

Reliability of retrieval from semantic memory: Information about people

FRANCIS S. BELLEZZA
Ohio University, Athens, Ohio

An estimate of the reliability with which information about specific people could be retrieved from semantic memory was made. College-aged subjects were asked to describe 10 friends and 10 famous people. They then performed the identical task 1 week later. For personal friends, they listed a mean of 8.02 attributes, and for famous people, 5.72 attributes. The estimate of recall reliability using the common-element correlation was .38 for friends and .55 for famous people. The between-subjects correlation for attributes of famous people was .21. Reasons for the modest reliability values, such as the inadequacy of the prompts and the complexity of the responses, are discussed.

Previous research has shown that the retrieval of information from semantic memory attains only modest levels of reliability. Bellezza (1984a) found that, in two tests of recall separated by 1 week, about 69% of the category instances generated in the first session were also generated in the second session. In a second experiment (Bellezza, 1984b), subjects gave definitions of the same common nouns in each of two sessions 1 week apart. The proportions of propositions common to the two definitions given for each noun were .55 for concrete nouns, .46 for category labels, and .43 for abstract nouns. As in the experiment measuring the reliability of retrieving category instances (Bellezza, 1984a), these values reflect only modest reliability in the retrieval of verbal information from semantic memory.

In Bellezza (1984a, 1984b), the words used were either nouns that referred to classes of objects or nouns that referred to abstract concepts. Although the word "apple" refers to a smaller set of objects than does the word "fruit," both these words denote a class of objects. If subjects listed instances and attributes of classes of objects by first recalling some subset of referents, then the modest reliability of retrieval of these words from one session to the next could be explained by the different referents thought of during each session. This process is an example of the availability hypothesis of Tversky and Kahneman (1973). Retrieval of different referents occurred when category instances were generated in Bellezza (1984a) and may have occurred when concrete nouns were defined in Bellezza (1984b). The purpose of the present

experiment was to use labels that had only one referent and to have subjects describe these. In each of two sessions separated by 1 week, the subjects first wrote down the names of 10 friends or relatives and then described each one using phrases of one or two words. They next described 10 famous people chosen by the experimenter. Of course, it could be argued that even though each word had only one referent, the subjects could recall the person in different contexts during each session and that this would result in reduced similarity between the two descriptions. This concern is valid, but the purpose of the experiment was to determine if the reliability values obtained when information about people is retrieved from semantic memory are comparable to those obtained in earlier experiments.

METHOD

Subjects

Sixteen undergraduates enrolled in introductory psychology courses at Ohio University participated for extra course credit.

Material

A norming group of 45 subjects was first tested. The subjects were asked to write down the names of as many famous people as possible. These people could be living or dead. They were given 20 min to do this. Except for the name Ronald Reagan, the 10 most frequently mentioned names were used as the famous names in the main experiment. These were: Johnny Carson, Adolf Hitler, Martin Luther King, Abraham Lincoln, Marilyn Monroe, Richard Nixon, Elvis Presley, Robert Redford, Barbara Streisand, and George Washington.

Procedure

The 16 subjects in the main part of the experiment were tested as a group. First, they were given 3 min to write down the names of 10 people whom they knew well. These could be friends, relatives, or a mixture of both. They wrote these names in the 10 blank spaces provided in their test booklets. Next, the subjects had to describe as accurately as possible the 20 people listed in their test booklets. The first set of 10 comprised their personal friends, and the second set of 10 were the famous names arranged in a random order. The descriptions of all 20 people were limited to phrases of one and two words, as used in the phrase condition of Bellezza (1984b). Two examples were given. One was of

Portions of this research were presented at the meeting of the Psychonomic Society in Minneapolis, MN, in November 1982. This research was supported in part by a grant from the Field-Wiltsie Foundation. Thanks go to Kathy Kamin, Sylvia Sims, Rosalind Whatley, and Cathy Young for their assistance in collecting and scoring the data. Thanks also go to Ohio University Computer and Learning Services for making computer time and their facilities available. Requests for reprints should be sent to Francis S. Bellezza, Department of Psychology, Ohio University, Athens, OH 45701.

a hypothetical friend named John Smith, who was described as "friendly, active, loves baseball, likes beer, hates dancing, former Marine, business major." The second example was Ronald Reagan, who was described as "American President, conservative, defeated Carter, cuts taxes, builds defense, cuts welfare." The experimenter paced all the subjects through the task and allowed them 1 min for each name. The procedure for the second session 1 week later was the same, and the same instructions were used. The only difference was that the famous names were listed in a different random order.

