Journal of Experimental Psychology: General
1987, Vol. 116, No. 3, 310–311

0096-3445/87/$00.75

Conceptual Representation of Pictures and Words: Reply to Clark

Mary C. Potter
Massachusetts Institute of Technology

Judith F. Kroll
Mount Holyoke College

Clark (1987) offers a dual coding alternative (Paivio, 1971, 1986) to the conceptual hypothesis that Potter, Kroll, Yachzel, Carpenter, and Sherman (1986) proposed to explain the ease with which people can read and understand rebus sentences in which a picture replaces a noun. We present theoretical and empirical reasons for positing a conceptual representation that is distinct from the representation of an object's name and from a mental image of it. The hierarchical conceptual model has greater explanatory and predictive power and is more parsimonious overall than Clark's alternative.

Because there is always more than one way to account for a given set of experimental observations, the choice among theories must be based on parsimony as well as explanatory adequacy. Clark (1987) suggests that the conceptual model of Potter, Kroll, Yachzel, Carpenter, and Sherman (1986) is less parsimonious than an alternative dual coding model because the former posits three representational codes whereas the latter posits only two, albeit with a richer set of associative connections. The third code in Potter et al.'s (1986) model adds hierarchical structure to the representational architecture in that it takes input from two or more lower level codes (here, lexical and imaginal codes), abstracting the common concept represented in those codes.

Both the theoretical approaches considered here posit the existence of distinct lexical (verbal) and imaginal codes. The lexical code represents the sound, orthography, and articulation of words and the imaginal code represents the shape and perhaps other perceptual characteristics of perceptible entities. In dual coding theory, all other information is also represented in one or the other of these codes. According to the conceptual model, however, most of one's knowledge is represented in the third, amodal or neutral code; the lexical and imaginal codes act as access codes and play certain other roles in working memory such as providing an articulatory buffer or a mental sketchpad (Baddeley, 1986).

Both theories predict a processing advantage for stimuli presented in a form that matches the surface code in which the relevant information is stored. For example, both theories predict that written words will be named faster than pictured objects because both agree that articulatory-phonological information about words is stored in the verbal-lexical coding system, which is the first representational code activated by word stimuli. A picture first makes contact with the image code, then with the conceptual code, and only subsequently with the lexicon, so picture naming is relatively slow. (Potter et al., 1986, do not posit a direct link between images and the lexicon as shown in Clark, 1987, Figure 1B, but an indirect link via the concept: see Potter, 1979.)

The two theories disagree as to how conceptual information about an entity—such as a carrot—is stored, information such as the fact that carrots are edible, are plants, cost less than a dollar a bunch, and so on. The dual coding model as originally put forward (Paivio, 1971) proposed that abstract information of that kind was represented in the verbal system, implying that the information would be retrieved more rapidly if the critical stimulus was a word rather than a picture. The conceptual model proposes that all such information is represented in the amodal or conceptual code and hence is equally accessible (or roughly so) whether the stimulus is a word or picture. A large body of research measuring the time to respond to pictures and words in a variety of tasks (for reviews see Kieras, 1978; Potter, 1979; Snodgrass, 1984) has borne out this prediction of the conceptual model.

Paivio has modified his theory since 1971 (see Paivio, 1986) by attenuating the functional separation between codes; this modification makes unambiguous predictions difficult. The problem is illustrated in Clark's model (1987, Figure 1C) of the picture-word results reported in Potter et al. (1986). Clark proposes that an image code for a carrot may be as directly associated as the word *carrot* with the verbal code "eat" or "food." But, because the relative strengths of links depend on a person's experience, they are difficult to predict. Thus, in the dual coding model it is simply a coincidence that word stimuli generate about the same reaction times as pictures in categorization tasks (carrots are food), in pragmatic-judgment tasks (carrots can be eaten), and in many other tasks requiring conceptual knowledge about the object. In contrast, the conceptual model predicts exactly this outcome, whenever the task requires conceptual information. Only when the task requires information that is represented in one of the surface codes (for example, phonological information about the object's name) will a word have an advantage over a picture or vice versa.

