# Reliability of retrieving information from knowledge structures in memory: Self information

FRANCIS S. BELLEZZA
*Ohio University, Athens, Ohio*

It has been proposed that information about the self is better organized and more salient in memory than is information about other domains of knowledge. Yet, the data reported by Greenwald, Bellezza, and Banaji (in press) suggest that the reliability with which information about the self can be retrieved from memory is not particularly high. In the present experiment, the verbal prompts used in requesting self information from subjects were made more explicit than those used by Greenwald et al. by the addition of subcategories and examples of the self information requested. Using a test-retest procedure, only .58 of the items from six self categories that were retrieved from memory during an experimental session were again retrieved 1 week later. This value is comparable to the value of .56 found by Greenwald et al. The reliability value for the two common categories, *fruit* and *fish*, included in the experiment was .76, which was significantly greater than the value found for the self categories. These results suggest that information about the self is not particularly well-organized and salient in memory.

Cognitive psychologists believe that many aspects of thought, language, and behavior are mediated by organized sets of knowledge in memory. These *knowledge structures* represent information about various real-world objects and domains and are variously identified as natural categories, mental images, memory schemas, prototypes, and so forth (Norman, 1982; Rumelhart, 1980; Wyer & Srull, 1984). An obvious prerequisite for a set of information to be considered a knowledge structure in memory is that it be well-organized, so that the information contained in it can be easily and reliably accessed and retrieved.

Bellezza (1984a) proposed a procedure by which the reliability of retrieving information from knowledge structures could be estimated. College students were presented verbal labels of knowledge structures and were requested to retrieve as much information from memory about each structure as they could. The type of information to be recalled was usually constrained in some way. For example, Bellezza's (1984a) subjects had to retrieve from memory instances of common categories such as *fruit* and *clothing*, whereas Bellezza's (1984c) subjects had to provide definitions of concepts such as *chair* and *advantage*. In Bellezza's (1984c) experiment the concepts defined included category labels such as *tree*. Enough time was provided in this procedure, usually 3 min, so that at the end of this period subjects reported that they usually could retrieve no further relevant information. The entire retrieval process was repeated 1 week later, and the reliability of retrieval was measured by the proportion of information that was the same in the two sessions.

Using this measure of the proportion of retrieval overlap, the reliabilities of retrieval determined for various types of knowledge structures has not been particularly high. When recalling instances from common categories such as animals, birds, and body parts, the mean proportion across categories and subjects was .69 (Bellezza, 1984a). When defining categories, the mean proportion was .46 (Bellezza, 1984c). When defining concrete objects, the mean proportion of overlapping information was .55, and when defining abstract nouns it was .43 (Bellezza, 1984c). When describing famous people, the mean proportion was .55. However, when describing friends, the reliability was .38 (Bellezza, 1984b). It appears that knowledge structures that represent category instances have the greatest reliability of retrieval (.69), whereas knowledge structures containing information about friends have the lowest reliability of retrieval (.38).

These low reliabilities were obtained even though the retrieval cues used were considered as optimal. That is, subjects were instructed to retrieve information from knowledge structures, which were identified using their commonly used labels. In the language community these labels appear to specify these structures without ambiguity.

It is possible, however, that the reliability of the best-organized knowledge structures in memory have yet to be tested. A number of theorists have proposed that the self-concept or self-schema represents a particularly important and well-organized knowledge structure in memory (Epstein, 1973; Greenwald, 1981; Kuiper & Rogers, 1979; Rogers, 1981). If information about the self is particularly salient and well-organized in memory, then the reliability of retrieving information about the self from memory should be very good. Yet, in a study by

Greenwald, Bellezza, and Banaji (in press), the proportion of common information generated in two sessions regarding nine categories about the self taken from McGuire and Padawer-Singer's (1976) study averaged only .52. This value is no larger than the reliability found for retrieving attributes of concrete objects (.55) or famous people (.55). Furthermore, the reliability of self information seems to be considerably less than the reliability of retrieving instances of common categories (.69).

One possible problem with the self categories tested is that some of these categories may not be adequately identified using short verbal descriptions. For example, asking subjects to provide information about facets of the self such as good qualities, bad qualities, daily activities, activities enjoyed, and so forth (Greenwald et al., in press) may not communicate effectively to the subject what type of information is to be retrieved. If the prompt is not specific enough, then the subject may not understand from what knowledge structure in memory the information should be retrieved (Bellezza, 1984b).

