. Such themes can be practically anything—puns, witticisms, movie titles, names of politicians, . . . The theme may be given explicitly in the puzzle title, or it may have to be discovered.

Themes, when they are recognized as such, can be especially helpful clues, as, presumably, they are intended

to be. Consider the following example. The semantic clue for an eleven-letter target was *Star of “Stormy Weather”?*. (A question mark at the end of a clue generally is itself a clue, indicating that the target is a pun or some other type of play on words.) Upon reading the semantic clue, I made no effort to come up with a candidate target, thinking my time would be better spent working on orthogonal words, given the paucity of my knowledge of movies and movie stars.

When I returned to this clue later, several of the letters had been filled in from intersecting words. Now, in addition to the semantic clue, I had the structural information `_ _ UDE_A_N_`. I doubt that this would have brought the target to my mind, but I had also discovered that the target for *Star of “Run Silent, Run Deep”?* was ETHELWATERS and had finally realized that the puzzle’s title, *Typecasting*, was a clue to several of the longer targets, which were puns on the names of movie stars. All of this together was enough to evoke CLAUDERAINS, which turned out to be correct. Recognizing the theme made the finding of other theme targets, such as ZEROMOSTEL for *Star of “Much Ado about Nothing”?* and NATALIEWOOD for *Star of “The Petrified Forest”?*, much easier.

The semantic clue for a ten-letter word was *Vacant*. UNOCCUPIED seemed the obvious answer. However, the second, third, and fourth letters of the target word had already been identified as N, O, and U, respectively. N and O fit, but U in the fourth-letter position did not. Only after several other letters had been identified from intersecting targets did the resolution become clear. The target was UNOUPCCIED. The puzzle’s title was *Move Up*. At first this did not register as a thematic clue, and even if it had, I might not have given it the intended interpretation. With UNOUPCCIED in hand, however, its meaning became clear, because this was UNOCCUPIED with UP moved from its normal location. Recognition of the thematic clue in the title was essential to making much progress on this puzzle, inasmuch as it contained several target words in which UP had been moved. *It pinpoints the problem* called for ACUNCTURUPE, for example, and *Henson’s brood* for THEMPETUPS. Table 3 gives some examples of interpretations of semantic clues that are conditioned by puzzle themes.

The selection of puzzle themes is an art. Often the theme, even when announced, is cryptic, and discovering

its meaning in reference to the puzzle is a puzzle itself. In my experience, discovery of the connection between a theme and a puzzle is often a moment of insight during puzzle solving that greatly facilitates progress thereafter. Schulman (1996) gives many examples of extraordinarily clever and enigmatic themes that puzzle constructors have used and, more generally, provides a delightfully informative insider account of the process of puzzle construction.

The sparseness of word space

I once made a small bet with an erudite colleague that there are not more than 100 palindromic words (exclusive of proper nouns, hyphenated words, abbreviations, etc.) in English. It was a brash bet, with no better justification than the fact that I had not been able to think of as many as 100, despite considerable effort to do so. My colleague, perhaps because he found it easy to come up with a few tens of instances immediately, was quite certain that there must be many more than 100 and was confident that he would be able to demonstrate that with a little further thought. A few days later he dropped by my office to pay off the bet. He too was now of the opinion that there are probably not more than 100 such words.

I still do not know for certain whether there are as many as 100 palindromic words in English. Everyone whom I know to have tried to produce this many has failed. Perhaps there are more than 100 in the *Oxford English Dictionary* (hereafter, *OED*), but I would be greatly surprised to discover that there are many more than that. The website www.palindromelist.net/#words, which claims to have “The Biggest List of Palindromes Online,” gives only 40. It turns out that determining the number of one-word palindromes, even approximately, is not easy. One does not get far on the task before running up against the question of what to count as a word. Each of the individual letters can function as a word in context: “Beethoven’s Fifth Symphony is in the key of C minor;” “Z is the last letter of the English alphabet.” Should we count each of them as a palindromic word? Should we count *stats*, which is an abbreviation for *statistics* but appears to have been deemed a word in its own right by virtue of its widespread use? Is *racecar* one word or two? What about *testset*, or *spacecaps*?

Table 3 Interpretations of semantic clues that are conditioned by puzzle themes

Theme	Clue	Target
<i>White house moonlighters</i>	<i>President as painter</i>	JACKSONPOLLOCK
<i>Redundancies</i>	<i>Dark red rope</i>	MAROONSTRAND
<i>City slickers</i>	<i>Oregonian biologist</i>	EUGENETICIST
<i>The plot thickens</i>	<i>How to get an “exact amount”</i>	EXTORTION
<i>Capitalists</i>	<i>Two-time Pulitzer-winning playwright</i>	NASHVILLEWILLIAMS

Table 6 (in the Appendix) shows the 66 palindromic words of which I am currently aware that can be found in the 20-volume, 209,500-entry *OED*, Second Edition 1991. I would be very happy to receive additions to the list at r.nickerson@tufts.edu. Excluded are hyphenated words (*pull-up*, *tut-tut*), parts of hyphenated words (*non*), contractions (*ma'am*, *li'l*), abbreviations (*stats*), slang (*bub*), proper nouns (*Nan*, *Tet*), and all single letters except A and I. I have placed the table in the Appendix on the chance that the reader may wish to see how many palindromes he/she can generate. Casual experimentation leads me to believe that most literate adults can generate a sizeable percentage of these in a few minutes. In one such informal experiment, half of the members of a group of 12 high school graduates produced at least 28 palindromes in half an hour; the most productive person produced 37 (Nickerson, 1980). I find it interesting that people can search memory at all for words that satisfy such a criterion, and quite remarkable that they can quickly find such a respectable percentage of (presumably) all that there are.

The small number of palindromic words provides a striking illustration of the redundancy of the orthographic code that we use to represent words and of what I referred to in the heading for this section as the “sparseness of word space.” The number of possible palindromic combinations of 26 letters taken n at a time is $26^{n/2}$ when n is an even number and $26^{(n+1)/2}$ when n is odd. Thus, the number of possible five-letter palindromic combinations is 17,576. The number of possible palindromic combinations, considering all lengths from one to, say, eight letters, is 950,508; for word lengths up to ten letters, the number becomes 24,713,260. Even if there were as many as 1,000 palindromes in English, this would still represent a remarkably small fraction of the palindromic letter combinations that are possible.

That only a small percentage of possible letter combinations form words is not unique to palindromes, of course. This is true of written language as a whole. This property of language has very significant implications for crossword puzzle doers. It means that it usually is not necessary to identify more than a small fraction of the letters in a word—especially a long word—in order to identify the word uniquely, or at least to narrow the candidates to a very few. Imagine listing as many five-letter words as you can that begin with B within, say, 1 min: *bread*, *broad*, *blank*, *blink*, *black*, *brine*, *brown*, . . . Then do the same for five-letter words ending with M: *dream*, *cream*, *steam*, *scram*, *gloom*, *forum*, *alarm*, . . . In both cases, one is likely to be able to generate a fairly long list. Now make a list of five-letter words that begin with B and end with M: *broom*, *bloom*, *bream*. Can that be all there are? Probably not, but I leave it to the reader to extend the list, since I—at the moment—am unable to do so.

This exercise prompts the question of how a search of memory for a word with two or more specified letters (e.g., B and M) in specified positions (e.g., first and last) proceeds. The obvious brute-force possibility would be to search all of the words one knows that begin with B and look for those that end with M, or to search all those one knows that end with M and look for those that begin with B. It seems highly unlikely that we do that, even unconsciously.

Whatever the nature of the search process, one can often identify a word with certainty on the basis of knowledge of a relatively small fraction of its letters if one knows the positions of those letters. In general, the longer the target word, the smaller the fraction of its letters that will be necessary to make the identification. This is simply another way of expressing the fact that English is highly redundant at the level of word recognition.

Although this may be intuitively obvious to any language user who thinks about it, what may be less obvious is how great the redundancy is. Probably not more than 1 or 2 out of a million of the more than 200 billion combinations of one to eight letters will actually form a word. This fraction falls off rapidly as the length of the letter string increases. According to one casual estimate, increasing the length of the letter string by one decreases the fraction of combinations that are words by nearly one and one-half orders of magnitude; for example, whereas between 1 in 1,000 and 1 in 10,000 five-letter strings form words, only roughly 1 in 100,000 six-letter combinations do (Nickerson, 1980).

Figure 1 shows estimates of the percentages of distinct words of specified lengths in the lexicon, inferred from a corpus of 12,882,039 word tokens and approximately 96,000 word types (courtesy of Tom Landauer, Touchstone Applied Science Associates²). Given that the number of possible letter permutations increases extremely rapidly with the number of letters in a string, the ratio of the number of words of length n to the number of possible letter permutations of length n drops off precipitously with increasing n , as shown in Table 4.³ The most common word length in the corpus is seven letters; note that fewer than 2 millionths of the more than 8 billion permutations of seven letters form words.

It follows from these data that the longer a target word, the smaller the percentage of its letters that is needed to provide a basis for identifying it, on average. “On average” is a considered qualification, because there are words, even

² As provided to the University of Colorado for research purposes only.

³ The W/P ratio would be greater, of course, if based on a corpus of more than 96,000 words, but even with the largest plausible estimates of the number of words in the language, the drop-off would still be precipitous.

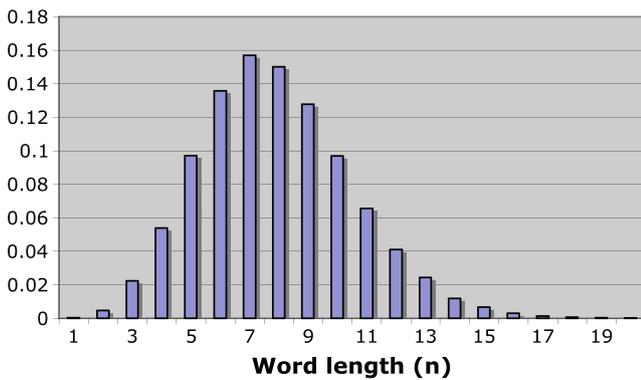


Fig. 1 The percentages of word types of specified lengths (*n*) in a corpus of 12,882,039 word tokens

long words, that differ from each other with respect to relatively few letters. We may think of all the permutations of *n* letters as a fully occupied *n*-dimensional Hamming (1950) space. (The Hamming distance between any two *n*-letter words is defined as the number of positions—first, second, third, etc.—in which the two words have different letters.) What the data in Table 4 show is that, except for very small *n*, only a very small percentage of the points in an *n*-dimensional space will represent words; the vast majority of points will represent nonword strings.

The data in Table 4 tell us that, on average, there is a considerable distance between any two words in a Hamming space. (That the Hamming distance between words tends to be large is what makes possible the development of error-detecting and error-correcting codes.) However, they do not tell us how the words are distributed—for example, whether they tend to cluster—thus leaving open the possibility that some words have near neighbors. And we know that there is such clustering, although I am not aware of any attempts to quantify this. There are games that exploit this property of words; examples include Scrabble, Anagrams, and Boggle. In another such game, which has no name of which I am

Table 4 Estimated ratio of number of words (*W*) composed of *n* letters to the number of permutations (*P*) that can be formed with *n* letters

<i>n</i>	<i>W</i>	<i>P</i>	<i>W/P</i>
1	26	26	1.0000000
2	433	676	0.6405325
3	2,123	17,576	0.1207897
4	5,154	456,976	0.0112785
5	9,285	11,881,376	0.0007815
6	12,989	308,915,776	4.2047 × 10 ⁻⁵
7	15,023	8,031,810,176	1.8704 × 10 ⁻⁶
8	14,371	2.0883 × 10 ¹¹	6.8818 × 10 ⁻⁸
9	12,234	5.4295 × 10 ¹²	2.2532 × 10 ⁻⁹
10	9,267	1.4117 × 10 ¹⁴	6.5646 × 10 ⁻¹¹

aware, players are given a word with the challenge to make a list as long as possible, such that each word in the list differs from its predecessor with respect to a single letter only; this can be played with or without the constraint that all words in the list must have the same number of letters.

