

The Effects of Substituting "Objects" for
"Forms" on the Revised Token Test (RTT)
Performance of Aphasic Adults

Lorraine Lohmann
Presbyterian Medical Center

Thomas E. Prescott
Veterans Administration Hospital
Denver, CO

At the 1976 Clinical Aphasiology Conference a discussion issue was raised following a paper by Dr. Berry, concerning the adequacy of using geometric forms for testing auditory performance. Following that discussion, we decided to test whether or not performance differences exist when geometric forms are used for test objects as compared to the use of so-called "real" objects for stimulus items.

The question has also been raised in the literature as to whether there is a difference in aphasic subjects' performance on certain nouns over certain other nouns. Among the two sets of nouns compared in these experiments were nouns representing the names of shapes for forms, such as circle, square, rectangle, etc., and nouns representing the names of objects that could be more readily manipulated, such as chair, table, button and box. This question is relevant to those tests employing forms for stimuli.

In 1966 Goodglass, Klein, Carey and Jones looked at aphasic subjects' ability both to recognize and name items from several different "semantic categories." Though the authors never defined their term "semantic category," they tested nouns of the following types: "colors," "numbers," "letters," "geometric forms" and "objects." On the auditory tasks, aphasic subjects were asked to point to the picture of a single word presented auditorally. In the second phase of the experiment the aphasic subjects were asked to name these pictures. In general, Goodglass, et al. found differences in both recognition and naming performance across the different "semantic categories." In particular, they found that aphasic subjects made fewer errors recognizing pictures associated with object names than pictures of geometric forms. However, as these authors were primarily interested in noting statistical differences among clinical types of aphasia (i.e., Broca's, Wernicke's and Amnesic), no statistical analysis was carried out on differences between word categories. It is therefore difficult to ascertain whether these differences were significant. Furthermore, the question still remained whether aphasic subjects' overall auditory comprehension of language structures would improve when object nouns were used.

A second study comparing "objects" and "forms" in an auditory recognition task was undertaken by Kreindler, Gheorgita and Voinescu (1971). Using pictures of circles and squares, as well as pictures of flowers and houses, they required their subjects to point to the picture identified by a verbal description. The description included two size, two color and two shape dimensions on a Token Test-type task. This task was designed to enable inferences to be made about aphasic subjects' relative impairment on "concrete" vs. "abstract" concepts. The notion of degree of abstractness as a variable in differential performance will be discussed later. The authors found that aphasic subjects identified flower and house descriptions with significantly fewer errors, again, apparently supporting the notion that "forms" were more difficult to identify than "objects". This study offers statistically significant evidence that aphasic adults, given two visually similar tasks with only the noun element varied, will recognize "objects" more easily. However, like the Goodglass et al. (1966) study, it investigates only auditory word recognition and therefore cannot be used to verify this phenomenon in auditory comprehension of language structures beyond single words.

The first study comparing performance on "objects" vs. "forms" which looked at both auditory recognition and language comprehension was that of De Renzi and Vignolo in 1962. Their Token Test consisted of a series of auditory commands of varying length and linguistic complexity, to be performed on an array of tokens. In their early testing, De Renzi and Vignolo (1962) found that, although adjectives occurred one and one half times as often as nouns on the TT, aphasic subjects made one and one half times as many errors on the nouns. They were curious whether it might be due to "the peculiar type of noun employed," and they designed an experimental version using common objects such as "pencils, thimbles, toothbrushes, and so on, of five different colors." They found that aphasic subjects did better on these nouns than on form nouns, apparently confirming their hypothesis. From these results they went on to speculate that "pencil" was easier than "circle" because of the kind of meanings that may be associated with these words.

Although De Renzi and Vignolo demonstrated a differential performance on "objects" vs. "forms" on language comprehension task, their results are inconclusive for two reasons: 1) they provide no statistical analysis of their results; and 2) the two tasks, the TT and the experimental TT, were not equivalent. In the original TT, two different shaped tokens in two sizes and five colors were used. In the experimental version, the authors used several (probably four) objects in five colors. No mention was made of size. As yet, the role that visual processing plays in TT performance has not been investigated. However, it is reasonable to assume that visual recognition of an item among circles and squares in five colors and two sizes will differ from recognition of an item from four different objects (of various shapes) in five colors. Visual similarity as a variable will be discussed in more detail later.

