

# Contextual imagery in meaning and memory\*

IAN BEGG

*McMaster University, Hamilton, Ontario, Canada*

and

JAMES M. CLARK

*University of Western Ontario, London, Ontario, Canada*

Three hundred homonyms were selected, and sentence fragments were written to emphasize two meanings of each. The words were rated on image-arousing capacity both in and out of context and on frequency of occurrence in context. Imagery values for the words out of context were predicted quite well by an average of the contextual imagery ratings, weighted by their relative frequencies. The finding is consistent with the hypothesis that words presented in isolation are interpreted in specific senses according to a frequency bias, with imagery ratings reflecting those senses. In a memory experiment, words were selected to vary orthogonally on both contextual and out-of-context imagery. Recall of words in isolation was a function of out-of-context imagery, while recall of words presented in context was a function of contextual imagery, further supporting the hypothesis.

The most common operation defining mental imagery derives from group ratings of the ease with which words evoke images (Paivio, Yuille, & Madigan, 1968). Imagery, so defined, is positively related to performance levels in a variety of memory tasks (e.g., Paivio, 1971). Since many English words have several different meanings, each of which is presumably characterized by some degree of image-arousing capacity, an important question arises concerning just what it is that Ss rate when they are presented with lists of isolated words. A second important question arises concerning whether the different meanings will be differentially memorable. In the present paper, ratings of the image-arousing capacity of words presented in sentence context are compared with ratings of the more usual sort in a correlational study, and the two types of imagery are contrasted in a free recall task.

The question of how isolated words are interpreted has been addressed both theoretically and experimentally. For example, Kintsch (1970) suggested that particular meanings of words are normally selected by the intraverbal context in which the word is encountered but that, in the absence of context, meanings are selected on the basis of a frequency bias. The suggestion was supported by Winograd and Conn (1971), who found that homographs initially presented without context were better recognized if the recognition test biased a relatively frequent rather than a relatively infrequent context. They concluded that "words are encoded by Ss in some specific semantic representation and that the likelihood of a particular encoding is a function of its familiarity [p. 706]."

Further, if words are initially encountered in contexts that bias one particular meaning and are later tested for recognition (e.g., Light & Carter-Sobell, 1970) or frequency estimation (Rowe, 1973) of the same words in different semantic contexts, levels of performance are lower than for a variety of control conditions. Thus it appears that Ss interpret words in terms of specific meanings; which particular meaning is selected is normally a function of the intraverbal context in which the word is encountered, but in the absence of such a context, S selects a specific meaning based on its familiarity.

The relationship between imagery and meaning has also been addressed theoretically and experimentally. As early as 1890, James suggested that the reawakening of sensory images was highly related to the meanings of concrete words. Other discussions of imagery and referential meaning can be found in Ausubel (1965) and Bugelski (1970). The most developed theory of imagery as reference, that of Paivio (1971), has received considerable experimental support. For example, O'Neill (1972) found a strong relationship between rated definability of words and their image-arousing capacity. Paivio and Begg (1971a) found that image-arousal latencies and comprehension latencies are highly correlated for sentences of varying concreteness and grammatical structure. Further, Ss are less likely to detect synonym substitutions in concrete sentences than changes in the meaning of the same sentences, a pattern that is markedly different with abstract sentences (Begg & Paivio, 1969); in fact, if Ss rate the semantic similarity of pairs of concrete and abstract sentences differing with respect to synonym substitutions or reversals of the subject and object, the ratings show the same pattern as the memory results (Johnson, Bransford, Nyberg, & Cleary, 1972).

Thus it appears that isolated words are interpreted or

\*The research was supported by Grant A8122 from the National Research Council of Canada to Ian Begg. Reprints can be obtained from Ian Begg, Department of Psychology, McMaster University, Hamilton, Ontario, Canada, L8S 4K1. We thank Larry Jacoby, Betty Ann Levy, and Allan Paivio for critical comments on early versions of the paper.

encoded in terms of specific, rather than general, meanings and that mental imagery is related to the referential meaning of verbal material. There is as yet, however, no evidence concerning whether imagery ratings reflect specific interpretations of isolated words or some more general meaning. Based on the idea of imagery as a referential meaning reaction to words, however, one would expect imagery ratings to reflect specific word meanings, with more familiar meanings more likely to be aroused than less familiar meanings. The question, then, of what imagery ratings reflect for words with multiple meanings has two plausible answers. On the one hand, all Ss could be rating the most frequent meaning of the word. On the other hand, each meaning of the word could be rated by some proportion of Ss, depending on the relative frequency of occurrence of that meaning. The overall rating would then represent a weighted average, or expected value, over the various meanings.