Clue interactions

Puzzle doers always have more than one clue for a given target word—the semantic clue and the number of letters—at a minimum. They may have several. How do the different clues interact? Do their effects combine linearly? This seems unlikely.

Some clues are sufficiently obscure that it is doubtful whether they, by themselves, would lead a person to their target words. The following are a few examples of such clues that include puns and other types of word play: *Cheese catchers*, *Minimum ends*, *Gsge*, *Long lunches*, *V, Y, Butter*, *Run in the heat*, *Bedding down*, *A cotillion in a pastry shop*, *A little lower*, *Little colonizer in torment*, *Moving picture*. (The target words are shown in Table 7.) If one has been primed to expect these types of clues, say because of a theme indicated by a puzzle name or discovered in the course of finding target words, one may have some chance of making the connections between them. If one has not been given a reason to expect them, they are likely to be very challenging.

Often, longer targets in crossword puzzles are composed of more than one word—books or movie titles, familiar sayings, poem segments, full names. Generally such targets can be identified only as a consequence of discovering constituent letters shared with orthogonal targets. Sometimes the discovery of a small percentage of those letters will suffice to identify a target; sometimes a large percentage will be necessary. Readers may wish to try their hand at solving the following sayings on the basis of the letter clues provided. In each case, approximately two-thirds of the constituent letters were removed at random: for each letter a die was cast, and the letter was retained if the die showed either 3 or 6.

1. _ _ _ _ _ N I _ S _ E _ I N E
2. _ _ _ _ _ Y I _ E T _ H W O _ _
3. _ N H N S M _ E I _ W _ K _
4. _ O O A N _ _ _ _ P _ L H _ _ _
5. A R L I _ _ _ _ G _ H _ S _ O _ _

I suspect that most readers will not find this to be a trivially easy task. The difficulty illustrates the facilitative role that the use of spaces between words plays in printed English and other alphabetic languages. The above targets are represented as they would appear in a crossword puzzle, where between-word spaces are not used. Readers who are

stumped by any of the examples may wish to try again with knowledge of where the between-word spaces would be if the sayings were printed conventionally. The following numbers give the number of letters in each successive word in each of the five sayings: (1) 1, 6, 2, 4, 5, 4; (2) 3, 5, 4, 4, 3, 4; (3) 4, 5, 4, 5, 4; (4) 3, 4, 5, 5, 3, 5; (5) 1, 7, 5, 7, 2, 4. Another indication of the redundancy of language is the ease with which such sayings often can be completed once a single constituent word has been identified. (The sayings are given in Table 8.) Of course, puzzle designers may intentionally select targets that are not readily identified in their entirety from a knowledge of a few constituent letters. But even when this is the case, the redundancy of language is sufficiently great that one almost invariably can infer many of the letters from knowledge of what some of the others are.

Sometimes a puzzle features an unusually lengthy target that is distributed in three, four, or more parts over the puzzle area. The clues to such a target may be as unrevealing as *Start of a verse*, *Second line of verse*, *Third line of verse*, *Last line of verse*. The most skillful puzzle doer has little hope of coming up with the targets for clues of these sorts until some of the letters have been identified as a consequence of filling in intersecting targets. In this example, the verse is not a familiar one—at least it was not familiar to me—and I was unable to complete it until well over half of the letters had been found. The reader may wish to try to fill in the letters missing from the following partially completed strings. Approximately half of the letters have been supplied, the specific half having been determined by consultation of a table of random numbers (Edwards, 1957). My finding of the solution was hindered by the fact that some of the letters initially identified from intersecting vertical targets proved to be wrong. In this illustration, all of the letters provided are correct. (The answer is given in Table 9).

```
R _IS _NGFR _ _S _ORP _ _ _ _T
I _ _ _L _I _ _N _ _ _ _P _EJOK _
_ _ORHO _ _AN _ _ _ _AK _ _ _N _ _
_ _EN _ _ _A _Y _ _THE _ _ROA _
```

Clue ambiguity and garden paths

Usually the clues that one encounters in crossword puzzles are the type that would be expected to elicit the target word, given a sufficiently knowledgeable puzzle doer. Often, however, especially in more difficult puzzles, clues are used that are intended to be abstruse, or, as Schulman (1996) puts it, “to induce plausible misreadings” (p. 310). In such cases, it may be obvious that the target word, when it is found, satisfies the clue; but the clue by itself is unlikely to be a sufficient

basis for eliciting the word. An example of such an intentionally abstruse clue is *power of attorney* for the target word SIGNIFICANT. Parsing SIGNIFICANT into SIGN IF I CANT makes the match obvious. I suspect that most puzzle doers are unlikely to see this relationship in the absence of any clues beyond the original semantic one.

Many semantic clues are inherently ambiguous, even when supplemented by knowledge of the number of letters in the target item. In some cases, the ambiguity is sufficiently great that the target could not be identified uniquely by a puzzle doer with total access to a lexicon containing the entire language. *Common contraction* for a four-letter target is a case in point. *Alphabetic sequence* for a three-letter target is another. Targets for such clues can be identified uniquely only with the help of knowledge of one or more of their constituent letters gained by discovering one or more of the targets with which they intersect. Such clues can restrict the search space considerably, however, even in the absence of supplementary clues. There are only eight possibilities for a three-letter target satisfying *compass point*, for example, the first letter of which can be N, S, E, or W, the second only N or S, and the third only E or W. Furthermore, the second letter must be the same as the first, if the first is N or S, and the third the same as the first, if the first is E or W.

Other semantic clues are ambiguous in the sense that they can be interpreted in more than one way. Many words can be used as both nouns and verbs, for example, and can have nouns or verbs as synonyms. If one first interprets such a clue as a particular part of speech, one may be led down a garden path in the search for a synonymous target. More generally, most words have more than one dictionary definition; many have several. If the subset of meanings the puzzle doer considers does not contain the one that points to the target, the search again can be taken down a garden path.

The clue for a six-letter word is *Volunteers*. The word seems harder to find than it should be. Nothing suggests itself, nor do I have the feeling that the right word is lurking around ready to pop into consciousness at any moment. Even after learning that the first letter is O and the last two are RS, I am still stumped. Only after learning that the second letter is F do I realize that the desired word is OFFERS. I had interpreted *Volunteers* as a noun and had been searching for a synonymous noun. *Offers* can be a noun, but as such it is not synonymous with *Volunteers*; only as verbs do these words have similar meanings.

The clue *Rose et al.* for a five-letter word proves to be useless until I discover from orthogonal entries that the first, third, and fifth letters are P_T_S, whereupon it dawns on me that the answer is PETES (for Pete Rose and namesakes). I had been searching with a flower in mind and coming up blank.

Keep in office fails to dredge up the target for __LE_T. I keep thinking of what I do in my office with stuff I do not wish to discard or send to someone else. Upon returning to the item some time later, it is obvious that the target is REELECT. I could not say, after the fact, whether realization that *office* in the clue could refer to a political position occurred before or after REELECT popped into mind.

The clue for a five-letter word is *Target of the Pioneer*. Nothing that occurs to me fits, until I discover that the last two letters are __US; whereupon VENUS immediately surfaces and I realize, for the first time, that *Pioneer* refers to the spacecraft and not to an early settler of the American west. (I had missed the clue in the fact that *Pioneer* was capitalized.) In this instance, it seemed to me in retrospect that I became aware of VENUS *before* interpreting *Pioneer* as the name of the spacecraft, and made that connection only as a result of VENUS having come to mind. But this is little better than a guess; we do not know much about the processes involved.

Sometimes the intonation with which one reads a clue (even silently) can seem to lock a particular interpretation of an ambiguous word or phrase into place so that one fails to see that another interpretation is possible. I read *Play parts* with emphasis on the second word, as a verb–noun phrase, and failed at first to note that it could also be read, with emphasis on the first word, as a noun–noun phrase, with the first noun being used adjectivally. The target for this clue was SCENES.

Especially clever puzzle builders sometimes use semantic clues for which there are two or more plausible candidate targets that have the correct number of letters and have some letters in common. These can be problematic, because if one fixes on an incorrect possibility that fits, and especially if one gets some corroborating evidence from orthogonal targets that it is correct, the hypothesis can be difficult to dislodge. The semantic clue for a five-letter word was *Jelly fruit*, and I knew already from orthogonal words that the first and third letters were G and A, respectively. GRAPE seemed so obviously to be the answer that I immediately put it down. Only after finding it impossible to make further progress on this section of the puzzle with GRAPE in place did it occur to me to consider whether it was the only jelly fruit I could think of that would fit the G_A__ constraint. A little effort brought to mind GUAVA, which happened to be correct. Although both GRAPE and GUAVA were in my lexicon, the former was much more readily accessible than the latter, and having found one candidate that fit the constraints, I had made no attempt to find another.

All of these cases are consistent with the idea that when attempting to solve a problem that requires hypothesis testing (diagnosing a medical illness, trouble-shooting a malfunctioning piece of equipment, explaining an unex-

pected turn of events), people tend to consider only one hypothesis at a time, even if there are several that merit consideration, and persist with the hypothesis in hand unless it is shown beyond reasonable doubt to be wrong (Barrows, Freightner, Neufeld, & Norman, 1978; Bruner, Goodnow, & Austin, 1956; Elstein, Shulman, & Sprafka, 1978; Legrenzi, Girotto, & Johnson-Laird, 1993; Mynatt, Doherty, & Dragan, 1993). Evans (2007) referred to this aspect of behavior as reflective of the “singularity principle,” which is one of three that he considers descriptive of hypothetical thinking. When attempting to solve a problem that can have more than one solution, people find it easy to accept the first solution they discover and believe it to be *the* solution, failing to consider the possibility that there may be others (Nickerson, 2005).

Puzzle makers often select targets that have synonyms with the same number of letters. When the nontarget member of such a pair is the more common of the two and is more strongly associated with the clue, it can be an effective distractor. It may induce the puzzle doer not only to put the inappropriate word in the blanks but to stop searching for a better alternative. There is evidence that anagrams are more difficult to find if the letters already spell a word than if they do not (Beilin & Horn, 1962; Ekstrand & Dominowski, 1968). I hazard the guess that something similar happens with crossword puzzles, and that it is more difficult to find the correct target word if the space has been filled with an incorrect word than if it has not.

Other aspects of anagram solving are suggestive with respect to crossword puzzle doing. The time required to find solution words for anagrams, for example, was shown long ago to be influenced by such variables as the number of letters in the anagram (Kaplan & Carvellas, 1968), the degree to which the order of the letters of the anagram differs from the order of the letters of the solution word (Mayzner & Tresselt, 1958), the frequency of occurrence in the language of the solution word (Mayzner & Tresselt, 1958), and the bigram transition probabilities of the anagram (Mayzner & Tresselt, 1959, 1962; Ronning, 1965). In particular, solutions are found faster when the number of letters in the anagram is small, when the difference between the letter order of anagram and solution word is small, when the frequency of the solution word is high, when the bigram transition probabilities of the anagram are low and those of the solution word are high, and when the anagram does not spell a word.

Having an incorrect word in place in the puzzle can also impede further progress by providing misleading clues for intersecting words. Eventually, of course, the puzzle doer may be forced to reconsider this choice, because of problems encountered in filling in the orthogonal words, but the fact that one target candidate that fits the clue has

been found may decrease the effectiveness of the search for another. Following are examples of other semantic clues that have, in my experience, evoked incorrect possibilities. The first target possibility is the one that came first to mind; the second is the one that proved eventually to be correct: *Regarding* (ASTO, INRE), *Unshut* (OPEN, AJAR), *Takes nourishment* (EATS, SUPS), *Baking chamber* (OVEN, KILN), *Some speakers* (ORATORS, WOOFERS). Note that in each of the last three examples, the two possibilities not only have the right number of letters, but also have one or more letters in common in the same position(s).