In 1976 Martino, Pizzamiglio and Razzano proposed a version of the Token Test based primarily on the substitution of "objects" for "forms". They called it "A new version of the Token Test for aphasics: A concrete objects form" (hereinafter, referred to as the TTC). The TTC used five objects (envelope, cup, comb, brush, pencil) in two colors (black and white) and two sizes (large and small), and was administered in the same fashion as the TT. Aphasic adults demonstrated significantly higher mean scores on the TTC while non-aphasic brain-damaged and normal adult mean scores showed no significant improvement.

Though the above results seem to support the notion that differences exist on language comprehension tasks, this study also failed to control for the visual dimensions of the two arrays. Instead of substituting two objects in five colors for the two forms in five colors, the authors substituted five objects in two colors, keeping only the size dimension constant. This again raises the question whether performance improved because "envelope, cup, comb, brush, pencil" were "easier" to understand or just easier to see and pick out from an array, especially when there were more objects than colors from which to choose.

The variables which have been claimed to influence differential performance on "objects" vs. "forms" were, in most of the studies cited, discussed after testing. Since these variables were either not explained or not controlled for, their reported influence on aphasic subjects' auditory comprehension abilities amounts to speculation. Several variables will be reviewed.

1. Abstractness/Concreteness

Several of the authors previously cited who observed differences between "objects" and "forms" speculated that the principal difference lay in the relative abstractness of each noun set and the ability of the aphasic subject to handle it. However, none of these studies adequately defined the term "abstract". Furthermore the assumption was made that "forms" were abstract and "objects" were concrete. The conclusions that aphasic individuals do better with the concrete because they did better with objects are therefore questionable.

Gardner (1973) questioned the feasibility of testing abstractness as a variable in relative performance. He maintained that abstractness is difficult both to define and operationalize. He further maintained that it is therefore difficult to explain differences in performance with one set of nouns over others in terms of relative impairment of abstraction abilities. Several attempts have been made to derive an operational definition of "abstractness" with the interest directed toward observing differences among normals (De Renzi and Vignolo, 1962; Martino et al., 1976; Kreindler et al., 1973). In general, this research has first assigned operational definitions to the terms "abstract" and "concrete", then employed a panel of judges to determine which words belonged to each category.

Gorman (1961) made one of the first attempts at classifying words along the abstract/concrete dimension. In doing this, he sought to test recognition memory for abstract and concrete words in normals. He proposed operational definitions for two judges to classify. Then, using the list of words thus derived on a short-term memory task, he tested normals and found that abstract nouns were significantly less well remembered. Gorman also found that word frequency was a variable in recognition, and that it was independent of abstractness.

Nahinsky, Slaymaker, Aamiry and O'Brien (1973) offered another operational definition of "concrete". They developed their definition when testing for the occurrence of different processing strategies in normals. These authors defined "concrete" as follows:

An attribute or dimension is considered concrete to the extent that its values can be directly derived from a visual representation of the stimulus (Nahinsky et al., 1973, p. 307).

They continued, using "forms" in their example:

Thus, observing "two black circles," gives the S direct information about the shape, shading and number of figure attributes. Such attributes ... (are)... concrete (Nahinsky et al., 1973, p. 308).

Not only do these authors define concrete attributes as those attributes which are easily visualized, but they actually use forms as examples of concrete stimuli. This study raises questions about the "a priori" assumptions of previous investigators that forms are abstract.

Paivio, Yuille and Madigan (1968) refined Gorman's classification process by developing values for "relevant attributes" of 925 nouns. Twenty eight judges used a 7 point rating scale of concreteness. Where 7 represented the most concreteness, "circle" scored a 6 and "square", a 5.7. Their results confirm what the various definitions of "concrete" previously mentioned suggested, i.e., that forms are conceptually concrete. Therefore, the entire dimension of concreteness/abstractness may not be a useful one in describing differences between "forms" and "objects".