### EXPERIMENT I

In order to decide between the alternative possibilities outlined above, a correlational study was carried out. Sentence fragments were written to emphasize two meanings for each of 300 homonyms. It should be noted that most high-frequency words are homonyms, since the same graphemic and phonemic signal can occur in many different senses, as evidenced by the high correlation between word frequency and the number of dictionary meanings (cf. Reder, Anderson, & Bjork, 1974; Schnorr & Atkinson, 1970). However, the current word pool was selected from Roget's *Thesaurus*, since contextual, rather than denotative, meaning was the manipulation of interest. Each word was rated for the ease of image arousal (I) in isolation and in each of the two contexts. Additionally, Ss rated the frequency of occurrence of the words in the contexts. All ratings were made on 7-point scales, following the procedure of Paivio et al (1968).

#### Method

**Subjects.** For the rating tasks, 50 students from the University of Western Ontario and 44 from McMaster University served as Ss. No payment was given for participation. Between 8 and 10 Ss served in each condition.

**Materials.** Initially, 300 words were selected from Roget's *Thesaurus*, with the criterion that each word was mainly a noun with at least two distinct meanings. For each word, two sentence fragments were written, one emphasizing each of two meanings. For example, "stamp on a letter" vs "stamp of the foot," "spring shower" vs "bridal shower," and "pairs of hose" vs "garden hose" are three of the pairs. Other examples can be found in Table 1. Ratings were collected on three measures which will be described separately below.

(1) *Background Imagery.* The 300 words were rated for I on a 7-point scale, with 7 being high, following the procedure of Paivio et al (1968). Ten Ss rated 100 of the words while 8 rated the other 200. As with all the ratings to follow, 100 words were rated by Ss from the University of Western Ontario and 200

were rated by Ss from McMaster University.

(2) *Contextual Imagery.* The sentence fragments, with the crucial words underlined, were presented to Ss in four groups of 9 or 10. They were instructed to rate I as above, but to rate the particular sense of the word emphasized by the sentence fragment. No S rated both senses of a given word.

(3) *Contextual Frequency (F).* The same materials as immediately above were presented to an additional four groups of 9 or 10 Ss, who were instructed to rate the frequency with which they thought the word occurred in the particular sense in which it was used in the sentence fragment. Again, the ratings were on a 7-point scale with 7 being high.

### Results and Discussion

**Imagery Ratings.** For each word presented without sentence context, the mean I rating was calculated across Ss. Fifty-one of the words appeared in the Paivio et al (1968) norms, and an additional 27 words appeared in an expanded version of the norms provided by Paivio. For the two sets of nouns, the respective correlations between the present norms and those of Paivio were .917 and .894. Across the 78 nouns, the correlation was .858. Therefore the present norms are in good agreement with more established norms, despite the fact that small groups from two different universities provided the present ratings. From this relationship we infer that the present norms can be used with confidence concerning their reliability.

The overall distributions of I ratings were very similar for both the words presented without context and those

Table 1  
High- and Low-Imagery Words Used in Concrete and Abstract Contexts

Concrete	Contextual Concreteeness	
	Concrete	Abstract
	High Background Imagery	
a final EXAMINATION		a careful EXAMINATION
play a RECORD		set a RECORD
PINCH on the bottom		make do in a PINCH
wooden CROSS		hybrid CROSS
HAND and foot		off HAND
enter the ROOM		no more ROOM
admire her FIGURE		calculate the FIGURE
wolf and his MATE		a MATE in three moves
chicken WING		left WING politics
return to NATURE		human NATURE
STAKE in the ground		STAKE in the future
a sharp ANGLE		a legal ANGLE
	Low Background Imagery	
coffee GROUNDS		GROUNDS for divorce
RECOGNITION of a friend		RECOGNITION for service
the COUNT in his castle		by the latest COUNT
facial EXPRESSION		everyday EXPRESSION
VIEW from the balcony		in VIEW of recent evidence
the alley WAY		the WAY to do it
pretty young MISS		hit and MISS
rockets in SPACE		SPACE of time
facial FEATURE		special FEATURE
art OBJECT		OBJECT of inquiry
JUSTICE of the peace		love of JUSTICE
violent DEMONSTRATION		DEMONSTRATION of a fact