There is a point to be made here about memory search strategies that not only applies to the doing of crossword puzzles, but may also have more general applicability. Usually when one finds a plausible candidate for a target word, it does not pay to spend a lot of time searching for additional candidates that fit the constraints, because usually the first one that is found is the one that is needed. This could be for either of two reasons: (1) In most cases, there is only one word in the language that fits, or (2) the one that occurs to the puzzle solver is likely to be the one that occurred to the puzzle designer, because it was considerably more accessible to both of them than the alternative possibilities. In any case, if the first candidate that one thinks of that fits the constraints is highly likely to be the one the puzzle requires, then, if one wishes to minimize total effort, it may not make sense to try hard to think of additional possibilities, except when there is compelling evidence that the first one is not going to work. Of course, if puzzle doers recognize the author of a puzzle as someone who habitually uses obscure target words and provides clues for them that are likely to evoke more accessible candidates that also fit, they may—with good reason—be less prone to settle immediately for the first candidate that comes to mind, but instead work a little harder to come up with less apparent alternatives.

Knowledge in puzzle doing

Puzzle doing is a knowledge-based activity. At least three kinds of knowledge contribute to success at it: linguistic knowledge, general knowledge, and knowledge that is relatively specific to the doing of crossword puzzles.

Linguistic knowledge

The experience of doing crossword puzzles convinces me that I have a lot of knowledge (not all completely accurate) *about* language, or, more specifically, English, that I was not aware I had. Much of this knowledge is not easily articulated, but it is readily accessed, given the necessary

evoking situation. Given, for example, the pattern B__M, I am able to say, with moderate confidence, that there are few words that fit it.

Linguistic knowledge that is useful includes semantic knowledge (knowledge of word meanings, synonyms, antonyms, and word associations), syntactic knowledge (knowledge of parts of speech, tenses, contractions, and word spellings), and statistical knowledge (knowledge of the relative probabilities of specific letters occurring in specific positions within words, and of specific letter combinations).

If one sees a Q at the beginning of a word, one can be almost certain that the next letter is U and that the one following that is a vowel. If the first letter of a word is R, the next one quite probably is not a T, or any other consonant, except perhaps H. If the final two letters of a word are NG, it is worth considering the possibility that the letter preceding N is I. If the penultimate letters are BL, CL, DL, GL, KL, PL, SL, or TL, it is a good bet that the final letter is either E or Y. If the final two letters are GH, it is highly likely that the preceding letter is either I or U.

It is not necessary that one be able to articulate such rules, or even to be aware of them at a conscious level, in order to use them. Moreover, while such rules are very useful in general, one's thinking must not be overly constrained by them; crossword puzzle designers are impishly clever at finding words that do not fit expectations based on the statistical properties of language.

One of the things one frequently does when working on a crossword puzzle is rule out the possibility of letter strings on the grounds that they are not words. This is likely to happen, for example, when most of the letters of a target word are known as a consequence of having filled in intersecting words. In such cases, it is sometimes possible to rule out an emerging target by being quite sure that a letter string (e.g., KLQZ) does not occur in English words; however, sometimes it is also possible to rule out orthographically reasonable possibilities on the grounds that they are nonwords. Doing so without consulting the dictionary would seem to require that one knows all the words in the language. But, in fact, puzzle doers do it all the time, and it is unlikely that any of them knows all the words in the language. Of course, sometimes one rules out a combination that actually is a word that one does not recognize as such, but my guess is that the frequency with which this happens is small relative to the frequency with which the combinations people rule out really are nonwords.

Perhaps this can be attributed to the sparseness of word space, as noted above, on the assumption that most orthographically reasonable letter combinations are non-

words, so the probability that an orthographically reasonable letter combination that one does not recognize as a word is not a word is relatively high, even for an individual with a limited vocabulary.

This does not account, however, for the speed with which people can make word–nonword decisions. If one made the nonword decision on the basis of randomly searching one’s lexicon for a specific entry and not finding it, the decision “nonword” would be expected to take considerably longer than the decision “word” on the average, and to be less variable with respect to time. The expectation that it would take longer follows from the fact that, assuming a random search, finding an item that is there would require checking half of the items on average, whereas determining that an item is not there would require checking all of them. The expectation of lesser variability comes from the fact that the number of items that would have to be checked in order to find a given item would vary randomly from one to the number of the entire set, whereas the items that would have to be checked to determine that a particular item was not there would invariably be the entire set.

General knowledge

Some targets are identifiable from their clues on the basis of the kind of world knowledge that people would be expected to have acquired from everyday life and the clue–target connection is simple and direct. Examples are *Threesome after Q* (RST) and *180 degrees from SSW* (NNE).

Designers of relatively challenging puzzles, like those found in the *Sunday New York Times*, like to use clues that will *not* suggest their targets immediately to the average reader and to base many of the solutions on knowledge that not everyone is likely to have. I am not sure how to characterize this knowledge, beyond saying that it is not common in the sense of being normally acquired by all, or nearly all, people of average intelligence in the course of daily living. Some of it might be called *academic* knowledge, because it is likely to be acquired as a consequence of formal education; some might be called *literary*, because it is acquired mainly by reading books; some is specialized in the sense that it is most likely to be possessed by people who are active, or at least actively interested, in a specific field, or topical area (e.g., sports, movies, astronomy, mythology, rock music). Examples include *Cleaned up Walden well* (DIDATHOREAUJOB); *Start of a best seller’s title: 1936* (GONEWITH); *Shoulder shrugger* (TRAPEZIUSMUSCLE).

While it seems likely that the more knowledge one has that relates to the relationship between a clue and its target, the better, this rule is not without exception. The semantic clue for a five-letter target was *Rodrigo Diaz de*

Vivar. I had no idea, so went on to other parts of the puzzle. I returned to this clue after discovering from an intersecting word that the third letter of the target was C. Recognizing Rodrigo Diaz de Vivar as a Spanish name, albeit one that I did not recall having encountered before, I surmised that it was the name of a well-known Spaniard, possibly a celebrity or important historical figure. C in the third-letter position was enough to bring El Cid to mind, which (as ELCID) turned out to be correct. My knowledge of Spanish history is very limited, and El Cid is one of very few names that a search of my lexicon on Spanish history would discover. This is perhaps an illustration of the point made by Gigerenzer and Goldstein (1996, 1999; Goldstein & Gigerenzer, 1999) that knowledge being greatly limited can sometimes work to the advantage of the problem solver.

Puzzle-specific knowledge

In addition to the linguistic and general world knowledge that can be brought to bear on crossword puzzles, another useful body of knowledge, about puzzles and their construction, comes from experience in doing them. This knowledge is hard to describe, but any habitual puzzle doer acquires it over time.

Certain words, especially certain short ones, appear with a much greater frequency in crossword puzzles than in the language in general. Words that I would guess fall in this category include ISIS, ORIEL, ORT, AMAH, NENE, THOLE, SLOE, and OAST (*Goddess of fertility, Bay window, Leftover, Oriental nurse, Hawaiian goose, Oar fulcrum, Wild plum, Hop-drying kiln*). Puzzle addicts are likely to have acquired quite a few such items in their lexicons, perhaps more so than people who do not do puzzles but have similar linguistic experience in other respects.

Lexical search

Memory can be searched on the basis of essentially any criterion that can serve to classify words, no matter how arbitrary or bizarre that criterion may seem to be. Readers who are willing to try their hand at thinking of words that satisfy the following criteria will have little difficulty in doing so: words that begin with FL, words that end with PT, words that have an R in the third-letter position, three-syllable words that have the stress on the second syllable, homophones (wait–weight, sale–sail), homographs (wind, lead), stress-differentiated homographs (minute, attribute), palindromes (tenet, eye), words that spell other words backward (emit–time, lever–revel), anagrams of other words (canoe–ocean,

tame–mate), and names of members of semantic categories (vegetables, composers, birds). We can also do searches on the basis of combinations of such criteria; if this were not the case, crossword puzzles would be a boring diversion.

Several investigators have asked people to generate lists of words that satisfy specific criteria, such as those mentioned above (Bousfield, 1953; Bousfield & Sedgewick, 1944; Indow, 1980; Indow & Togano, 1970; Johnson, Johnson, & Mark, 1951; Nickerson, 1980; Rundus, 1973; Shiffrin, 1970). For many criteria, the rate of word production typically drops off roughly exponentially with time. A plot of the total number of words produced as a function of time is often reasonably well fitted by the function

$$n(t) = n(\infty)(1 - e^{-\lambda t}) \quad (1)$$

where $n(t)$ is the number of words produced by time t , $n(\infty)$ is the total number that can be produced in an unlimited time, and λ is a parameter that determines the rate at which the curve approaches asymptote. Equation 1 would not be expected to be descriptive of performance when the criterion defines a well-known set of few members (e.g., months of the year) or when people are asked, and are able, to follow a linear search strategy in identifying category members. As an example of the latter case, Indow and Togano (1970) asked Japanese people to list the names of major Japanese cities starting with the northernmost city and working south; in this case, $n(t)$ was linear.

For those cases in which performance is described by Eq. 1, both $n(\infty)$ and λ vary depending on the criterion that defines the target word set and also vary for different people working with the same target sets. Moreover, plots of $n(t)$ for individual people often display departures from the smooth curve defined by Eq. 1, of the kind that would be obtained if people sometimes produced words in bursts or clusters.

Equation 1 is consistent with a very simple stochastic model of the process of finding target words. Let us assume that the “region” of search contains a total of N items, $n(\infty)$ of which would be recognized by the searcher as belonging to the target set. Now suppose that in one time unit, the searcher draws a random sample of S items from the N -item set. On the average, the number of targets, τ , contained in such a sample will be

$$\tau = (S/N) n(\infty). \quad (2)$$

If $S = 1$, then $\tau = n(\infty)/N$ is the probability that the single item sampled is a member of the target set.

Suppose that all of the drawn items are replaced before the sample for the next time unit is drawn (which is to say that sampling within a single time unit is done without

replacement, but sampling across units is done with replacement). The average number of new (previously undiscovered) targets in a one-unit time sample will be the difference between the average number of targets in that sample and the average number of old (already discovered) targets in the sample. Given that $n(t)$ represents the number of targets found by time t , the number of remaining undiscovered targets at time t is $n(\infty) - n(t)$, and the average number of new targets in a sample will be

$$\tau_{\text{new}} = (S/N)[n(\infty) - n(t)] \quad (3)$$

Such a model was proposed by Kaplan, Carvellas and Metlay (1969) to account for the performance of people who had been asked to produce as many four-letter words as they could from sets of letters varying in number from five to ten. Indow and Togano (1970) referred to this model as the *constant rate and exhaustive scanning* (CRES) model, for obvious reasons. The “constant rate” here refers to the rate at which items are inspected, not the rate at which new targets are found; the latter decreases exponentially as the total number of found targets increases and the remaining pool of as-yet-unfound targets shrinks.

I once developed a discrete-trial variation of the CRES model in which a “trial” was defined as the drawing at random of a single item from the search set (Nickerson, 1980). Sampling was assumed to be with replacement, independently of the outcome of each draw. This approach permits one to calculate the number of trials it will take, on average, to produce any specified number of targets, given search sets and target sets of specified sizes. A weakness in this model is that the time required to inspect a single potential target item—that is, to execute a trial—is not specified.

This could be inferred from curves fitted to data if one were willing to take the asymptote of such a curve as an index of the total number of targets in the searcher’s lexicon and had some independent basis for estimating the size of the total search set—the number of items in the “region” of the lexicon that is searched. There are good reasons for not taking the asymptotes of data curves as indications of the number of targets of specified types that are in one’s lexicon, and how to produce credible estimates of the total number of items that are contained in a lexical search space is not known.