2. Word Frequency

Frequency of a word's occurrence in English has been verified as a variable in the naming ability of aphasic individuals by numerous authors including Rochford and Williams (1965); Oldfield and Wingfield (1965); Halpern (1965);

Goodglass, Hyde and Blumstein (1960); and Gardner (1973). In general these studies found that the higher the frequency of a word, the fewer errors aphasic individuals made on it, and that in spontaneous speech, they tended to use high frequency words more than low frequency words. However, there has been little work done on the influence of word frequency upon recognition. Gorman (1961), working with normals, found a negative correlation between word frequency and recognition. Gardner (1973), in testing for the influence of "operativity" on naming, manipulated word frequency while looking at recognition of certain items and found that word frequency may not be as strong a factor in recognition (and perhaps comprehension) as it is in naming.

Though word frequency has not been tested or controlled for in most of the recognition and comprehension tasks reported, it is reasonable to assume that, since word frequency operates as a variable on naming tasks as well as on recognition tasks for normals, it may also influence the recognition and comprehension performance of aphasic adults. The question of its influence needs further investigation.

3. Operativity

In an attempt to determine the contribution of a number of variables to naming facility in aphasia, Gardner (1937) identified and defined "operativity." His study grew out of informally observing that aphasic adults had greater success naming items that could be handled or manipulated in some way. He sought an explanation for this phenomenon in the writings of Piaget:

... individuals construct their knowledge of the world by acting upon the objects in the world, by manipulating, ordering and otherwise transforming them (Gardner, 1973, p. 213).

He further stated that:

Objects, which can be readily "operated upon" will be known through a variety of actions and sensory modalities (Gardner, 1973, p. 213).

He tested the contribution of "operativity" by presenting four pictures, each with a variety of items for aphasic adults to name. The items were designated "operative," i.e., "discrete and readily manipulated," or "figurative," i.e., "equally picturable but not readily manipulated."

Gardner found that when aphasic patients were given a multiple-choice identification task on items that they had had difficulty naming, they made only 44 more errors than normals. Gardner concluded that, though certain items may be difficult for aphasic adults to name, it does not mean they will have difficulty understanding or identifying them. This suggested to him that "operativity" may not be as strong an influence on comprehension skills as it is on naming. The selection of "objects" in the current study was based in part on the items' relative "operativity."

4. Visual Array

Another factor that may be related to the differences demonstrated in previous studies is the equivalence of the visual array of stimulus items. Though the tasks being discussed here are primarily tests of auditory comprehension, on each test the patient is required to visually recognize the stimulus heard.

Two of the experimental tests discussed (De Renzi and Vignolo, 1962; Martino, Pizzamiglio and Rezzano, 1976) in fact, are not equivalent to the original TT visually, though auditorally they presumably differ only in the substitution of "objects" for "forms." Since the questions being asked by each had to do with comparative performance on "objects" vs. "forms," the two forms should have been substituted with two objects, preserving the relationships of all other visual variables--five colors two sizes and even two shapes.

In view of the possible influence of operativity, word frequency and similarity of visual arrays, it was the purpose of this study to reexamine whether aphasic adults demonstrated statistically significant improvement on auditory comprehension scores when "operative objects" in this case (buttons and boxes) were substituted for "figurative forms" (circles and squares). To accomplish this, ten left-hemisphere brain-damaged adults were each administered two versions of the Revised Token Test (McNeil and Prescott, 1973). One version, which we will call RTT Objects, employed buttons and boxes. The study was designed to determine whether aphasics' auditory comprehension improved significantly when "operative" objects (buttons and boxes) were used in place of "figurative" test stimuli (circles and squares). Word frequency and size, shape color and array of the tokens were controlled and reasonably equated between the two versions.

The ten subjects were given the RTT Forms and the RTT Objects. The order in which subjects received each test was counterbalanced. Five subjects were randomly selected to receive the RTT Forms first, the other five were first administered the RTT Objects. Each of the subjects was then given the other test version within ten days of completing the first.