The purpose of the experiment reported here was to determine whether the reliability of retrieval of information about the self improves if additional subcategories and examples are provided to the subject when the self information is requested.

## METHOD

### Subjects

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

### Procedure

The procedure was that used by Bellezza (1984a, 1984b, 1984c). All subjects were tested twice in a group with the sessions separated by 1 week. Six of the nine categories of information about the self used by Greenwald et al. (in press) were arbitrarily selected. In each session the subjects were given 3 min to write down as much information as they could for each of the self categories. The materials used were similar to those of Greenwald et al. However, in addition to the question about the self stated at the top of each page in the question booklet, a set of four more specific subcategories and examples were provided with each question to make the information requested less ambiguous. The following six questions and hints were provided: (1) Write down a list of the *activities* that you enjoy. Hints: hobbies, entertainment, sports, other social events. (2) Write down the names of people you *dislike*. Hints: relatives, popular stars, politicians, acquaintances. (3) Write down what

you think are your *good qualities*. Hints: traits, virtues, special talents, attitudes. (4) Write down a list of your *daily activities*. Hints: health, thoughts, school/work, physical activities. (5) Describe yourself *physically*. Hints: looks, style, stature, appearance. (6) Write down what you think are your *bad qualities*. Hints: habits, temperament, problems, attitudes.

These questions were presented in the same order each week. In addition, the following retrieval task was used as a warm-up task presented before questions about the self: Write down the names of as many different kinds of *fish* as you can think of. Hints: sport fish, pet fish, edible fish, deep sea fish. Similarly, the following retrieval task was used at the end of the questions concerning the self to prevent selective rehearsal of the last recalled information about the self: Write down the names of as many different *fruits* as you can think of. Hints: citrus, tropical, canned, dried.

## RESULTS

Almost all subjects answered the questions using single words or short phrases. The answers of those few subjects who responded using full sentences were easily decomposed into simple idea units (Bellezza, 1984c).

### Self Categories

**Number of items recalled.** The mean number of responses recalled for each self category combined over subjects and sessions is shown in Table 1. An analysis of variance showed no significant effects of session on the number of items recalled, but there was a significant effect of self category [$F(5,145) = 78.32, p < .001, MSe = 16.52$]. A Tukey HSD test (Kirk, 1982) indicated that any two means differing in value by more than 2.9 items were significantly different from one another. As can be seen from Table 1, the subjects recalled significantly more items for the categories *daily activities* and *liked activities* than they did for the other self categories.

In this experiment the subjects recalled more items for five of the six self categories than did subjects in the Greenwald et al. (in press) experiment. The mean number of items recalled over the six self categories was 11.0 items, whereas in the Greenwald et al. experiment the mean number recalled for these six categories was 9.4. It is possible that the additional subcategories and examples provided with the questions in the present experiment increased the amount of information retrieved from memory compared with the Greenwald et al. experiment.

## Table 1
### Recall Characteristics of Self Categories and Common Categories

| Category | Number of Items Retrieved | Within-Subject Overlap | Between-Subjects Overlap |
|---|---|---|---|
| Self Categories | | | |
| Physical Appearance | 9.5 | .67 | .09 |
| Daily Activities | 14.1 | .63 | .29 |
| Liked Activities | 18.8 | .61 | .24 |
| Disliked Persons | 7.0 | .57 | .02 |
| Bad Qualities | 7.5 | .54 | .13 |
| Good Qualities | 9.1 | .46 | .12 |
| Common Categories | | | |
| Fruit | 18.9 | .80 | .60 |
| Fish | 14.6 | .71 | .30 |

**Reliability of retrieval.** The proportion of items from each self category that were the same in each session is shown in Table 1 for each category of the self. An analysis of variance showed the category of self information to be a significant factor [$F(5,145) = 6.11, p < .001, MSe = .027$]. A Tukey HSD test indicated that any two mean overlap scores differing by more than .12 were significantly different from each other. It appears that the reliability of retrieval of characteristics of physical appearance is significantly better than the reliability of retrieval of bad qualities and good qualities. Similarly, the reliability of retrieving information about daily activities and liked activities seems to be better than the reliability of retrieving information about one's good qualities.