One reason for not considering $n(\infty)$ to be the number of targets of a specified type in one’s lexicon is that when people are asked to list members of the same category on different occasions, they typically produce a few more words on each successive attempt (Indow & Togano, 1970). Another reason for not taking $n(\infty)$ as an index of the number of targets in one’s lexicon would be people’s

ability, after having produced all of the items from a specified category that they can, to recognize as members of that category items that they did not produce. I am not aware of formal experimental data on this question but surmise that, unless the category had very few members, people would be able to do this.

If one accepts the argument that $n(\infty)$ does not indicate the total number of targets in a searcher's lexicon, this means that people typically do not produce all of the targets that they know, even when given unlimited time to do so. This prompts two questions. First, what percentage of the targets in one's lexicon does one typically produce, and how does this depend—if it does—on the nature of the target category? Second, why does one not produce all of the targets that one's lexicon contains? These questions prompt others. What constitutes a lexical search space? How much control does one have over the portion of one's memory that is searched? From what kind of data might one infer the contents of the space that is being searched? What does it mean to have a word in one's lexicon? What is a word? We will return to the last two questions presently.

Bases for search

No one would question that it is possible to retrieve words from memory on the basis of meaning. Given a definition, one can search memory for a word or words that match it. Crossword puzzle doers know that it is also possible to retrieve words from memory strictly on the basis of structure. If, for example, I know from the filling in of intersecting words that a target word for which I am looking has the structure `_ _PL_N_ _ION`, I can search memory for words that have the specified letters in the indicated positions without reference to meaning at all. How such a search is conducted is not at all clear.

The letter combination GH is an interesting one, especially as it occurs at the end of English words. It almost always follows one of a few vowels or vowel combinations: I, EI, OU, AU. There are several instances of most of these combinations, including the following examples: NIGH, THIGH, SLEIGH, WEIGH, DOUGH, BOUGH, and COUGH. (I am aware of only one common instance in which terminal GH follows AU; can you think of it?) The particular end-word combination OUGH has a remarkable variety of pronunciations—to wit, BOUGH, DOUGH, THROUGH, TOUGH, COUGH, TROUGH (which can be pronounced either as “trof” or “troth”), and HICCOUGH.

This is interesting because it permits a distinction between orthographic and phonetic similarities. Words ending in OUGH are more similar orthographically to each other than they are to words ending in IGH or EIGH, but

they fall into a variety of phonetically-defined categories. Suppose that one is given the task of listing as many words as one can that end in GH. Will the resulting lists show clustering in terms of phonetic properties? Orthographic properties? Both?

My conjecture is that lists produced by people given such a task would show clustering in terms of both phonetic and orthographic properties. I would expect whether the GH is silent or pronounced as /f/ to be a major, but not the only, determinant of clustering. I would expect to see COUGH and TOUGH in the same cluster, or BOUGH and DOUGH, more often than COUGH and BOUGH, or TOUGH and DOUGH. I do not know how I would bet on the question of which two of the following three are most likely to appear together: THOUGH, ROUGH, and WEIGH. THOUGH and WEIGH have the common phonetic feature of a silent GH, whereas THOUGH and ROUGH have much in common orthographically.

A study focused on phonetic or orthographic clustering of retrieved words that was intended to exploit the fact that GH is sometimes, but not always, silent would have a considerably larger population of target words with which to work if the task were to produce words that contained the GH combination within them, but not necessarily in the final two positions. One gains here several more categories of words that contain silent GH but that differ in other interesting ways. OUGHT, BOUGHT, THOUGHT, NAUGHT, FRAUGHT, and TAUGHT, for example, are quite similar phonetically but fall into two obvious categories orthographically. How might one expect the following words to cluster: WEIGHT, FREIGHT, HEIGHT, SLEIGHT, NIGHT, and FLIGHT?

An argument can be made that although we can search our lexicons on the basis of either phonological or orthographic features of words, for most of us a phonological search is the more natural one. We might expect this to be the case simply on the basis of the fact that children with normal hearing and vocal potential invariably become competent users of oral language long before they learn to read. Some people never learn to read, but presumably they can produce words that have specified sound patterns—rhymes with “red,” begins with an “ess” sound, ends with “ing”.

Here is an informal experiment that relates to this point. Try to think as quickly as you can of a four-letter word that ends with ANY. I suspect that most readers will think of one very quickly, without being aware of conducting any systematic search. Now try to think as quickly as you can of a four-letter word that ends with INY. Of one that ends with ONY. Of one that ends with UNY. Of one that ends with ENY. Is there a word in each of these cases?

Did any of them give you trouble? Did you find yourself resorting to a letter-by-letter search in any cases—AINY,

BINY, CINY, DINY, . . . ? My guess is that, in most cases, a word came to mind quickly and you did not have to do a systematic search, at least at a conscious level. It is a safe bet, however, that ENY proved to be more difficult than the others for many readers; you may have come to the conclusion, after doing a letter-by-letter search, that there is no four-letter word ending with these letters.

What makes ENY a less effective clue than the other letter combinations? Consider the words that match the other clues (MANY, ZANY, TINY, BONY, PONY, PUNY). In all cases, stress is on the first syllable, and the Y has the short-vowel pronunciation; and this is true not only of the word but of the way the three-letter clue would be pronounced by itself. This means that if one tries to find a word that sounds like—rhymes with, has the same stress pattern as—the clue, one is likely to succeed. (Note that the sound match is better in some cases than in others—MANY matches the usual way of pronouncing ANY better than does ZANY, for example, but the stress pattern matches in both cases.)

ENY differs from the other clues in that the only common four-letter word that ends in these three letters has a different pronunciation—stress on the second syllable and a long-vowel pronunciation of Y. This probably is not the way most of us would pronounce ENY, so this letter combination does not serve as an effective clue for a phonological search. If the search were strictly visual, it should be as effective as all of the others; the word it clues is not a rare one.

Among the many bases for a search of one's lexicon, none is more interesting, in my view, than the word or concept that links two ostensibly unrelated words. The clue *Kind of license or justice* illustrates the case. Provided also with the knowledge that the target word has six letters, most puzzle doers, I am guessing, would turn up POETIC fairly quickly. Table 5 gives a few more examples of word or concept pairs of the sort that one is likely to see as crossword puzzle clues. (The targets for these clues are shown in Table 10 in the Appendix.)

Table 5 Clue pairs

-
- Word with *space* or *limits*
-
- Word with *blond* or *Wednesday*
 - Prefix for *scope* and *gram*
 - Suffix for *right* and *court*
 - *Lima* or *Orson*
 - Suffix with *fraud* or *flat*
 - *Slug* or *song* follower
 - Words after *loose as* or *silly as*
 - Kind of *song* or *park*
-

How might dual pointers work? An obvious possibility is that each of them identifies a set of candidates independently and one searches the two sets looking for a common item. For example, if one were asked to think of four-letter prefixes for *scope*, one might come up with PERI, GYRO, TELE, and HORO. The same request with respect to *gram* might produce MONO, TELE, KILO, and SONO. With both sets in hand, a quick scan reveals the common item.

A question of some interest is whether the process of retrieving items that satisfy one of the clues is influenced by the fact that one is searching for an item that fits two clues instead of only that one. Is the process that finds possible prefixes for *scope* affected by the fact that one wants a result that could also be a prefix for *gram*? My feeling is that the answer is yes. I am not aware of compelling empirical evidence on the question, but one can imagine an experiment in which some participants generate words (or parts thereof) suggested by single clues, and others generate words (or parts thereof) suggested by dual clues. The time required to produce specific words is taken in both cases, and the question of interest is whether the dual clues produce the words of interest in less time than would be predicted from the times taken to produce them in response to the single clues, appropriately combined.

An experiment that bears some resemblance to this imagined one, except that it deals with recently learned associations, was reported by McLeod, Williams, and Broadbent (1971). Participants learned two lists of paired associates in which the responses were the same on both lists, but the stimuli differed. Thus, two stimuli were paired with each response. When the two stimuli for a given response were presented simultaneously, recall of the response was more likely than would have been expected from performance with the stimuli presented separately. This finding, among others, has been taken as evidence that the effect of simultaneously activating two pointers to the same response is greater than the sum of the effects of activating each alone (Baron, 1985).

At least one test of creativity, the Remote Associates Test (Mednick, 1962), centers on the ease with which people can make remote word associations. Each item in the test is composed of three words that are not directly related in any obvious way. The test-taker's task is to find a fourth word that is closely associated with all three of the not-obviously-associated words. People who do well at the task are said to have relatively flat associative hierarchies—it is not much more difficult for them to call up a remote association to a stimulus word than to call up a close associate. Those who do poorly on the test are said to have relatively steep associative hierarchies—remote associates come to mind much more slowly for them than do close associates.

Strategies in target search

My own experience with crossword puzzles leads me to distinguish three types of search for a target word. The distinction is not a sharp one, inasmuch as the three types shade into each other, but the distinction may be conceptually useful, nonetheless. The first type of search seems hardly like a search at all: One looks at the semantic clue and the number of letters required and waits, as it were, for the target word to pop into mind. The second type of search seems, introspectively, like a search. I am not only aware of trying to generate words for consideration, but candidates come readily to mind. These words typically fit the semantic clue but may be rejected because they are not consistent with the other constraints (number of letters or known letters in specific positions).

This distinction is similar to the one that Indow (1980) makes in the context of a discussion of list generation tasks. He notes that when people try to generate names of members of familiar natural categories (e.g., flowers, animals), they do so with little effort or awareness of a search that involves consideration and rejection of possibilities that do not qualify for category membership: “usually it is not necessary to conceive of any irrelevant words in order to make a relevant word available. Relevant words seem to pop up one by one directly” (p. 624). In contrast, when the target category is arbitrarily defined and difficult (one example Indow gives is Japanese nouns with a specified ending sound), one is more likely to be aware of consciously thinking of several words in order to find one that fits the criterion. Indow refers to these two cases as *direct* and *indirect* retrieval, respectively. He does not argue that all arbitrarily defined categories evoke indirect retrieval, but only those that are difficult.

I made a two-way distinction similar to Indow’s in a discussion of several list generation tasks.

In subjects’ reports of how they perform list-generation tasks, there is often the suggestion of a dual-mode retrieval process: a relatively passive mode in which one waits for possibilities to come to mind, and an active mode in which one consciously attempts to “find” possibilities. It appears that subjects often use the passive mode until it no longer produces, and then switch to the second, more structured mode. (Nickerson, 1980, p. 117)

The idea that people process information in two distinctly different ways has many proponents among cognitive psychologists. The two types are referred to variously as *intuitive* (or *heuristic*) and *analytic*, or simply *Type 1* and *Type 2*, or *System 1* and *System 2* (Beller & Kuhn, 2007; Evans & Over, 2004; Hammond, 1978; Reyna, 2004; Sloman, 2002; Wason & Evans, 1975). The

first type of process is described as preconscious, fast, automatic, heuristic, and pragmatic, and the second as conscious, slow, deliberate, analytic, and abstract. I suspect that most crossword puzzle doers would find this distinction meaningful.

The third type of search that I wish to distinguish relative to the doing of crossword puzzles is perhaps appropriately considered an extreme instance of the second type, and may be characterized as “grasping at straws.” In this case I use clues, including indirect and tentatively inferred clues, in a desperate attempt to find candidates that, if they are in my lexicon at all, are proving to be very difficult to access. Sometimes the desperation is sufficiently great to evoke mechanically stepping through some set of possibilities. Trying every letter in every unfilled position is usually practically feasible only when all but one or two of the letters of a target word have already been discovered; however, sometimes it can be useful to do a letter-by-letter search for a single position, even when several other positions are still blank. There are also situations in which enough is known to narrow the set of possibilities for a particular position to, say, a vowel, or to one of a subset of consonants. Methodical searches of the type just described are frowned upon by serious puzzlers: “A systematic search through a problem space may be the first refuge of a simulation program, but it is the last resort of the expert: no puzzler will be methodical if he can help it” (Schulman, 1996, p. 300).