Although any three-code, hierarchical model could be mimicked by a two-code model by multiplying the associative links between units and giving them appropriate weights, such a

Preparation of this article was supported by National Science Foundation Grant BNS83-18156 to Mary C. Potter. We thank Virginia Valian and Janellen Huttenlocher for helpful comments.

Correspondence concerning this article should be addressed to Mary C. Potter, Department of Brain and Cognitive Sciences, E10-039, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139.

model would be ad hoc unless the number of links and their weights could be specified in a principled manner. Moreover, each time a new piece of information such as a fact about carrots is acquired, two new links are needed in the dual coding model to attach the information to the two codes for "carrot," but only one new link is required in the conceptual model (see Clark, 1987, Figures 1B and C). The difference in parsimony becomes still more marked when there are more than two surface codes, as when a person knows more than one language. The many-one-many structure of the conceptual model minimizes the number of new connections required to represent the added language while successfully predicting (among other results) the relative time to translate from one language to the other, to name pictures in each language, and to judge the relation between a category name in one language and an exemplar in the second language (Potter, So, Von Eckardt, & Feldman, 1984). Thus, on grounds of explanatory power as well as parsimony, the conceptual model is to be preferred to Clark's dual coding model.

The evidence we have discussed so far concerns isolated words and pictures. Even if the conceptual model is accepted as descriptive of the architecture of mental representation, one can ask whether there is more in the lexical entry than simply orthographic, phonological, and articulatory information. It seemed possible to Potter et al. (1986) that lexical entries might contain certain kinds of information designed specifically to permit word meanings to be combined, or syntactic structures to be built. This specialized, modular information might not be needed when the task required processing of a word or picture in isolation, but might come into play when a sentence was processed. It was this lexical hypothesis (not the more general issue of dual coding) that Potter et al. (1986) investigated by substituting a picture for a word, in sentences presented very rapidly. As in the earlier work with single pictures, subjects had little difficulty understanding the rebus sentences. This suggested that amodal, conceptual information, not specifically lexical information, is used in structuring and combining the meanings of words during sentence processing.

Although Potter et al.'s (1986) research was not intended as a further test of Paivio's dual coding model, the results are in fact difficult to explain within that model. The verbal associations that are the basis for sentence processing in the dual coding model would be absent in most cases for inserted pictures. Hence, the dual coding model should predict that in such a context rebus pictures would have a marked disadvantage compared with the words they replace; yet there was no difference between the two when the task was to judge the plausibility of the sentence (see Clark, 1987, Table 1). This result strengthens the conclusion that the conceptual coding model of Potter et al. (1986) has greater explanatory power and is more parsimonious than the dual coding model put forward by Clark as an alternative.

References

Baddeley, A. (1986). *Working memory.* Oxford, England: Clarendon Press.

Clark, J. M. (1987). Understanding pictures and words: Comment on Potter, Kroll, Yachzel, Carpenter, and Sherman (1986). *Journal of Experimental Psychology: General, 116,* 307–309.

Kieras, D. (1978). Beyond pictures and words: Alternative information-processing models for imagery effects in verbal memory. *Psychological Bulletin, 85,* 532–554.

Paivio, A. (1971). *Imagery and verbal processes.* New York: Holt, Rinehart & Winston.

Paivio, A. (1986). *Mental representations: A dual coding approach.* New York: Oxford University Press.

Potter, M. C. (1979). Mundane symbolism: The relations among objects, names, and ideas. In N. R. Smith & M. B. Franklin (Eds.), *Symbolic functioning in childhood* (pp. 41–65). Hillsdale, NJ: Erlbaum.

Potter, M. C., Kroll, J. F., Yachzel, B., Carpenter, E., & Sherman, J. (1986). Pictures in sentences: Understanding without words. *Journal of Experimental Psychology: General, 115,* 281–294.

Potter, M. C., So, K.-F., Von Eckardt, B., & Feldman, L. (1984). Lexical and conceptual representation in beginning and proficient bilinguals. *Journal of Verbal Learning and Verbal Behavior, 23,* 23–38.

Snodgrass, J. G. (1984). Concepts and their surface representations. *Journal of Verbal Learning and Verbal Behavior, 23,* 3–22.

Received April 16, 1987
Accepted April 21, 1987 ■