Four measures of performance which are standard to the RTT were derived: overall scores, individual subtest scores, scores on the noun element (i.e., shape words in the commands) and overall test time. Overall and individual subtest scores were then converted to percentile equivalents as prescribed and provided for by the RTT.

The following measures of performance were statistically compared: overall percentile scores, individual subtest percentile scores, raw scores on the noun element, and overall test times. An analysis of variance with repeated measures (Winer, 1962), applied to each of these multidimensionally scored measures, revealed no significant differences between the scores obtained for forms vs. objects. Table 1 summarizes these results.

The final two items represented on Table 1 show two additional analyses. The percent total errors analysis represents mean plus/minus values for both test versions. The difference of 14.1% and 12.7% was not statistically significant at the .05 level of confidence.

Only the difference on percent of errors on the noun element (the last item in Table 1) was significant at $P < .05$. When plus/minus scoring was used, performance on the RTT Objects was significantly better than that on RTT Forms.

From this analysis it appears that there is no difference between objects and forms on all measures that are standard to the RTT, i.e., overall percentile scores, individual subtest percentile scores, raw multidimensional scores on the noun elements and overall test time. On the two measures that were added to standard RTT scoring, the following results were obtained: the percentage of total errors (when errors were simply counted and not classified multidimensionally) was not significantly different for either test condition. However, the percentage of errors on the noun element (when errors were simply counted) demonstrated a significant difference between the two test conditions. The means on this measure were as follows: 15% of the nouns on the RTT Forms contained errors, while 8.9% of the nouns on the RTT Objects contained errors, a difference of 6.1%, significant at $P < .05$.

TABLE 1
GROUP MEANS FOR MEASURES OF PERFORMANCE

Measure of Performance	Condition I (RTT Forms)	Condition II (RTT Objects)
Overall Percentile Scores	47.4	47.4
Subtest Percentile Scores:		
Part 1	52	57.5
Part 2	46.3	50.6
Part 3	43.2	46.7
Part 4	41.3	39.7
Part 5	46.6	48.9
Part 6	51.7	50.8
Part 7	49.3	48.1
Part 8	53.1	54.1
Part 9	50.7	52.5
Part 10	51.6	56.9
Raw Scores on Noun Element I	11.96	12.3
Raw Scores on Noun Element II	11.64	11.95
Overall Time (minutes)	35.7	32.7
% Total Errors	14.1	12.7
% Noun Element Errors	15	8.9 *

* Significant at $P < .05$

These results suggest that when word frequency and the size, shape, color and array of tokens are equated, the use of "objects" in place of "forms" does not significantly improve the RTT performance of aphasic adults when multidimensional scores are used. Furthermore, under the same control conditions, when total number of errors is tallied (as is done in some versions of the original TT), aphasic subjects' performance is not significantly improved.

This study has also demonstrated, however, that under the controlled conditions mentioned above, use of "operative" objects can significantly reduce the group mean number of errors which occur on the noun element (6.1%) when plus/minus scoring is used for these elements. This result suggests several things. First, it suggests that the type of scoring used can influence whether differences are observed, at least on the noun element. In this case, the tally method of scoring revealed a difference, whereas the multi-dimensional scoring system did not. This result, therefore, may be relevant to tests employing the tally scoring method even though differences in noun element errors between objects and forms were small (6.1%), even for the tally method.

It is our observation that these findings do not support the contentions of previous investigators that "objects" improve performance on tests of auditory comprehension like the TT, at least when other variables are controlled. Furthermore, the results of this study do not support the argument that grew out of the belief that objects were easier than forms and that results of tests of auditory comprehension using objects are consequently more "clinically useful" or are a better indicator of "functional" abilities.