presented in sentence contexts. For example, the respective 66th percentiles fell at ratings of 5.29 and 5.28, the 50th at 4.61 and 4.52, and the 33rd at 3.60 and 3.64. The sentence contexts thus did not generally bias responses one way or the other.

In the introduction, two hypotheses were outlined concerning the relationship between ratings of contextual I and ratings of the more usual sort. Either such ratings reflect (a) the contextual I of the higher-F context, or (b) the F-weighted average of the contextual-I values. Across the 300 homonyms, the correlation between I rated without context and the I of the higher-F context was .477 ( $p < .001$ ), providing some support for the first hypothesis. The second hypothesis is based on the assumption that I ratings in the absence of context represent an expected value over the probabilities of encountering a word in all its meanings,

$$\xi(I) = \frac{\sum p_i I_i}{\sum p_i}$$

where  $i = 1, 2, \dots, n$  and  $p_i$  is the probability of encountering the word in a context emphasizing a particular meaning,  $i$ ,  $\sum p_i$  is the probability of encountering the word at all, and  $n$  is the number of meanings assumed by the word. Since the words appeared in only two contexts in the present case, the predicted I value,  $I^*$ , was derived from

$$I^* = \frac{\sum F_i I_i}{\sum F_i}$$

where  $i = 1, 2$  and  $F_i$  and  $I_i$  represent ratings of word frequency and imagery in the given contexts. Over the 300 homonyms, the correlation between I ratings without context and  $I^*$  was .693 ( $p < .001$ ). The second hypothesis is preferred over the first for two reasons. First, .693 is reliably greater than .477,  $t(297) = 3.96$ ,  $p < .001$ . Second, the partial correlation between  $I^*$  and I remained .613 after I of the higher-F context was partialled out, while the correlation between I of the higher-F context and usual I became  $-.269$  with  $I^*$  partialled out.

For the 51 words appearing in both the present norms and those of Paivio et al (1968), the correlation between I from the present norms and  $I^*$  was .834, and between I from the Paivio et al norms and  $I^*$  it was .809. With I of the higher-F context partialled out, the  $r$ s remained .536 and .638, respectively ( $ps < .001$ ). For the 27 words appearing in Paivio's extended norms and the present norms,  $I^*$  correlated .702 with the present norms and .726 with Paivio's, with respective partial correlations of .654 and .717 ( $ps < .001$ ). Across the 78 words, the respective correlations were .788 and .799, with partial correlations of .605 and .612 ( $ps < .001$ ). In none of the

above cases did the correlation between I and I of the higher-F context remain reliable at  $p < .05$  after partialling out  $I^*$ .

Clearly, the expectancy model of I ratings receives strong support from the present data. The general idea is that images represent specific meanings of words and that Ss rate the image-arousing capacity of those meanings in the normal tasks. Additionally, as suggested by Kintsch (1970) and supported by Winograd and Conn (1971), words presented in isolation are more likely to be interpreted in terms of more, rather than less, familiar meanings.

**Subsidiary Analyses.** In addition to the ratings discussed above, additional groups of Ss, again in small groups, rated the semantic similarity (SS) of the two contexts for each of the words on a 7-point scale. Since Paivio and Begg (1971b) found that low-I words share more associations with each other than do high-I words, it was expected that different meanings of the lower-I words would be more similar to each other than the different meanings of the higher-I words. In order to assess this possibility, the 300 words were divided into deciles of background I. The mean SS values, from the highest to the lowest decile, were 2.62, 2.82, 2.31, 2.54, 2.55, 3.39, 3.48, 3.16, and 3.42, with standard errors ranging from .16 to .24. Two analyses supported the apparent negative relation between I and SS. First, a Scheffé comparison following a reliable F ratio ( $p < .001$ ) from a one-way analysis of variance showed that the lowest four deciles of I received higher SS ratings than the highest six ( $p < .05$ ). Second, the correlation for the mean I and mean SS rating over the 10 deciles was  $-.787$  ( $p < .01$ ). Thus, not only do abstract words have more similar association patterns than do concrete words, but even different meanings of abstract words are rated as being more similar than different meanings of more concrete words.