These estimates of the reliability of retrieving information about the self from memory are quite similar to those obtained by Greenwald et al. (in press). The largest difference in overlap scores involves *bad qualities*, which had a value of .54 in the present experiment, but .44 in Greenwald et al. Nevertheless, the mean overlap score for the six self categories used in the present experiment was .58, and the mean for these same six categories in Greenwald et al. was .56. It appears that the extra hints and cues provided with the questions in the present experiment did not measurably enhance the reliability of retrieving information from memory about the self.

## Common Categories

Even though the two common categories *fish* and *fruit* were not of primary interest in this study, the performance measures obtained for these categories were similar to those obtained by Bellezza (1984a) and Greenwald et al. (in press). The mean number of instances of fish recalled was 14.6 compared to 13.7 for Greenwald et al. The mean number of instances of fruit recalled was 18.9 compared to 17.2 for Greenwald et al. and 17.8 for Bellezza (1984a). Although the addition of hints and cues in the present experiment seem to have increased recall slightly, the effect was not great.

The mean overlap score for fish was .71, whereas for Greenwald et al. (in press) this value was .72. The mean overlap score for fruit was .81, whereas for Greenwald et al. this value was also .81 and for Bellezza (1984a) it was .79. As was the case for the self categories, one may conclude that the additional subcategories and examples provided to subjects did not enhance the reliability of retrieving common category information from memory.

## Between-Subjects Reliability

Also shown in Table 1 are the mean between-subjects overlap scores computed for the six categories of the self and for the two common categories. The overlap scores were obtained by randomly pairing subjects and then determining what proportion of their responses was the same. This was done for each session, and in each session the subjects were paired in a different way. An analysis of variance showed that session was not statistically

significant, but that category was significant [$F(7,196) = 82.47, p < .001, MSe = .012$]. A Tukey HSD test indicated that between-subjects overlap scores differing in magnitude by more than .09 were significantly different. For the categories of the self, the subjects agreed most for daily activities and liked activities. This is not surprising when one remembers that all the subjects tested were college students. The between-subjects overlap value of .60 for the category of fruits is identical to that obtained by Bellezza (1984a). However, previous experiments did not compute between-subjects overlap scores for the other categories used here.

## DISCUSSION

The use of hints and cues in addition to questions requesting information from knowledge structures in memory about the self seemed to increase slightly the amount of information recalled compared with that recalled in earlier studies using the same categories (Bellezza, 1984a; Greenwald et al., in press). More important, however, the use of additional hints and cues did not increase the reliability of retrieving information from memory. Compared with earlier studies, there was no increase in the overlap scores for either the self categories or the common categories. The overlap scores were almost identical to those obtained by Greenwald et al. and Bellezza. Bellezza (1984b) suggested that the reliability of retrieving information from permanent memory might be increased by using more elaborate and specific retrieval cues, but the results of the present experiment provide no evidence to support this hypothesis.

These results and those of Greenwald et al. (in press) also suggest that information about the self is not particularly well-organized in memory, as suggested by Epstein (1973), Greenwald (1981), Kuiper and Rogers (1979), and Rogers (1981). The mean overlap score for the six categories pertaining to information about the self was .58. This value means that only 58% of the information retrieved regarding some aspect of the self during a session was also retrieved 1 week later. This value was significantly less than the overlap mean, .76, which was found for the category instances generated for the common categories of fish and fruit [$F(1,29) = 121.50, p < .001, MSe = .004$]. A greater overlap score was found for common categories, even though more responses were generated for the common categories than for most of the self categories. This is of interest because one might expect knowledge structures containing large amounts of information to be more poorly organized (Bellezza, 1984a).

The reliability of retrieval regarding information about the self, as measured by overlap scores, is comparable to the reliability of retrieving information about concrete objects (.55; Bellezza, 1984c) and about famous people (.55; Bellezza, 1984b), and lower, in general, than the mean reliability of recalling instances of a variety of common categories (.69; Bellezza, 1984a).

Bellezza (1984c) found that definitions of concrete nouns were retrieved more reliably than were definitions of abstract nouns. The reliabilities of retrieving information about the self found here also reflect the effect of concreteness. The categories of the self that represent directly perceived and experienced aspects of the self, such as physical appearance, daily activities, liked activities, and persons disliked had a mean overlap score of .62, whereas the more abstract good qualities and bad qualities had a lower mean overlap value of .50.