At the most general level, the strategy in both the second and third types of search might be described as “generate and test,” a general search strategy commonly noted in the computer science and artificial intelligence literatures. The objective is to generate hypothesized solutions (candidate words) and to test them against known constraints. But this is not very revealing. The interesting question is, What determines the hypotheses that are generated? What guides the search for candidate words? Many strategies that puzzle doers can use can be identified at a level of specificity somewhat greater than that of generate and test. I will mention some of them here, but I suspect there are many more.

If the target word is believed to be a verb in the past tense, there is a reasonable chance that its last two letters will be ED. One may then hypothesize that the target word ends in ED and see if this helps find the orthogonal word that contains the hypothesized E or the one containing the hypothesized D. If the clue is a present participle or gerund (ends in ING), one may guess that the target word is of the same class, tentatively consider ING to be its final three letters, and see whether this helps find any of the intersecting target words. If the clue suggests a third-person singular present-tense verb, the target is likely to end with S. Examples could be multiplied.

That puzzle doers use strategies and are aware of doing so is beyond doubt; when asked, they report doing so (Hambrick, Salthouse, & Meinz, 1999). However, there are many questions regarding strategies and their use. How effective are specific strategies? How important are specific strategies relative to vocabulary and general knowledge? It seems a safe bet that puzzle doers develop increasingly effective strategies and become more proficient in strategy use with experience in puzzle doing. This represents one way in which effective puzzle doing is knowledge dependent; in this instance, strategic knowledge is the specific type involved.

The feeling of knowing—and of not knowing

Crossword puzzle doers are very familiar with the feeling of knowing, and with the feeling of not knowing. It is not unusual, in my experience at least, to be unable to think of a target word and, at the same time, to be very confident that the word is in one's lexicon and will come to mind in time. My inclination, in this situation, is to attempt to find one or more of the target words that intersect with the one I cannot access, in the belief that identification of one or more of the letters of the elusive word will bring it to mind. If that is not possible, I may simply leave the word and work on other parts of the puzzle, with the intention of coming back to it later for a fresh, and perhaps more productive, look.

The feeling of knowing is not an either-or state of mind. How difficult one expects it to be to access a word that one feels one knows can vary over a considerable range. How long I am apt to spend trying to find an elusive, but believed known, word before moving on to other parts of a puzzle depends on how hard I think it will be to access the target without the help of additional clues—that is, how close to the “tip of the tongue” I think it is. If it seems to be close, I will work at it; if it seems to be far away, I will move on and come back to it later. Sometimes I am confident that I do not know the target at all, in which case I see little point in trying to think of it.

When people are asked general-knowledge questions of varying difficulty, how long it takes them to respond, either with what they think to be the answer to a question or an indication that they cannot produce it (“I don't know,” “I can't remember”), appears to depend not only on whether what they strongly believe to be the answer comes quickly to mind but, if it does not, on the likelihood they attach to being able to come up with the answer if they keep trying. Smith and Clark (1993) found a positive correlation between the feeling of knowing and the time people took before giving up on questions they

could not answer; more generally, they found that, when people were able to answer a question, the higher the confidence in the answer, the more quickly it was produced, whereas when they could not produce an answer, the stronger the feeling of knowing, the longer they took before giving up. This makes intuitive sense. As Smith and Clark pointed out, “[t]hey should only continue [searching] as long as they believe they might retrieve an acceptable answer” (p. 27).

Not only does one's feeling of knowing vary when one cannot come up with a target to satisfy a clue or set of clues, but when candidate items come to mind, they can evoke different degrees of confidence that they are correct. At one extreme are those candidates that one feels sure are correct as soon as one thinks of them. At the other are instances that feel like little more than wild guesses. And all possible gradations lie between these extremes. It often happens that one thinks of a word that one recognizes as a plausible possibility but that one is not sure enough to write down (at least with a pen) until getting some corroborating evidence from orthogonal words.

The feeling of *not* knowing can take the form of believing that one would recognize a target as correct if it were given, but that one will be unable to produce it oneself. One may also feel that one would not even recognize a target as correct if one saw it. I find it embarrassingly easy to produce a long list of clues that have left me with the latter feeling. I knew, for example, that I did not know the target for *Absquatulated*; the clue definitely was not in my lexicon. (It is not even in my dog-eared Webster's, but it is in the *OED*.) I guessed, however, with a bit more than middling confidence, that it was a past-tense verb. The target was a four-letter word, and I discovered from filling in orthogonal words that its last two letters were *_ _ED*. On the assumption that the conjecture about the target being a past tense verb was correct, the range of possibilities had now been narrowed sufficiently that it was reasonable to begin considering possibilities on a trial-and-error basis: SPED, BLED, PLED, TIED, LIED, VIED, . . . This strategy did not work in this case, however, because the clue was so completely foreign to me that I realized I would not recognize the answer, which happens to be FLED, even if I stumbled upon it. The example illustrates that the inability to recognize a correct item as correct does not imply an inability to identify an incorrect item as incorrect; incorrect items sometimes can be identified as such on the basis of violations of linguistic rules. The assumption that *absquatulated* is a past-tense verb, if correct, rules out any candidate for *_ _ED* (SLED, DEED, FEED, HEED, NEED, . . .) that is not a past-tense verb.

As numerous studies have shown, when people feel they have knowledge in memory that they cannot retrieve, the

strength of this feeling is a reasonably good indication of the probability that they will be able to recall it eventually or to recognize what they cannot produce (Blake, 1973; Read & Bruce, 1982; Smith & Clark, 1993), or even to produce it with the help of additional retrieval clues, such as the first letter of the sought-for word (Gruneberg & Monks, 1974).

Semantic priming

The clue for a six-letter word was *Former Dolphins quarterback*, and from words already filled in I believed the fourth and sixth letters both to be E. Nothing came to mind, and I did not have a strong feeling of knowing the answer. I did not finish the puzzle, but went off to other pursuits. Several days later, the name GRIESE came, uninvited, to mind. I was not thinking about the puzzle at the time, and have no recollection of ever consciously trying to think of the name of the former Dolphins quarterback after my brief attempt when working on the puzzle. Only after the name came to mind did I recall that I had tried unsuccessfully to think of it several days before.

This experience of having the target of a memory search pop into mind days after having tried and failed to find it is not uncommon. Another instance in my experience involves an attempt, already mentioned, to list palindromic words. For many days after trying to write as many one-word palindromes as I could think of, other such words would spontaneously present themselves. Often I could not be sure, without checking, whether a word that came to mind was already on my list—sometimes it was, and sometimes it was not. One instance stands out in my memory, now several years after the fact. I was jogging early in the morning, not thinking about palindromes, and suddenly in my head was the word REPAPER, large as life, and it was not on my list. The structure of this palindrome—RE . . . ER—led me to wonder whether there might be others that begin with RE and end with ER. A little thought brought RELEVELER to mind (one who makes things level again) but, alas, LEVELLER has adjacent Ls, so it does not work.

I suspect that most readers will have had similar experiences, often, perhaps, involving the later emergence of a name that could not be recalled when sought. My most recent such experience involved an anagram. My wife and I stopped for dinner in a small restaurant in Maine that had paper placemats featuring ads from local businesses and a variety of puzzles to occupy guests while waiting for their orders. Among the puzzles were several anagrams, one of which—*tiperacu*—stumped me. When the food arrived, I put the puzzles away to get on

with the main purpose of being there. An hour or so after leaving the restaurant, the solution popped into mind when I was not consciously thinking about it. (The solution appears at the end of the [Appendix](#).) Such experiences lend credence to the idea that the mind continues to work on problems below the level of consciousness after one has given up focused efforts to solve them. Some readers may see other support for this idea in the experience of having an insight regarding how to solve a problem only some time after having failed in a focused attempt to find a solution and having walked away from the problem to concentrate on other things.

Many examples can be drawn from science and mathematics of people who report having suddenly realized the solution to a problem on which they had been working intensely but unsuccessfully for a long time. This phenomenon is what led Graham Wallas (1926/1945) to distinguish several phases of creative problem solving, one of which is a period of “incubation,” during which one’s mind continues to work on a problem below the level of awareness. Sometimes such insights appear to have been facilitated by events or thoughts that relate to the problem in some analogical or metaphorical way. Friedrich Kekule’s dream of a snake swallowing its tail, which provided him the clue to the structure of the benzene ring, is a famous—if disputed—case in point. Dmitri Mendeleev had the insight that finally yielded his periodic table of elements in a dream, after exhausting himself by working on the problem in the waking state. It is claimed that his insight was facilitated by his recognition of the similarity of the task of arranging the elements in a table in such a way as to reveal important relationships among them and the card game Patience (a form of solitaire) that he liked to play (Strathern, 2000).

The list of examples of insights that have occurred to scientists and mathematicians regarding solutions to problems on which they have spent considerable time and effort, but on which they are not consciously working when the insight occurs, could easily be extended. Examples involving Henri Poincaré, Carl Frederick Gauss, William Hamilton, Alan Turing, Paul Halmos and Andrew Wiles are described briefly in Nickerson (2010, chap. 6). Undoubtedly, similar examples could be noted in other contexts as well. And crossword puzzle doers know from experience that a similar phenomenon occurs, if on a more pedestrian level, with garden-variety folk.

Some theoretical questions and conjectures

The experience of doing crossword puzzles, and playing related word games, prompts a variety of questions and

conjectures about memory search and about how the mind works more generally. The following few, some of which have already been mentioned directly or indirectly, come readily to mind.

When there are two or more clues, can search be guided by more than one of them at the same time?

Can one search simultaneously on two or more clues of the same type? On two or more clues of different types? When searching for a five-letter word that means X, does the search process consider only five-letter words, looking for one that means X; or does it consider all words that satisfy the semantic clue, while looking for one that has five letters; or is it guided by both clues simultaneously?

Clearly, it must be assumed that lexical searches can be “localized,” in the sense that they do not all go through the entire lexicon. How the “to-be-searched” locale is delimited is a question that remains to be answered.

When a clue has more than one meaning, can memory be searched with respect to more than one meaning simultaneously?

Missing a word because of searching on the wrong part of speech is a common problem in my experience. This suggests that one does not search one’s lexicon, at least consciously, for words that have the same meaning as, say, *pitch*, but for words having the same meaning as *pitch* when used as a noun, or for those having the same meaning as *pitch* when used as a verb. It also suggests that when searching on one part of speech, one is unlikely to find words that are synonymous with respect to a different part of speech. I suspect that the search is narrower even than this, and that when searching for a word that means the same as, say, *pitch* as a noun, one searches for something that is synonymous with *pitch*₁ (slope), *pitch*₂ (tonal frequency), *pitch*₃ (thrown ball), *pitch*₄ (sales talk), or some other meaning that *pitch* can have as a noun.

What can be said about the difference between more and less effective clues in general, or about what makes an effective clue effective?

Every crossword puzzle doer is keenly aware that some clues are more helpful than others. As already noted, knowledge of specific letters in specific positions can be more or less helpful, depending on what the letters are and which positions they occupy. From filling in orthogonal words, I learn that the last two letters of a

four-letter word are BT; immediately, before looking at the semantic clue, DEBT springs to mind. Why is this clue so effective? One possibility is that there is only one four-letter word in my lexicon that ends with BT. This is consistent with my introspection, for when I try hard to think of others, I am unsuccessful. More interestingly, I am reasonably confident that there are not many such words in the language. This does not really explain why the clue is effective, however. I suspect it would be possible to find another such structural clue that pointed unambiguously to a single target word that would not be nearly as effective. How is it that _ _BT gets so quickly to the (presumably) only four-letter word ending in BT that is in my lexicon? It seems unlikely that a search of my entire lexicon, or anything close to that, is required. I am aware of only one common five-letter word ending in BT; I suspect most readers will bring it to mind easily. It is not at all clear, however, how one goes about retrieving this word. (There is also at least one seven-letter word and one eight-letter word that end in BT, but they are considerably less common and undoubtedly more difficult to identify).