The final analysis concerned the likelihood, given only the context, that Ss would supply the correct word. Groups of 9 or 10 Ss were given the sentence fragments with the crucial words omitted and were asked to fill in any words that would fit the context. The first point to note is that the probabilities were very low overall, with 375 of the 600 contexts being zero. Two sets of analyses were conducted to compare the relative contributions of contextual I and F to the probability S would generate the appropriate word to the contextual cue,  $Pr(W/C)$ . First, the 600 individual contexts were subdivided into four groups based on high- or low-I or F. Contexts rated above 5.00 on either measure were considered high. For contexts high on both measures,  $Pr(W/C)$  was .160; for high-F, low-I contexts,  $Pr(W/C)$  was .131; for low-F, high-I, it was .070; and for contexts low on both measures, it was .044. All standard errors, as well as all those reported below, were less than .02, with  $n$ s ranging from 110 to 200. Three independent  $t$  tests were conducted. First, high-F contexts were associated with

higher Pr(W/C) values than low-F contexts,  $t(598) = 6.20$ ,  $p < .001$ , with means of .15 and .05. Second, high-I contexts were associated with higher Pr(W/C) values than low,  $t(598) = 2.81$ ,  $p < .005$ , with means of .12 and .08. Third, contexts high in F but of low I [mean Pr(W/C) = .13] were contrasted with those of high I but low F (mean = .07),  $t(243) = 2.49$ ,  $p < .02$ . Therefore, F; more than I, is related to the likelihood, given a particular context, that a particular word will be generated.

In the second set of analyses, both contextual I and F were considered within individual words. Thus, any given word had one context that was higher in F than the other; the respective means, .133 and .063, differed reliably,  $t(299) = 4.86$ ,  $p < .001$ . Similarly, the higher-I and lower-I contexts were associated with different values of Pr(W/C), .118 and .079,  $t(299) = 2.64$ ,  $p < .001$ . Finally, for 96 words in which contextual I and F were negatively related, the high-F, low-I words (mean = .121) were associated with higher values of Pr(W/C) than the lower-F, higher-I contexts (mean = .071),  $t(95) = 2.06$ ,  $p < .05$ . Thus, whether the analysis is considered in terms of absolute or relative values of I and F, the latter is a more positive correlate of the probability that the context will elicit the word than is the former, although both variables are related to the probability.

**Summary.** The most notable result of Experiment I is that normal ratings of I can be quite accurately predicted from contextual I and F. Further, the relationship found is at best an underestimate of the true relationship, since the correlation is attenuated through imperfect reliability and since only two contexts were used to form the prediction. Thus, not only do Ss interpret words presented without context in terms of specific meanings, but those meanings themselves are a function of relative familiarity, and I ratings reflect those meanings. Consequently, a memory experiment was conducted to determine whether memorability of the words would be more a function of contextual or usual I values.

## EXPERIMENT II

In the second experiment, Ss were presented with normally high- and low-I words, either with no particular context or with contexts biasing either a high- or low-I sense of the words. In order to make predictions concerning the relative memorabilities of the words in each of the conditions, some consideration of what is remembered in a memory task must be given. From the first experiment, it was concluded that Ss interpret words in specific senses. If the further assumption is made that Ss remember words in terms of interpretations, predictions follow readily. That is, for words presented without any particular context, normal I will be positively related to recall; for words presented

in sentence context, however, contextual I should be a more potent correlate of recall than should normal I. The most critical contrast is thus between normally high-I words used in a low-I sense and normally low-I words used in a high-I sense. Words of the latter sort should be more memorable than those of the former.

### Method

**Subjects.** Sixty Ss were paid \$2/h for their participation. Twenty Ss recruited from Canada Manpower in Hamilton, Ontario, served in the free recall of isolated words, with 10 Ss receiving each of the two orders required to counterbalance I with list position. For the recall of words in sentence fragments, 40 students from McMaster University served as Ss, with 10 Ss receiving each of the four orders required to counterbalance both types of I with list position.