The frequency, contiguity, and similarity of perceived objects, people, and events may determine the nature and permanence of memory organization. Because the more abstract descriptors used in describing good and bad qualities may not be used frequently and have no direct physical-world counterparts, they may be more poorly organized in memory. In spite of the poor organization of self-knowledge evidenced here, some investigators have argued that the self as a knowledge structure is particularly useful in helping to organize and store new knowledge in memory. The relatively high level of recall following the relating of new information to the self has been called the *self-reference effect*

(Kuiper & Rogers, 1979; Rogers, 1981). Bellezza (1984d, 1986) has provided evidence that the self-reference effect is limited in its generality and has argued that, because the reliability of retrieving information about the self is not particularly high, the self as a knowledge structure does not provide a reliably retrievable set of mental cues for the acquisition and recall of new information.

One could argue that self-knowledge could not be completely or accurately obtained using the measures discussed here. The low reliabilities of self information found using these techniques might have resulted from the fact that the self information collected here was invalid, incomplete, or both. Perhaps the categories of self information were too large, too general, and the labels too ambiguous (A. G. Greenwald, personal communication, April 13, 1987). For example, rather than asking the subject to name disliked persons, one should have asked for people disliked at college, people disliked in high school, or people disliked in general, such as celebrities.

A second argument could be made along the following lines: All knowledge in memory is either declarative or procedural (Anderson, 1983), but only declarative knowledge can enter short-term memory and be verbalized (Ericsson & Simon, 1984). Therefore, if much of the knowledge of the self is procedural and is expressed in behavior that is not verbal, the procedures used in the experiment described here would not accurately reflect the reliability of retrieving information about the self from memory. What are needed are some additional performance measures. Perhaps in the future, performance measures or a measure of affect will be developed that demonstrate more reliable self information than the verbal measures used here.

A third argument against the unreliability of self information is that a single knowledge structure called the self should be rejected in favor of a working self, an on-line self, a momentarily accessible self, or what Markus and Wurf (1987) label the dynamic self-concept. To explain the results of the present experiment, one could suggest that the knowledge structures accessed by a subject on two separate occasions represented two different, but related, facets of the self.

## REFERENCES

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

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: Information about people. *Bulletin of the Psychonomic Society, 22,* 511-513.

Bellezza, F. S. (1984c). Reliability of retrieval from semantic memory: Noun meanings. *Bulletin of the Psychonomic Society, 22,* 377-380.

Bellezza, F. S. (1984d). The self as a mnemonic device: The role of internal cues. *Journal of Personality & Social Psychology, 47,* 506-516.

Bellezza, F. S. (1986). Mental cues and verbal reports in learning. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 20, pp. 237-273). New York: Academic Press.

Epstein, S. (1973). The self-concept revisited: Or a theory of a theory. *American Psychologist, 28,* 404-416.

Ericsson, K. A., & Simon, H. A. (1984). *Protocol analysis: Verbal reports as data.* Cambridge, MA: MIT Press.

Greenwald, A. G. (1981). Self and memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 15, pp. 201-236). New York: Academic Press.

Greenwald, A. G., Bellezza, F. S., & Banaji, M. R. (in press). Is self-esteem a central ingredient of the self-concept? *Personality & Social Psychology Bulletin.*

Kirk, R. E. (1982). *Experimental design: Procedures for the behavioral sciences* (2nd ed.). Belmont, CA: Wadsworth.

Kuiper, N. A., & Rogers, T. B. (1979). The encoding of personal information: Self-other differences. *Journal of Personality & Social Psychology, 37,* 449-514.

Markus, H., & Wurf, E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology, 38,* 299-337.

McGuire, W. J., & Padawer-Singer, A. (1976). Trait salience in the spontaneous self-concept. *Journal of Personality & Social Psychology, 33,* 743-754.

Norman, D. A. (1982). *Learning and memory.* San Francisco: W. H. Freeman.

Rogers, T. B. (1981). A model of the self as an aspect of the human information processing system. In N. Cantor & J. F. Kihlstrom (Eds.), *Cognition, social interaction, and personality* (pp. 193-214). Hillsdale, NJ: Erlbaum.

Rumelhart, D. E. (1980). Schemata: The building blocks of cognition. In R. Spiro, B. Bruce, & W. Brewer (Eds.), *Theoretical issues in reading comprehension* (pp. 33-58). Hillsdale, NJ: Erlbaum.

Wyer, R. S., Jr., & Srull, T. K. (1984). *Handbook of social cognition* (Vol. 1). Hillsdale, NJ: Erlbaum.