Any clue, by definition, delimits a subset of the lexicon—namely, that subset of items whose members are consistent with the clue. The example just given illustrates that a clue can delimit a very small subset of one’s lexicon indeed. This does not account for the effectiveness of such clues, because it begs the question of how one manages to focus one’s search in the “region” of the lexicon that contains the item(s) delimited by the clue.

In principle, it should be possible to determine precisely how much information any specified structural clue provides to a person with complete knowledge of a given (*OED*’s, his/her own) lexicon. It requires nearly 18 bits to specify a word in the 1991 *OED*’s corpus of 209,500 words. A clue, or set of clues, that would reduce the number of possible targets to, say, about 50 would convey approximately 12 bits of information. It seems fair to say that such a clue, or set, would be quite an informative one. Consider, for example, the set of clues: five letters, first and third letters C and D, respectively—that is, C _D __. If one knew the size of the lexicon, the percentage of words in it that have five letters, the percentage of the five-letter words that begin with C, and the percentage of the five-letter words beginning with C that have D in third-letter position, one could easily compute the number of items in the set of possibilities. One can imagine a set of not-implausible assumptions that would make the target possibilities relatively small. If, for example, one were to assume that about .1 of the words in the lexicon have five letters, about .05 of the five-letter words begin with C, and about .05 of the five-letter words

that begin with C have D in the third-letter position, the set of possibilities would be $.1 \times .05 \times .05 \times 209,500 \approx 52$.

It may strike the reader as likely that there are more than about 50 five-letter words in the language that begin with C and have D as the third letter, and, of course this exercise, with the arbitrary assignment of percentages, provides a very tenuous basis for expecting there to be so few. There are, after all, 17,576 ways to fill in the blanks of C_D_ _ . A moment's thought makes it clear that a small percentage of these possibilities form words; realization that the second letter and at least one of the final two must be vowels reduces the number of possibilities to 936, but this is still a large number relative to 52. In fact, a search of the *OED* yielded a list of 42 five-letter words with C and D in first- and third-letter positions, 16 of which are designated as obsolete or archaic. (The list is available by e-mail on request to the author.)

So it is the case that, given knowledge of the language as represented in the *OED*, the set of clues embodied in C_D_ _ would convey between 12 and 13 bits of information, thereby reducing the search space to roughly .0002 of its original size. Different structural clues would convey different amounts of information to an observer with full knowledge of the lexicon, and the amount of information conveyed by any particular structural clue is computable in principle. Presumably, no one has as complete a knowledge of language as is represented in the *OED*, but it is obvious that structural clues serve the purpose of reducing the size of the search space, and they often reduce it to a surprising degree.

Quantifying their effects for different people would require complete knowledge of the lexicons that individuals carry in their heads. However, it is possible to make some plausible conjectures about the relative informativeness of specific clues on the basis of what is known about the statistics of language and the assumption that language users have some knowledge of what those statistics are. Probably few puzzle doers know that the letter E occurs in written English about 140 times as frequently as does the letter Z, but most certainly know that E is a far more common letter than Z, and learning that a target word contains a Z is much more informative than learning that it contains an E. I suspect that most would consider _ _ _ _ B to be a much more helpful clue than _ _ _ _ S, recognizing that there are many fewer five-letter words that end in B than there are that end in S, in part because a terminal S is used to indicate the plural of so many four-letter nouns as well as the third-person singular of present-tense verbs. People know that certain letter combinations are common in certain letter positions and that others seldom occur, if ever: They expect to see TH, CH, and SP at the beginnings of

words, but not SR, CM, or WT; they would be surprised to see a long string of consonants or a long string of vowels, because they know such strings are highly unlikely.

In short, different clues can convey different amounts of information to people who have less than complete knowledge of the lexicon. Lower-frequency letters are likely to be more informative as clues than higher-frequency letters, and letters appearing in positions in which they infrequently appear are likely to be more informative than letters occurring in positions in which they often appear. More generally, it seems reasonable to assume that the relative informativeness of clues to real puzzle doers is roughly approximated by their relative informativeness to an ideal observer whose knowledge of the lexicon is complete.

Goldblum and Frost (1988) interpreted one aspect of their results to be an indication that the amount of information provided by a cluster of (adjacent) letters is greater than the sum of that provided by each of the cluster's constituents alone. This is a particularly interesting conclusion, because it can be true in an information-theoretic sense only if the occurrence of the constituent letters is negatively correlated. Consider a two-letter cluster, say AB. If the probability of these two letters occurring in combination is the product of the probabilities of their respective occurrences, $p(AB) = p(A)p(B)$ —which is to say that the occurrence of one is independent of the occurrence of the other, or their correlation is 0—then the information conveyed by their joint occurrence is exactly the sum of the information conveyed by their separate occurrences.

If the correlation is positive— $p(AB) > p(A)p(B)$ —then the information conveyed by their joint occurrence is less than the total of that conveyed by their individual occurrences. Perhaps the most obvious example of a letter combination illustrating this relationship is QU: Given the knowledge that Q has occurred, one can be almost certain that U follows it, and so knowing QU is not much better than knowing Q. If the correlation is negative— $p(AB) < p(A)p(B)$ —then the information conveyed by their joint occurrence is greater than the sum of that conveyed by their individual occurrences. The combination BT as the penultimate and final letters of a word illustrates this case; if B in the penultimate position conveys x bits and T in the final position conveys y bits, BT in the final two positions conveys more than $x + y$ bits.

What are the implications of the fact that one can search memory effectively for words that contain a specified silent letter or letter group?

Some words contain silent letters that affect their pronunciation, and some contain silent letters that have

no such effect. Words with a terminal E (BITE, FATE) illustrate the former case; those with a silent initial K (KNOT, KNIGHT) illustrate the latter. That we can retrieve words of both types from memory is obvious. However, it is not clear, in the absence of data, whether one of these types of clue is more effective than the other. In the examples given, effect on pronunciation is confounded with position within the word, and any study of the importance of a pronunciation effect would have to control for this factor.

The terminal E generally changes the pronunciation of the preceding vowel from short to long, as is illustrated by BITE versus BIT. GH at the end of a word may affect pronunciation, too, as illustrated by THOUGH versus THOU. That GH has an effect is considerably less clear in the cases of WEIGH and DOUGH, inasmuch as there is room for doubt as to whether we would pronounce WEI and DOU as we do these two words.

The list of questions prompted by the doing of crossword puzzles is easily extended. To wit: Is it easier to search memory on the basis of letters, phonemes, syllables, or morphemes? How does one characterize the size of an individual's vocabulary? How does one count polysemous words or different forms (tense, number) of the same word? What is stored in one's mental lexicon: Words? Morphemes? Semes? Syllables? Among the more interesting questions, in my view, are some that relate to the fundamental concept of a word: What is a word? What does it mean to have a word in one's vocabulary? What does it mean for a word to be in the language?

What is a word?

Throughout this article, the notion of a word has been taken for granted. Everyone knows what a word is, or at least so it would seem. But is that really the case? It can be very difficult to identify individual words in a speech sound stream. Their beginnings and endings are not nearly as clearly marked as they are in written language. If one looks at a spectrographic representation of "We were away in Europe," for example, one sees no clear beginnings and endings of the words that comprise the utterance. If we did not come to such a representation with the knowledge that the utterance that is represented is composed of five separate words, we would see little, if any, evidence of that in the representation itself. Similarly, if we did not already have models of the individual words in mind, there would be no way to segregate them auditorily within the sound stream.

When one listens to an unfamiliar language for the first time, one does not hear words, as such. A major difficulty with which an adult learner of a new language must contend is that of trying to parse continuous speech into

individual words; one may acquire a sizeable vocabulary by paired-associate learning and then be totally befuddled when exposed to the language in use by not "hearing" any of the words one has so laboriously learned.

Confining one's attention to written language, one might say that a word is that which is represented by a sequence of letters bounded by spaces; such a definition would suffice at least to provide the basis for counting the number of words (tokens) in this essay, say. Alternatively, one might define a word as that which is represented by a sequence of letters that can be found as an entry in a dictionary of the language, with the qualification that nonword entries are typically explicitly identified as such.

But many words, in this sense, have many dictionary definitions. Should such a word be counted as one word, or many? Should we think of the *pen* in "He signed the letter with a pen" as the same word as that in "He put the pig in the pen," or does it make more sense, from a psychological point of view, to consider them to be two different words?

Typically, we do not consider members of a homophonous word set (*meet, mete, meat; pair, pare, pear; vain, vane, vein*) to be the same word, even though they are acoustically identical. Nor, I think, do we usually consider homographs such as *sewer* (one who sews) and *sewer* (where waste water goes), or *lead* (the element) and *lead* (the frontmost position) to be the same word, even though they are orthographically the same. Why, then, should we consider *pen* (a writing instrument) and *pen* (an enclosure) to be one word just because they are pronounced and spelled the same way? What do we do, for example, with words with alternate spellings (*sceptic, skeptic; sulfur, sulphur; theater, theatre; enquire, inquire*); should they be counted as one word or two?

How we answer these questions has implications for how one would estimate the number of words in an individual's vocabulary or the number of words in the language. Miller (1951/1963) pointed out that the *OED* contains (or did contain at the time of his writing) 317 definitions of the word *take*, and that 171 of these meanings were found by Thorndike and Lorge (1944) in their corpus, which contained 3,504 tokens of *take*. This many definitions for one "word" is undoubtedly unusual, but entries with multiple definitions are common. The question of what constitutes a word prompts other closely related questions.

What does it mean to say that one has a word in one's vocabulary?

Does it mean that one understands all of its meanings? (I suspect that few people could satisfy this criterion with respect to more than a very few words.) Most of them? A reasonable subset of them? At least one of them?

What, in fact, does it mean to understand a word's meaning? To be able to state it in the form of a definition? To be able to use the word (in accordance with one or more of its definitions) appropriately in various contexts? To be able to interpret it correctly when one encounters it? In all cases in which one encounters it? At least in most cases?

What does it mean for a word to be “in the language?”

Does the fact that *absquatulate* is in the *OED* mean that it is in the language? What if the vast majority of the users of a language, say 99.9%, would not recognize a dictionary entry as a word; in what sense can such an entity be said to be “in the language”? I am not suggesting that *absquatulate* is necessarily in this category, although I would not be surprised if that were the case; the point is that there undoubtedly are “words” in dictionaries, especially such comprehensive dictionaries as the *OED*, that the vast majority of users of that language would not recognize as words. Whether one considers such entities to be words in the language is, perhaps, a matter of perspective. And what about obsolete or archaic words? (Recall that 16 of the 42 five-letter words listed in the *OED* that have C and D in first- and third-letter positions were designated as obsolete or archaic.) Should they be considered to *be* in the language, or only as having *been* in it?

The reader will note, no doubt, that the word *word* has been used throughout this article without much evidence of concern as to whether its intended meaning would be understood. My guess is that the question of intended meaning did not often surface in the reader's mind. This illustrates what strikes me as one of the more interesting aspects of language; we use it naturally, easily, and effectively for most purposes, and become aware of its ambiguities and limitations only when we focus on it and press for a degree of precision that usually is neither necessary nor, perhaps, even desirable for most purposes. In looking back over what has been said in this essay, one will see that the word (there it is again) *word* (and again) has been used in a variety of ways, and I have not been careful to distinguish among them.

When I have spoken of target words for crossword puzzles, for example, I have not been careful to note that some of them may have many dictionary definitions, whereas others have only one. In general, I have spoken as though any string of letters (not beginning or ending with a hyphen) that would be found as a dictionary entry is a word; I have treated *feet* and *feat* as different words, despite that they are pronounced the same way, but I have treated *sewer* and *lead* each as one word, despite that each has more than one meaning and is pronounced in more than one way. It seemed natural to do this in the context of this essay because, for purposes of designing and solving crossword puzzles, *feet*, *feat*, *sewer*, and *lead* are all distinct

and single words. For other purposes, one might count differently.