RESULTS

A 2×2 analysis of variance was performed on the number of propositions used to describe each person. The two factors were type of name (friend vs. famous person) and session (Session 1 vs. Session 2). Both factors were within-subjects factors. Type of name was significant [$F(1,15)=43.28$, $p < .001$, $MSe=1.96$]. For personal friends, the subjects listed a mean of 8.02 attributes, and for famous people, they listed 5.72 attributes. The mean number of attributes given for each of the famous people is shown in Table 1. A separate analysis showed no difference among the individual famous names. The only other significant source of variation was the type of name \times session interaction [$F(1,15)=11.12$, $p < .005$, $MSe=3.11$]. For famous people, the mean number of propositions recalled during Sessions 1 and 2 were 5.78 and 5.65, respectively. However, for personal friends, the means were 8.53 and 7.51, respectively. These latter two means were significantly different according to a Cicchetti (1972) test. It may be that during Session 2 the subjects spent time trying to remember what personal friends they had described during Session 1 and how they described each one. This may have interfered with their writing of descriptions. The mean number of repetitions of friends' names was 7.89 of a possible 10.

A common-element correlation was used to determine the degree of overlap in the descriptions given in the two sessions. For famous people, this value was .55, and for personal friends who appeared on both lists, this value was .38. The difference between the two values was significant [$F(1,15)=51.18$, $p < .001$, $MSe = .005$]. The

overlap scores for the 10 famous people are shown in Table 1. A separate analysis showed no significant differences among overlap scores for the 10 famous people.

To determine the degree to which descriptions of famous people were similar among subjects, the 16 subjects were formed into eight pairs for Session 1 and a different set of eight pairs for Session 2. An overlap score for each famous person was then computed for each pair. The mean values obtained are shown in Table 1. The mean between-subject overlap score was .21, which was significantly greater than 0 [$F(1,115)=239.43$, $p < .001$, $MSe = .030$]. There was no significant effect for the different names.

DISCUSSION

The purpose of this experiment was to determine if the reliability of retrieval of information concerning specific referents, that is, specific people, would be higher than the reliabilities obtained for common nouns (Bellezza, 1984b). The answer seems to be no. The mean reliability for famous names was .55, which was approximately the same as the value of .54 found for concrete nouns by Bellezza (1984b). The mean reliability for the names of personal friends was .38, which was somewhat lower than the value of .46 found for category labels and the value of .43 found for abstract nouns. Perhaps this lower reliability value found for personal friends reflects the fact that the subjects have a greater deal of information in semantic memory about people whom they know well. Also, the subjects may have acquired new information or renewed old information regarding these people in the week intervening between the two sessions.

The mean between-subject overlap score for famous people was .21, which is somewhat lower than the value of .29 computed for the concrete nouns by Bellezza (1984b). This difference may indicate that some subjects may have had little knowledge about some of the famous names presented.

Adequacy of Prompts

The conclusion that follows from the modest reliabilities reported here and in Bellezza (1984a, 1984b) is not necessarily that all retrieval from semantic memory is unreliable. One hypothesis that can explain these results is that a single word is not an adequate prompt for retrieving information from semantic memory. Perhaps more information must be added to the retrieval cue if reliability is to improve. A number of investigators have proposed that *context* is important in comprehending and remembering (Anderson & Ortony, 1975; Barclay, Bransford, Franks, McCarrell, & Nitsch, 1974; Bransford & Johnson, 1973; Bransford & McCarrell, 1974). If subjects were presented the sentence "Dogs like Lassie and Benji seem to be as intelligent as people" and were then asked to define the word "dog," the reliability in week-to-week recall might increase.

Single-word prompts may be inadequate because many models of semantic memory imply that memory search must have two starting points. An example of a semantic-memory prompt that occurs in conversation is "Do you believe that Americans eat the right foods?" Here, retrieval takes place using the two concepts "Americans" and "food." In this case, a higher reliability in responding may occur because two concepts jointly define a smaller amount of information in semantic memory than does one (Anderson & Bower, 1973; Shiffrin, 1970). Therefore, information may be more reliably retrieved. In some semantic-network models of memory, search takes place until an intersection occurs, indicating that a decision and a response can take place (Collins & Loftus, 1975; Collins & Quillian, 1969). Searching from only one concept does not result in a well-defined retrieval process.