**Materials.** From the 300 homonyms described in the first experiment, 24 were selected as experimental items for the recall tasks. Twelve were normally high-I and 12 normally low, with mean I values of 5.68 and 2.69. Both the high- and low-I words had meanings that were high in contextual I, respective Is = 6.12, 5.59, or low, Is = 2.77, 2.69. The sentence fragments are presented in Table 1. The mean rated F values for the four conditions were between 4.70 and 4.96. It should be noted that we conducted an additional experiment in which Ss recalled words that were normally of high, medium, or low I, but which were presented in high- or low-F contexts. The only effect of F was that the effect of I was slightly larger in low- than in high-F words. Consequently, controlling F is probably not crucial over the range of F values in the present norms.

The 24 words were randomly arranged in a list with the restriction that two out of every four words were high-I. A second form of the list was then constructed, with the restriction that every ordinal position occupied by a high-I word in the first list was occupied by a low-I word in the second. Over the two lists, each of the 24 ordinal positions was occupied once by a high-I and once by a low-I word, thus counterbalancing I and ordinal position. The two lists were used as material in the no-context condition.

In the above lists, each block of four words contained two high- and two low-I words. One of each was selected at random to be a member of Set A, with the remaining ones being Set B. For the first list, words of Set A were written in the high-I context and words of Set B were written in the low-I context. In the second list, the procedure was reversed. For both the first and second lists, the order of words was identical to the first list in the no-context condition. Similarly, the words from the second list in the no-context condition were also arranged in two lists, with Set A and Set B words treated as above. Over the four lists, then, each of the 24 ordinal positions was occupied once by each of the four types of words. No homonym appeared twice in the same list.

**Procedure.** The two lists of words alone were each read to 10 Ss at the rate of 5 sec/word. Following presentation, 5 min were allowed for free written recall. Each of the four lists of words in context was read to two groups of 5 Ss, at the rate of 5 sec/sentence fragment. For one group, the homonym was stressed in each fragment and Ss were instructed only to recall the stressed word. For the other group, each fragment was read with a normal intonation and Ss were instructed to recall as much as they could. Thus, eight groups of 5 Ss served in the context conditions, with 20 Ss having stressed presentation and 20 nonstressed and with 10 Ss hearing each of the four lists. All Ss were given 5 min for free written recall.

In the context conditions, each S recalled words of four types, namely, words high and low in both normal and contextual I. Consequently, each S contributed four scores to the analysis. However, the contextual-I manipulation for any particular word

was a between-Ss variable. In order to avoid the "language-as-fixed-effect" fallacy (Clark, 1973), then, two analyses are possible. On the one hand, the data can be analyzed with Sets A and B as a within-S factor and the context in which each was used as a between-S factor. An interaction between the factors, such that the high-I context exceeded the low-I context for each set, could not be interpreted as a fixed-effect fallacy, since variability due to the Set by Context effect is nested within variability due to Ss. However, it is conceptually simpler to treat contextual I as a within-S factor, since each S did have words of each type in his condition, and, following Clark's (1973) suggestion, to calculate a quasi-F ratio ( $F'$ ). Although only the latter approach is mentioned in the results section, the former was also calculated, with exactly the expected results.

### Results and Discussion

**No-Context Condition.** The first analysis concerned recall of the words presented without sentence context. As usual, more high- than low-I words were recalled, 4.90 vs 3.80 out of 12,  $F(1,19) = 5.08$ ,  $p < .05$  (respective standard errors of means are .36 and .34). It should be noted that the difference is smaller than usual, probably because the 20 Manpower Ss varied extremely with respect to age, education, etc. However, it can be concluded that in the absence of context, the words used here are usual in their effects (and that the effect of word I on recall can be obtained in a "real-world" sample of unemployed people).

**Context Conditions.** For the words presented in sentence context, the first analysis consisted of analyzing the number of critical words recalled by each S with a 2 by 2 by 2 analysis of variance, with presentation type (stressed or normal) as an independent factor and both types of imagery (contextual or no-context) as repeated factors. Note that although each S contributed four scores to the analysis, the words whose recall constituted the score for high contextual I for a given S were different from the words whose recall constituted the score for low contextual I, and vice versa. This question will be addressed further below. At any rate, in the present analysis, words were treated as a fixed effect and Ss as a random effect.