In short, *word*, like many other entities of its kind, has a variety of meanings. The only way to avoid, or at least decrease, its ambiguity, I suppose, is to invent several new words that can be used in place of *word* when greater precision is required than is typically necessary in everyday communication. Thus, one might use *word*₁ when one wishes to connote an acoustic event of a certain type, *word*₂ to designate a specific letter string, *word*₃ to represent a letter string associated with a specific dictionary definition, and so on. It is quite remarkable that we are able to communicate passably well without going to such lengths. Words, whatever they are, are truly amazing things.

The crossword puzzle as a vehicle for studying cognition

Goldblum and Frost (1988) argued that the use of a crossword puzzle paradigm has some advantages over traditional lexical decision tasks, in which people must decide whether letter strings comprise words, as a method of exploring certain aspects of lexical content and access.

If only a fragment of a word is presented, and the subject is asked to retrieve the whole word containing this fragment, the extent to which a particular fragment facilitates retrieval may reflect the functional role of this fragment in the lexicon. (p. 159)

Baron, Freyd, and Stewart (1980) used partial-word clues of the type found in crossword puzzles to study individual differences in memory retrieval.

H.M. is well known to students of amnesia as a much-studied individual who had normal memory for events preceding 1953 but severe amnesia for events that occurred after that time (Gabrieli, Cohen, & Corkin, 1988; Kensinger, Ullman, & Corkin, 2001). The cause of the amnesia was surgical resection of medial temporal lobe structures to control otherwise uncontrollable epileptic seizures. Among the numerous published studies of the effects of the procedure on H.M.'s cognitive abilities, one by Skotko et al. (2004) was prompted by the fact that H. M., then a man in his early 70s, had made a hobby of crossword puzzles over his entire adult life. He regularly solved them before and after his surgery. Skotko et al. used specially constructed puzzles to test H.M.'s ability to use clues that referenced information that would have been available before, or only after, 1953. H.M.'s performance on the puzzles that referenced information that would have been available before 1953 was on a par with healthy volunteer puzzle doers, but his performance was considerably poorer on puzzles that referenced information not available before 1953. Among the puzzles that Gabrieli et al. had designed for their

study were some that used post-1953 clues for pre-1953 targets; these were items that.

(a) referenced events or story lines that were particularly newsworthy after H.M.'s operation and (b) had answers that related to knowledge gained before his operation (e.g., clue: childhood disease successfully treated by Salk vaccine [postoperative knowledge]: answer: polio [preoperative knowledge]). (Skotko et al., 2004, p. 759)

Surprisingly, H.M. showed considerable improvement in solving these puzzles over several days, suggesting to the experimenters that H.M. was “capable of learning some new factual information when it can be fixed to already acquired knowledge” (Skotko et al., 2004, p. 767), which could be hopeful news for others with amnesia due to injury to medial temporal lobe structures.

How effective one is likely to be at solving crossword puzzles can be predicted to a considerable degree from scores on tests of vocabulary and of word generation (Underwood, Diehim, & Batt, 1994). Not surprisingly, proficiency at solving crossword puzzles also correlates positively with skill at anagrams (Underwood et al., 1994; Witte & Freund, 1995).

When one thinks of using crossword puzzles—or crossword-puzzle-like tasks—to study cognition, one is likely to have in mind the possibility of shedding light on processes involving the search of memory, especially lexical memory. But crossword puzzles can engage aspects of problem solving more generally. I have already mentioned the use of themes in puzzles, as well as the fact that the themes are sometimes given explicitly and sometimes have to be discovered. Even when they are given explicitly, however, they may be cryptic, thus posing a problem for the puzzle doer to solve. Consider, for example, a *New York Times* puzzle by Bette Sue Cohen with the title *Altogether now*. This puzzle gave me much trouble, especially because there appeared to be several cases of a potential target almost fitting, but not quite. At some point it dawned that *Altogether* provided a critical clue if parsed as *Al together*, signaling that some cells of the puzzle were to contain both of the letters *a* and *l*. With this realization, the puzzle became considerably easier. In another example from the *New York Times*, a puzzle by Jim Page had the title *Clueless*, and, for several of the targets, no semantic clue was given. What the puzzle doer had to discover was that in those instances the clue was the number identifying the puzzle square for the target's first letter. Thus, the target for the word beginning in square 21 was GAMBLERSCARDGAME; that for the word beginning in square 13 was ROMANXIII. These and countless other examples that could be given illustrate that crossword puzzles can provide cognitive challenges beyond those of searching lexical memory.

Crossword puzzle doing and mental aging

Presumably people do crossword puzzles for a variety of reasons: the momentary escape it provides from other claims on one's mind; the opportunity to meet a challenge, and hopefully to experience a feeling of modest accomplishment; or perhaps to engage in a form of mental calisthenics with the purpose of helping preserve one's cognitive assets—by preventing or postponing the onset of Alzheimer's disease or other causes of mental decline. Some crossword puzzle doers—a small but enthusiastic minority—do them competitively. The American Crossword Puzzle Tournament has been held annually for 33 years, from 1978 to 2007 in Stamford, Connecticut, and since 2008 in Brooklyn, New York. (The 33rd was held in February 2010).

It is a common belief that an effective way to ward off, or at least to slow down, the ravages of time on aging brains is to exercise them regularly with mentally challenging tasks, of which doing crossword puzzles qualifies as one. One finds claims to this effect both in the popular media (Doraiswamy, 2010) and in the scientific literature (Schaie & Willis, 1996; Sorenson, 1933). My sense is that the evidence either way is more suggestive than compelling.

Whether or not doing crossword puzzles postpones dementia, aging puzzle addicts can take some comfort in evidence that whatever skill that doing such puzzles requires appears to be relatively immune to the mental abuses of time, at least for long-term puzzle doers (Rabbitt, 1993; Witte & Freund, 1995). It appears that the experience and knowledge that comes with age more than compensate for declines in other abilities involved in the task (Hambrick et al., 1999). In any case, whatever the cognitive effects of regularly doing crossword puzzles, I feel relatively certain that committed puzzle doers will endorse the claim that the practice makes the abuses of age on mentation more tolerable than they might otherwise be.

What motivates people to do crossword puzzles is not the topic of this article, but it is an interesting question. I think I would like to understand my addiction better—but then again, I am not so sure. My true motivation could turn out to be some peculiar Freudian quirk of which I would do better to remain ignorant.

Author Note The puzzle designers from whose puzzles were taken examples used in this article include Virginia P. Abelson, Nancy W. Atkinson, Dale Burgener, Roger Coburn, Bette Sue Cohen, Adam Crosse, Charles M. Deber, Gloria Evans, Matt Gafney, Henry Hook, Nancy Nicholson Joline, Bert H. Kruse, Tap Osborn, Jim Page, Henry Quarters, Merle Reagle, Richard Silvestri, and Tom Underhill. I am grateful to Thomas Landauer for making available the data represented in Fig. 1 and for encouragement with this line of reflection, and to puzzler Arthur Schulman for helpful comments on a draft of the manuscript, which he returned with a personalized vowel-less (vwillss) puzzle of his own design.

Appendix: Solutions

Word represented in the Specific Letters in Specific Positions section: VINDICTIVE

Table 6 English palindromic words

a	deified	gig	mum	pop	sagas	terret
aha	devoved	hah	noon	pup	sees	tit
alula	did	huh	nun	radar	sexes	toot
anna	dud	I	oho	redder	shahs	tot
bib	eke	kayak	pap	refer	sis	tut
bob	ere	kook	peep	repaper	solos	wow
boob	essse	level	peeweep	reviver	tat	
civic	ewe	madam	pep	revver	tattarrattat	
dad	eye	minim	pip	rotator	teet	
deed	gag	mom	poop	rotor	tenet	

The list of palindromes in Table 6 is instructive in several ways. There are a few words in it that many readers may not recognize as words. I am guessing that among them are ALULA, ANNA, DEVOVED, ESSSE, PEEWEEP, and TATTARRATTAT. All appear in the *OED*, according to which an ALULA is a particular cluster of bird wing feathers, an ANNA is a sixteenth part of an East Indian rupee, DEVOVED means vowed, ESSSE is an archaic word for ashes, a PEEWEEP is a bird, and TATTARRATTAT is a “nonce word” coined by James Joyce to represent a knock on a door.

Perhaps more interesting is the fact that several words are missing from the *OED* that one might have expected would be there. REDIVIDE, REIFY, and REV are there, but REDIVIDER, REIFIER, and REVVER are not. One wonders why, if redividing, reifying, and revving are recognized as bona fide actions, the people who perform them are not acknowledged to be redividers, reifiers, and revvers. This is especially puzzling in view of the fact that REVIVER *is* listed as a word. Another omission that seems a little strange is EVITATIVE. EVITATE (“shun”) is there, as is EVITATION (“shunning”), but not EVITATIVE, which, according to Wikipedia, is a grammatical case found in Australian languages but, in view of the meanings of EVITATE and EVITATION, might be thought to be an adjective meaning “inclined to shun.” A tartrate is a salt of tartaric acid, which is a principal acid of wine and an ingredient of baking powder; DETARTRATED, one might think, would be the past tense of DETARTRATE, the action of removing tartrates from substances, but the *OED* does not recognize it as such. Also not in the list because not in the *OED* is my all-time favorite palindrome, AIBOHPHOBIA, coined within the past few decades, perhaps as a joke, to mean “irrational fear of palindromes”.

Table 7 Some clues illustrating word play

Clue	Target
<i>Cheese catchers</i>	CAMERA
<i>Minimum ends</i>	EMS
<i>Gsge</i>	SCRAMBLEDEGGS
<i>Long lunches</i>	SUBS
<i>V</i>	CENTEROFGRAVITY
<i>Y</i>	HAPPYENDING
<i>Butter</i>	RAM
<i>Run in the heat</i>	MELT
<i>Bedding down</i>	EIDER
<i>A cotillion in a pastry shop</i>	ABUNDANCE
<i>A little lower</i>	CALF
<i>Little colonizer in torment</i>	ANTAGONISTIC
<i>Moving picture</i>	TEARJERKER

Table 8 Solutions to the maxims in the Clue Interaction section

ASTITCHINTIMESAVESNINE
THEEARLYBIRDGETSTHEWORM
MANYHANDSMAKELIGHTWORK
TOOMANYCOOKSSPOILTTHEBROTH
AROLLINGSTONEGATHERSNOMOSS

Table 9 Answer to the clue for a four-line verse

RAISINGFROGSFORPROFIT
ISAPLAINANDSIMPLEJOKE
FORHOWCANONEMAKEMONEY
WHENSOMANYOFTHEMCROAK

(From a Sunday *New York Times* puzzle by Tap Osborn)

Table 10 Targets for clues in Table 5

-
- Word with *space* or *limits* (OUTER)
-
- Word with *blond* or *Wednesday* (ASH)
 - Prefix for *scope* and *gram*. (TELE)
 - Suffix for *right* and *court* (EOUS)
 - *Lima* or *Orson* (BEAN)
 - Suffix with *fraud* or *flat* (ULENT)
 - *Slug* or *song* follower (FEST)
 - Words after *loose as* or *silly as* (AGOOSE)
 - Kind of *song* or *park* (THEME)
-