It should be kept in mind, however, that people often do have to retrieve information on the basis of only a single concept. Examples are "What do you think of the President?", "What do you think of dogs?", and "What do you think of modern technology?" On the basis of the evi-

Table 1
Mean Recall and Correlation Values for Friends
and Each of the Famous Names

	Number of Propositions Recalled	Within- Subjects Overlap	Between- Subjects Overlap
Names of Friends	8.02	.38	
Famous Names			
Johnny Carson	5.41	.52	.17
Adolf Hitler	5.97	.44	.20
Martin Luther King	5.47	.47	.18
Abraham Lincoln	5.69	.54	.15
Marilyn Monroe	5.56	.59	.21
Richard Nixon	5.44	.57	.23
Elvis Presly	6.81	.52	.28
Robert Redford	5.69	.58	.16
Barbara Streisand	5.84	.64	.31
George Washington	5.28	.57	.20

dence presented here and assuming that people use the availability heuristic (Tversky & Kahneman, 1973), it would be expected that people's judgments and their justifications for those judgments would vary from one occasion to the next. This is especially likely to occur if these occasions are widely separated in time.

Complexity of the Response

The reliability of retrieval from semantic memory may also increase if the amount of information required in the response is reduced. Subjects could be instructed to produce only one proposition as a definition for each presented word. Responding could be restricted even further by limiting subjects to "yes" and "no" answers in response to prompts such as "Do dogs bark?" Some data relevant to the question exist in the literature on word association. Gekoski and Riegel (1966) recorded single-word associations to single words presented as stimuli. When repeating the task 1 week later, they found only 43% of the associations to be the same. This measure of reliability is comparable to the values reported here. The results of Gekoski and Riegel suggest that reducing the complexity of the response would not greatly increase reliability. Fox (1970), using a similar word-association procedure with a 60-day interval, found 48% of the responses repeated. This result seems to indicate that the reliability values found when a 1-week intertest interval is used may be similar to the values obtained with longer intertest intervals.

REFERENCES

- ANDERSON, J. R., & BOWER, G. H. (1973). *Human associative memory*. Washington, DC: Hemisphere Press.
- ANDERSON, R. C., & ORTONY, A. (1975). On putting apples into bottles—A problem of polysemy. *Cognitive Psychology*, 7, 167-180.
- BARCLAY, J. R., BRANSFORD, J. D., FRANKS, J. J., MCCARRELL, N. S., & NITSCH, K. (1974). Comprehension and semantic flexibility. *Journal of Verbal Learning and Verbal Behavior*, 13, 471-481.
- BELLEZZA, F. S. (1984a). Reliability of retrieval from semantic memory: Common categories. *Bulletin of the Psychonomic Society*, 22, 324-326.
- BELLEZZA, F. S. (1984b). Reliability of retrieval from semantic memory: Noun meanings. *Bulletin of the Psychonomic Society*, 22, 377-380.
- BRANSFORD, J. D., & JOHNSON, M. K. (1973). Considerations of some problems of comprehension. In W. G. Chase (Ed.), *Visual information processing*. New York: Academic Press.
- BRANSFORD, J. D., & MCCARRELL, N. S. (1974). A sketch of a cognitive approach to comprehension: Some thoughts about what it means to comprehend. In W. B. Weimer & D. S. Palermo (Eds.), *Cognition and symbolic processes*. New York: Winston & Sons.
- CICCHETTI, D. V. (1972). Extension of multiple-range tests to interaction tables in the analysis of variance: A rapid approximate solution. *Psychological Bulletin*, 77, 405-408.
- COLLINS, A. M., & LOFTUS, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review*, 82, 407-428.
- COLLINS, A. M., & QUILLIAN, M. R. (1969). Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior*, 8, 240-247.
- FOX, P. W. (1970). Patterns of stability and change in behaviors of free association. *Journal of Verbal Learning and Verbal Behavior*, 9, 30-36.
- GEKOSKI, W. L., & RIEGEL, K. F. (1966). A study of the one-week stability of the Michigan restricted association norms. *Psychological Reports*, 18, 397-398.
- SHIFFRIN, R. M. (1970). Memory search. In D. A. Norman (Ed.), *Models of human memory*. New York: Academic Press.
- TVERSKY, A., & KAHNEMAN, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207-232.

(Manuscript received for publication April 20, 1984.)