The findings of this study support Gardner's observation that "operativity", at least by itself, does not significantly influence aphasic subjects' comprehension abilities as it does their naming abilities. This study does, however, raise questions about what variables, uncontrolled for in the earlier studies of this type, actually contributed to the differences observed. As mentioned earlier, word frequency and visual variables, such as the array of items, may also greatly influence relative performance. However, the influence of these variables has not been satisfactorily investigated. Another possibility is that two or more of these factors (word frequency, visual array and operativity) may have interacted in earlier experiments to produce the results obtained.

Finally, it has been suggested that the level of abstractness of the nouns used could be influencing aphasic patients' performance (De Renzi and Vignolo, 1962; Martino et al., 1976; Kreindler et al., 1971). However, this view was speculation, and when subjected to psychological definitions and rating scales for abstractness, "form" names were found to be as "concrete" as "object" names.

REFERENCES

- Benton, A. Problems of test construction in the field of aphasia. Cortex, 3, 32-53 (1967)
- Berry, W. A psychometric reconstruction of the Token Test. Proceedings of the Third Conference on Clinical Aphasiology. Albuquerque (1973).
- Berry, W. Testing auditory comprehension in aphasia: a clinical alternative to the Token Test. (Unpublished).
- Brookshire, R. Differences in auditory processing deficits among aphasia patients. Paper presented to the 48th Annual Convention of the American Speech and Hearing Association. San Francisco (1972).
- Clark, H., Carpenter, P. and Just M. Semantics and perception. In W. Chase (ed.) Visual Information Processing. New York: Academic Press, Inc. (1973).
- Cronbach, L. Essentials of Psychological Testing. New York: Harper and Row (1972)
- De Renzi, E. and Vignolo, L. The Token Test: a sensitive test to detect receptive disturbances in aphasia. Brain., 85, 665-678 (1962). In M. Sarno (ed.). Aphasia: Selected Readings. New York: Appleton-Century-Crofts (1972), pp. 320-329.
- Gardner, H. The contribution of operativity to naming capacity in aphasic patients. Neuropsychologia, 11, 213-220 (1973).
- Goldstein, K. Language and Language Disorders. New York: Grune and Stratton (1948).
- Goodglass, H., Hyde, M. and Blumstein, S. Frequency, picturability and availability of nouns in aphasia. Cortex, 2, 74-89 (1966).
- Goodglass, H., Klein, B., Carey P. and Jones, K. Specific semantic word categories in aphasia. Cortex, 2, 74-89 (1966).
- Gorman, A. Recognition memory for nouns as a function of abstractness and frequency. Journal of Experimental Psychology, 61, 23-29 (1961).
- Halpern, H. Effects of stimulus variables on dysphasic verbal errors. Perceptual Motor Skills, 21, 291-298 (1965).
- Kreindler, A., Gheorgita, N. and Voinescu, I. Analysis of verbal reception of a complex order with three elements in aphasics. Brain., 94, 375-389 (1971).
- McNeil, M. and Prescott, T. Assessment of auditory deficits associated with aphasia: the Revised Token Test. Proceedings of the Third Conference on Clinical Aphasiology. Albuquerque (1973).

- Martino, A., Pizzamiglio, L. and Razzano, C. A new version of the "Token Test" for aphasics: a concrete objects form. Journal of Communication Disorders, 9, 1-5 (1976).
- Nahinsky, I., Slaymaker, R., Aamiry, A. and O'Brien, C. The concreteness of attributes in concept learning strategies. Memory and Cognition, 1 (3), 307-318 (1973).
- Oldfield, R. and Wingfield, A. Response latencies in naming objects. Quarterly Journal of Experimental Psychology, 17, 279 (1965).
- Osgood, C. and Miron, M. (Eds.). Approaches to the Study of Aphasia. Urbana, Illinois: University of Illinois Press (1963).
- Osgood, C., Suci, G. and Tannenbaum, P. The Measurement of Meaning. Urbana, Illinois: University of Illinois Press (1957).
- Paivio, A., Yuille, J. and Madigan, S. Concreteness, imagery and meaningfulness values for 925 nouns. Journal of Experimental Psychology Monograph Supplement, 76 (1), 1-9 (January 1968).
- Porch, B. Porch Index of Communicative Ability: Theory and Development. Vol. I. Consulting Psychologist Press, Palo Alto (1967).
- Porch, B. Multi-dimensional scoring in aphasia testing. Journal of Speech and Hearing Research, 14, 776-792 (1971a).
- Porch, B. The use of homogenous subtest items in assessing aphasia. Paper presented at the 47th Annual Convention of the American Speech and Hearing Association. Chicago, Illinois (1971b).
- Rochford, G. and Williams, M. Studies in the development and breakdown of the use of names: Part VI the effects of word frequency. Journal of Neurology, Neurosurgery and Psychiatry, 28, 407-413 (1965).
- Thorndike, E. A Teacher's Word Book of the Twenty Thousand Words. New York: Gale Research Co. (1931).
- Winer, B. Simple analysis of variance with repeated measures. In Statistic Principles in Experimental Design. New York: McGraw Hill (1962), p. 105.
- Yamadori, A. and Albert, M. Word Category aphasia. Cortex, 9 (1), 112-125 (1973).