The presentation-type variable did not contribute to any main effect or interaction, so it will not be discussed further in the present section. Ss recalled more words encountered in high-I contexts than in low-I contexts,  $F(1,38) = 7.66$ ,  $p < .01$ . The more usual measure of I did not contribute to a reliable main effect,  $F = 1.03$ , although the two variables interacted,  $F(1,38) = 6.35$ ,  $p < .05$ . The effect of contextual I was larger for normally abstract words, whose respective high- and low-I contexts led to recall scores of 2.65 and 1.73 words of a possible 6 (respective standard errors of the means were .19 and .16), than it was for normally concrete words, whose respective scores were 2.53 and 2.25 (standard errors were .23 and .19).

Post hoc *t* tests were conducted with the following conclusions. Normally low-I words presented in low-I contexts were less well recalled than words of the other

three types. In the critical contrast between normally abstract words used concretely, such as "coffee GROUNDS," and normally concrete words used abstractly, such as "no more ROOM," Ss were more likely to recall words of the former than the latter variety. That is, the variation in contextual I is potent enough to reverse the normal I effect. The final point is that it is apparently easier to increase memorability of abstract words by using them concretely than it is to reduce the memorability of concrete words by using them abstractly.

In the second analysis, words were treated as a random effect, with Ss fixed. That is, each high- or low-I word was recalled after presentation in two contexts, with the word either stressed or nonstressed. The dependent variable, the number of Ss recalling a word, was analyzed with a 2 by 2 by 2 analysis of variance, with no-context I as a between-word factor and contextual I and presentation type as within-word factors. As in the first analysis, presentation type contributed to no reliable main effects or interactions. Words were recalled by more Ss when the words were presented in high-I than low-I contexts,  $F(1,22) = 9.99$ ,  $p < .01$ . Combining the *F* ratio from the first analysis with that from the second, the minimum value of the quasi-*F* ratio (cf. Clark, 1973) is  $F'(1,51) = 4.34$ ,  $p < .05$ . The interaction effect in the second analysis was marginal,  $F(1,22) = 3.94$ ,  $p < .06$ , yielding a nonsignificant  $F'(1,60) = 2.43$ ,  $p = .12$ . Thus the effect of contextual I remains reliable when treating both words and Ss as random effects, while the interaction effect does not.

Over the two analyses of recall of words encountered in either high- or low-I contexts, it is justifiable to conclude that words are recalled at higher levels when used concretely than when used abstractly. A further test of the same conclusion can be derived from the data. Since no S saw the same word in both contexts, 10 of the Ss in the stressed condition saw six concrete words used abstractly, while the other 10 Ss saw the same six words in concrete contexts, and vice versa. The same is true in the nonstressed condition and for the abstract words. Consequently, there are eight comparisons in the data between words presented in concrete and abstract contexts, with 10 Ss for each context in each comparison. That is, there are four comparisons in each of the stressed and unstressed conditions, four for each of concrete and abstract words, and four for each of two sets of words, allowing direct looks at replications in the experiment. In seven of the eight cells, words were more often recalled after presentation in concrete than abstract contexts (*Z* for sign test = 4.24,  $p < .001$ ). The discrepant cell was that one set of normally concrete words was recalled better in abstract than concrete contexts, but only when the words were stressed on presentation, and Ss were not required to recall the contexts.

## GENERAL DISCUSSION

The general point of this paper concerns meaning, imagery, and memory. That is, words are interpreted and remembered as specific meanings. If the words are presented in the context of other words, the context biases the interpretation of the word towards a specific sense. If the word is presented without specific context, S provides a likely context by selecting a relatively frequent sense of the word. To the extent that the specific senses of words are differentially image-arousing, they are differentially memorable. The familiar difference in recall between concrete and abstract words can therefore be interpreted as reflecting differential image-arousing capacity of relatively frequent specific senses of each word class.