Anagram of *tipercu*: PICTURE

References

- Balota, D. A., & Lorch, R. F. (1986). Depth of automatic spreading activation: Mediated priming effects in pronunciation but not in lexical decisions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *12*, 336–345.
- Baron, J. (1985). *Rationality and intelligence*. New York: Cambridge University Press.
- Baron, J., Freyd, J., & Stewart, J. (1980). Individual differences in general abilities useful in solving problems. In R. S. Nickerson (Ed.), *Attention and performance VIII* (pp. 763–778). Hillsdale, NJ: Erlbaum.
- Barrows, H. S., Freightner, J. W., Neufeld, V. R., & Norman, G. R. (1978). *Analysis of the clinical methods of medical students and physicians*. Hamilton, ON: McMaster University, School of Medicine.
- Beilin, H., & Horn, R. (1962). Transition probability effects in anagram problem solving. *Journal of Experimental Psychology*, *63*, 514–518.
- Beller, S., & Kuhnmünch, G. (2007). What causal conditional reasoning tells us about people's understanding of causality. *Thinking and Reasoning*, *13*, 426–460.
- Blake, M. (1973). Prediction of recognition when recall fails: Exploring the feeling-of-knowing phenomenon. *Journal of Verbal Learning and Verbal Behavior*, *12*, 311–319.
- Bousfield, W. A. (1953). The occurrence of clustering in the recall of randomly arranged associates. *Journal of General Psychology*, *49*, 229–240.
- Bousfield, W. A., & Sedgewick, C. H. (1944). An analysis of sequences of restricted associative responses. *Journal of General Psychology*, *30*, 149–165.
- Bowers, K. S., Farvolden, P., & Mermigis, L. (1995). Intuitive antecedents of insight. In S. M. Smith, T. B. Ward, & R. A. Finke (Eds.), *The creative cognition approach* (pp. 27–52). Cambridge, MA: MIT Press, Bradford Books.
- Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). *A study of thinking*. New York: Wiley.
- Doraiswamy, P. M. (2010, March/April). Boost your brain health. *AARP Magazine*, 39–42.
- Edwards, A. L. (1957). *Experimental design in psychological research*. New York: Rinehart.
- Ekstrand, B. F., & Dominowski, R. L. (1968). Solving words as anagrams: II. A clarification. *Journal of Experimental Psychology*, *77*, 552–558.
- Elstein, A. S., Shulman, L. S., & Sprafka, S. A. (1978). *Medical problem solving: An analysis of clinical reasoning*. Cambridge, MA: Harvard University Press.
- Evans, J. S. B. T. (2007). *Hypothetical thinking: Dual processes in reasoning and judgment*. New York: Psychology Press.
- Evans, J. S. B. T., & Over, D. (2004). If. Oxford: Oxford University Press.
- Farvolden, P. (1991). *Intuition and early cognitive processes in the solving of partial word puzzles*. Unpublished undergraduate honors thesis, University of Waterloo.
- Gabrieli, J. D. E., Cohen, N. J., & Corkin, S. (1988). The impaired learning of semantic knowledge following bilateral medial temporal-lobe resection. *Brain and Cognition*, *7*, 157–177.
- Gigerenzer, G., & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science*, *1*, 107–143.
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*, 650–669.
- Gigerenzer, G., & Goldstein, D. G. (1999). Betting on one good reason: The Take the Best heuristic. In G. Gigerenzer & P. M. Todd (Eds.), *Simple heuristics that make us smart* (pp. 75–95). New York: Oxford University Press.
- Goldblum, N., & Frost, R. (1988). The crossword puzzle paradigm: The effectiveness of different word fragments as clues for the retrieval of words. *Memory & Cognition*, *16*, 158–166.
- Goldstein, D. G., & Gigerenzer, G. (1999). The recognition heuristic: How ignorance makes us smart. In G. Gigerenzer & P. M. Todd (Eds.), *Simple heuristics that make us smart* (pp. 37–58). New York: Oxford University Press.
- Gruneberg, M. M., & Monks, J. (1974). “Feeling of knowing” and cued recall. *Acta Psychologica*, *38*, 257–265.
- Hambrick, D. Z., Salthouse, T. A., & Meinz, E. J. (1999). Predictors of crossword puzzle proficiency and moderators of age-cognition relations. *Journal of Experimental Psychology: General*, *128*, 131–164.
- Hamming, R. W. (1950). Error detecting and error correcting codes. *Bell System Technical Journal*, *29*, 147–160.
- Hammond, K. R. (1978). Toward increasing competence of thought in public policy formation. In K. R. Hammond (Ed.), *Judgment and decision in public policy formation* (pp. 11–32). Boulder, CO: Westview Press.
- Indow, T. (1980). Some characteristics of word sequences retrieved from specified categories. In R. S. Nickerson (Ed.), *Attention and performance VIII* (pp. 621–636). Hillsdale, NJ: Erlbaum.
- Indow, T., & Togano, K. (1970). On retrieving sequence from long term memory. *Psychological Review*, *77*, 317–331.
- Jenkins, J. J., & Palermo, D. S. (1964). *Word association norms*. Minneapolis, MN: University of Minnesota Press.
- Johnson, D. M., Johnson, R. C., & Mark, A. L. (1951). A mathematical analysis of verbal fluency. *Journal of General Psychology*, *44*, 121–128.
- Kaplan, I. T., & Carvellas, T. (1968). Effect of word length on anagram solution time. *Journal of Verbal Learning and Verbal Behavior*, *7*, 201–206.
- Kaplan, I. T., Carvellas, T., & Metlay, W. (1969). Searching for targets in letter sets of varying size. *Journal of Experimental Psychology*, *82*, 377–380.
- Karwoski, T. F., & Schacter, J. (1948). Journal of psychological studies in semantics: III. Reaction times for similarity and difference. *Journal of Social Psychology*, *28*, 103–120.
- Kensinger, E. A., Ullman, M. T., & Corkin, S. (2001). Bilateral medial temporal lobe damage does not affect lexical or grammatical processing: Evidence from amnesic patient H. M. *Hippocampus*, *11*, 347–360.
- Kent, G. H., & Rosanoff, A. J. (1910). A study of association in insanity. *American Journal of Insanity*, *67*, 37–96.
- Legrenzi, P., Girotto, V., & Johnson-Laird, P. N. (1993). Focusing in reasoning and decision making. *Cognition*, *49*, 37–66.

- Mayzner, M. S., & Tresselt, M. E. (1958). Anagram solution times: A function of letter order and word frequency. *Journal of Experimental Psychology*, *56*, 376–379.
- Mayzner, M. S., & Tresselt, M. E. (1959). Anagram solution times: A function of transition probabilities. *Journal of Experimental Psychology*, *47*, 117–125.
- Mayzner, M. S., & Tresselt, M. E. (1962). Anagram solution times: A function of word transition probabilities. *Journal of Experimental Psychology*, *63*, 510–513.
- McClelland, J. L., & Rumelhart, D. E. (1981). An interactive activation model of context effects in letter perception: Part 1. An account of basic findings. *Psychological Review*, *88*, 375–407.
- McLeod, P. D., Williams, C. B., & Broadbent, D. E. (1971). Free recall with assistance from one and from two retrieval cues. *British Journal of Psychology*, *62*, 59–65.
- McNamara, T. P. (1992a). Priming and constraints it places on theories of memory and retrieval. *Psychological Review*, *99*, 650–662.
- McNamara, T. P. (1992b). Theories of priming: I. Associative distance and lag. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *18*, 1173–1190.
- McNamara, T. P., & Altarriba, J. (1988). Depth of spreading activation revisited: Semantic mediated priming occurs in lexical decisions. *Journal of Memory and Language*, *27*, 545–559.
- Mednick, S. A. (1962). The associative of the creative process. *Psychological Review*, *69*, 220–232.
- Metcalfe, J., & Wiebe, D. (1987). Intuition in insight and noninsight problem solving. *Memory & Cognition*, *15*, 238–246.
- Miller, G. A. (1963). *Language and communication*. New York: McGraw-Hill. (Originally published 1951)
- Mynatt, C. R., Doherty, M. E., & Dragan, W. (1993). Information relevance, working memory, and the consideration of alternatives. *Quarterly Journal of Experimental Psychology*, *46A*, 759–778.
- Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (1998). The University of South Florida word association, rhyme, and word fragment norms. Available at www.usf.edu/FreeAssociation/
- Nickerson, R. S. (1977). Crossword puzzles and lexical memory. In S. Dornic (Ed.), *Attention and performance VI* (pp. 699–718). Hillsdale, NJ: Erlbaum.
- Nickerson, R. S. (1980). Motivated retrieval from archival memory. In H. E. Howe, Jr., & J. H. Flowers (Eds.), *Motivation and cognitive processes: 1980 Nebraska Symposium on Motivation* (pp. 73–119). Lincoln, NE: University of Nebraska Press.
- Nickerson, R. S. (2005). Bertrand's chord, Buffon's needle, and the concept of randomness. *Thinking and Reasoning*, *11*, 67–96.
- Nickerson, R. S. (2010). *Mathematical reasoning: Patterns, problems, conjectures, and proofs*. New York: Psychology Press.
- O'Connor, J. (1928). *Born that way*. Baltimore, MD: Williams & Wilkins.
- Rabbitt, P. (1993). Does it all go together when it goes? The nineteenth memorial Bartlett lecture. *Quarterly Journal of Experimental Psychology*, *46A*, 385–434.
- Read, J. D., & Bruce, D. (1982). Longitudinal tracking of difficult memory retrievals. *Cognitive Psychology*, *14*, 280–300.
- Reyna, V. F. (2004). How people make decisions that involve risk: A dual-process approach. *Current Directions in Psychological Science*, *54*, 60–66.
- Ronning, R. R. (1965). Anagrams solution times: A function of the "ruleout" factor. *Journal of Experimental Psychology*, *69*, 35–39.
- Rundus, D. (1973). Negative effects of using list items as recall clues. *Journal of Verbal Learning and Verbal Behavior*, *12*, 43–50.
- Schaie, K. W., & Willis, S. L. (1996). Psychometric intelligence and aging. In F. Blanchard-Fields & T. M. Hess (Eds.), *Perspectives on cognitive change in adulthood* (pp. 293–322). New York: McGraw-Hill.
- Schulman, A. (1996). The art of the puzzler. In M. P. Friedman & E. C. Carterette (Eds.), *Cognitive ecology* (pp. 293–321). New York: Academic Press.
- Selfridge, O., & Neisser, U. (1960). Pattern recognition by machine. *Scientific American*, *203*, 60–68.
- Shiffrin, R. M. (1970). Memory search. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press.
- Skotko, B. G., Kensinger, E. A., Locascio, J. J., Einstein, G., Rubin, D. C., & Tupler, L. A. (2004). Puzzling thoughts for H.M.: Can new semantic information be anchored to old semantic memories? *Neuropsychology*, *18*, 756–769.
- Sloman, S. A. (2002). Two systems of reasoning. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 379–398). Cambridge: Cambridge University Press.
- Smith, V. L., & Clark, H. H. (1993). On the course of answering questions. *Journal of Memory and Language*, *32*, 25–38.
- Sorenson, H. (1933). Mental ability over a wide range of adult ages. *Journal of Applied Psychology*, *17*, 729–741.
- Strathern, P. (2000). *Mendeleyev's dream: The quest for the elements*. New York: Berkley.
- Thorndike, E. L., & Lorge, I. (1944). *The teacher's word book of 30,000 words*. New York: Columbia University, Teachers College.
- Toglia, M. P., & Battig, W. F. (1978). *Handbook of semantic word norms*. Hillsdale, NJ: Erlbaum.
- Tulving, E., Schacter, D. L., & Stark, H. A. (1982). Priming effects in word-fragment completion are independent of recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *8*, 336–342.
- Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*, 207–232.
- Underwood, G., Diehim, C., & Batt, V. (1994). Expert performance in solving word puzzles: From retrieval clues to crossword clues. *Applied Cognitive Psychology*, *8*, 531–548.
- Wallas, G. (1945). *The art of thought*. London: Watts. (Original work published 1926)
- Wason, P. C., & Evans, J. S. B. T. (1975). Dual processes in reasoning? *Cognition*, *3*, 141–154.
- Witte, K. L., & Freund, J. S. (1995). Anagram solution as related to adult age, anagram difficulty, and experience in solving crossword puzzles. *Aging and Cognition*, *2*, 146–155.
- Woodrow, H., & Lowell, F. (1916). Children's association frequency tables. *Psychological Monographs*, *22*(97), 1–110.
- Woodworth, R. S. (1938). *Experimental psychology*. New York: Holt.