

A criticism of Wilkins' (1971) measure of category size, and its implication for the Smith, Shoben & Rips (1974) model of semantic memory

Peter E. Morris

Smith, Shoben & Rips (1974) proposed a model for information storage and decision making on the semantic properties of words. This model provides a more adequate account of the nature and latencies of such decisions than any other previous theory. However, there is one result in the literature which appears incompatible with their theory and which they were unable to conclusively fault. This is the finding by Wilkins (1971) that it takes longer to make a judgement that an item is an instance of a category, the larger the size of the category. Smith *et al.* (1974) suggested that Wilkins' measure of category size is inappropriate, but were unable to support their claim with quantifiable evidence. Such evidence is, however, available and is worth reporting since Wilkins' (1971) paper is one of the most frequently quoted of those in the area of semantic memory.

Wilkins measured category size in terms of the number of items listed as having been produced in each category in the Cohen, Bousfield & Whitmarsh (1957) category norms. These norms were derived by asking 400 subjects to list four items under each category heading. The question is, does the range of items so produced bear any relationship to the true size of the category, measured in terms of the items that would be classified or classifiable as falling into the category by ordinary people? This question is relevant not only to the validity of Wilkins' study but also more generally, since much research on semantic memory has concerned the relationship of category size and latency of retrieval of items from the category (see Johnson-Laird, 1974; Smith *et al.* 1974, for reviews). Efforts at estimating category size have involved counting the number of items given by subjects in a 60 second period (Freedman & Loftus, 1971) while other studies have attempted to manipulate category size by using as small categories those which are subsets of larger categories (e.g. mammal and animal). The former method raises problems of item accessibility while the latter has the confounded variable of the greater abstractness of the larger categories.

Was Wilkins' (1971) method an appropriate measure? A straightforward alternative method of estimating category size is to ask subjects how big the categories are. These reports can then be compared with the estimates obtained by Wilkins' method. A group of subjects can be asked to rate a set of categories on their size. Such subjective ratings in other areas of research have provided very useful and reliable estimates (e.g. the ratings of Imageability and Concreteness; Paivio, Yuille & Madigan, 1968; Morris & Reid, 1972). Ratings on the size of the categories of Cohen *et al.* (1957) were obtained by Battig & Montague (1969) when they were generating their category norms. It is therefore possible to examine the rated category size of those categories that were selected by Wilkins as being small and large. The result is surprising. The ratings were made on a seven-point scale, from 1 for very small to 7 for very large. The mean rating for Wilkins' large categories is 4.83 (range 4.12-5.62), while for his small categories it is 5.29 (range 4.59-5.83). It is clear that the rated size of the categories is opposite to that designated by Wilkins. What is more, the difference is actually significant at the 5 per cent level on a one-tailed *t* test ($t = 1.989$, *d.f.* = 14) even though only two sets of eight categories are involved. It seems that, given that the subjects can be trusted to know what they are talking about, Wilkins actually selected small categories when large ones were intended, and vice versa.

Does this mean that when subjects are asked to give just a few examples from a large category that they are more consistent with one another in the example that they give than when they give examples from a small category? This would be an interesting finding, but the answer seems to be no. If the size of the 43 categories of Cohen *et al.* as measured by Wilkins, is