DISCUSSION

- Q: I'd like to raise a number of issues. One is the issue of visual imagery which would co-occur with "concreteness". You did not mention visual imagery. The second concern is what happens for instance in Gardner's study, when he compared operativity vs. figurativity. Gardner compared within the same semantic class and you've gone across semantic classes. That is, when Gardner showed the subjects a picture of the body, he has a figurative element, which was the hip, let us say, vs. an operative element which was the finger and he compares within a semantic class. It seems to me that would be an effect that you may be compounding. You may have that issue that you may have to separate out.
- A: It would have been difficult to separate it out, I think, when we were looking for a variable to control in this study. Having done away with abstractness and concreteness, "operativity" seemed to be the best candidate. It was something that we could begin to describe the difference between forms and objects by. I think using operativity as a way of looking at these things is a beginning.
- Q: And the imagery issue?
- A: In the Paivio, Yuille and Madigan study, when they looked at concreteness and had their judges rate words on this rating scale of 1-7, they also

looked at imagery, and defined it in terms of a word's capacity to arouse non-verbal images. And, again, on a rating scale of 1-7, "circle" scored a 6.23, "square" scored a 3.37.

- Q: When you used the buttons did you follow a standard format like "show the large white button". What's the typical instruction?
- A: Absolutely. We changed nothing from the standard RTT to the new RTT except "button" for "circle" and "box" for "square".
- Q: I happen to think that the Token Test is a very good test, but one of the reasons for it is that it creates problems for the person that's having to do these actions. It's basically a memory test. They have to hold those things in memory and then they have to operate on those things as they're doing them. And to me one of the major difficulties in the test for anyone taking it is holding all these words that are so highly related. There is this high association value between the colors, the high association value between the sizes, the high association value between the forms. I think that the question is not necessarily one of concreteness or abstractness but perhaps a more basic issue is the one of the load that's placed on the system by having to handle items that are so closely related, so highly associated. Now in the one area which you did suddenly get something that was not highly related, buttons and boxes, you got a change.
- A: I will address several of your points pertaining to this study: 1) The change you mention was a small one. To review, there was no significant difference between the test conditions except on the noun element alone when scores were only tallied. Even here the results were not obtained by chance but did not appear to be clinically significant. 2) If indeed the RTT is partly a memory test, subjects were given the same load in both conditions so it would have little bearing on this study. 3) Association values among tokens are difficult to define and operationalize but whether they exist or not, little difference was demonstrated. 4) Concreteness/abstractness was discussed only because it has been raised as an issue in the past. My point was that it probably isn't relevant question.
- Q: Did you say that word frequency and word length were the same in the two methods, the RTT forms and the RTT objects?
- A: The word frequency was equated.
- Q: How?
- A: Using Thorndike and Lorge (1931).
- Q: You mean circles, squares, boxes fall in the same category in Thorndike and Lorge?
- A: Yes. They were equated. But we didn't discuss word length.
- Q: Did you present those live voice and did you do all the testing?
- A: Yes.
- Q: Do you have any reliability measures in terms of speech rate?
- A: No, only on scoring.