The results of the present study that are consistent with the above interpretation are twofold. First, imagery ratings of isolated words are well predicted by the frequency-weighted average of imagery ratings of specific senses of the words. Second, highly imageable senses of words lead to higher recall levels than do less imageable senses of the same words. Finally, the memory difference in free recall probably represents a conservative test of imagery effects, since the largest effects due to imagery are observed in cued recall (e.g., Begg, 1972, 1973; Begg & Robertson, 1973; Bower, 1970), especially when the imagery values of the cues are varied (e.g., Lockhart, 1969; Paivio, 1969; Philipchalk & Begg, 1971).

## REFERENCES

- Ausubel, D. P. A cognitive structure view of word and concept meaning. In R. C. Anderson and D. P. Ausubel (Eds.), *Readings in the psychology of cognition*. New York: Holt, Rinehart & Winston, 1965.
- Begg, I. Recall of meaningful phrases. *Journal of Verbal Learning & Verbal Behavior*, 1972, 11, 431-439.
- Begg, I. Imagery and integration in the recall of words. *Canadian Journal of Psychology*, 1973, 27, 159-167.
- Begg, I., & Paivio, A. Concreteness and imagery in sentence meaning. *Journal of Verbal Learning & Verbal Behavior*, 1969, 8, 821-827.
- Begg, I., & Robertson, R. Imagery and long-term retention. *Journal of Verbal Learning & Verbal Behavior*, 1973, 12, 689-700.
- Bower, G. H. Imagery as a relational organizer in associative learning. *Journal of Verbal Learning & Verbal Behavior*, 1970, 9, 529-533.
- Bugelski, B. R. Words and things and images. *American Psychologist*, 1970, 25, 1002-1012.
- Clark, H. H. The language-as-fixed-effect fallacy: A critique of language sampling in psychological research. *Journal of Verbal Learning & Verbal Behavior*, 1973, 12, 335-359.
- James, W. *Principles of psychology*. New York: Holt, 1890.
- Johnson, M. K., Bransford, J. D., Nyberg, S. E., & Cleary, J. J. Comprehension factors in interpreting memory for abstract and concrete sentences. *Journal of Verbal Learning & Verbal Behavior*, 1972, 11, 451-454.
- Kintsch, W. Models for free recall and recognition. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press, 1970. Pp. 331-373.
- Light, L. L., & Carter-Sobell, L. Effects of changed semantic context on recognition memory. *Journal of Verbal Learning & Verbal Behavior*, 1970, 9, 1-11.
- Lockhart, R. S. Retrieval asymmetry in the recall of adjectives and nouns. *Journal of Experimental Psychology*, 1969, 79, 12-17.
- O'Neill, B. J. Defineability as an index of word meaning. *Journal of Psycholinguistic Research*, 1972, 1, 287-298.
- Paivio, A. Mental imagery in associative learning and memory. *Psychological Review*, 1969, 76, 241-263.
- Paivio, A. *Imagery and verbal processes*. New York: Holt, 1971.
- Paivio, A., & Begg, I. Imagery and comprehension latencies as a function of sentence concreteness and structure. *Perception & Psychophysics*, 1971a, 10, 408-412.
- Paivio, A., & Begg, I. Imagery and associative overlap in short-term memory. *Journal of Experimental Psychology*, 1971b, 89, 40-45.
- Paivio, A., Yuille, J. C., & Madigan, S. A. Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology Monograph Supplement*, 1968, 76(1, pt. 2).
- Philipchalk, R. P., & Begg, I. Context concreteness and form class in the retention of nonsense syllables. *Journal of Verbal Learning & Verbal Behavior*, 1971, 10, 499-505.
- Reder, L. M., Anderson, J. R., & Bjork, R. A. A semantic interpretation of encoding specificity. *Journal of Experimental Psychology*, 1974, 102, 648-656.
- Rowe, E. J. Frequency judgments and recognition of homonyms. *Journal of Verbal Learning & Verbal Behavior*, 1973, 12, 440-447.
- Schnorr, J. A., & Atkinson, R. C. Study position and item differences in the short- and long-term retention of paired associates learned by imagery. *Journal of Verbal Learning & Verbal Behavior*, 1970, 9, 614-622.
- Winograd, E., & Conn, C. P. Evidence from recognition memory for specific encoding of unmodified homographs. *Journal of Verbal Learning & Verbal Behavior*, 1971, 10, 702-706.

(Received for publication June 10, 1974;  
accepted July 8, 1974.)