*TESTING FOR READING COMPREHENSION:
A BRIEF REPORT ON STIMULUS CONTROL¹*

In laboratory studies of stimulus control, the experimenter arbitrarily (from the subject's point of view) designates the stimulus that is to be "relevant". But when two or more stimuli are correlated with the same reinforcer, it is difficult to predict which stimulus or stimuli the subject will attend to. (This point has been discussed in detail by Ray, 1969; Ray and Sidman, 1970; Terrace, 1966.) Consequently, it is often possible for a subject to respond apparently correctly while actually attending to some irrelevant aspect of the stimulus situation.

For example, we have observed aphasic subjects who were unable to read letters and words greatly improve their scores in matching-to-sample experiments that involved these stimuli. We found several sources of irrelevant cues for their matching performances; two will serve as examples:

We were alerted to one possibility when we observed subjects placing their finger in front of the correct comparison key of the matching-to-sample apparatus during the intertrial interval, before the stimuli had actually appeared on the keys. Therefore, we repeated the experiment without the sample stimuli, presenting the subjects only with the display of comparison letters or words on each trial. They were still able to "match" correctly. Constant repetition of the same set of matching trials had permitted the subjects to memorize the sequence of display positions in which the correct comparison stimuli appeared. They had learned to "match" correctly without even observing the sample stimuli. Sequential locations of the correct choices were controlling their behavior. If we

had not discovered this irrelevant source of stimulus control, we would have been led to erroneous conclusions about these aphasic subjects' reading abilities.

A second possibility was revealed when we found that aphasic subjects who had improved nearly to perfection in matching printed words to identical printed words and pictures to identical pictures were still incapable of matching the words to their corresponding pictures (Sidman, 1971; Sidman, Stoddard, Mohr, and Leicester, 1971). When these subjects were presented with "language symbols" and asked to do identity matching, they succeeded, but when they were asked to do nonidentity matching with these same language symbols, they failed. They were not, in fact, responding to the same stimuli in the identity and nonidentity matching tasks. When the sample and comparison stimuli were identical to each other, they could be matched on the basis of their purely physical characteristics; the subjects could respond to these stimuli merely as geometric forms. This was not possible in nonidentity matching, where the stimuli could be matched or called "same" only if the subjects had previously learned and retained the arbitrary correspondences. Thus, even though "language stimuli" are involved in both types of matching procedure, language processes need not be involved in identity matching. Any conclusions about language ability that are drawn from tests of identity matching are likely to prove misleading.

Whenever a subject is for some reason incapable of responding to the stimuli the experimenter has specified as relevant, other stimuli in the situation will come to control his behavior. Thus, we find retarded subjects, brain-damaged subjects, and even normal subjects who have been inefficiently taught, responding not to experimenter-designated relevant stimuli,

¹This research was supported by Grants HD05124 and HD04147 from the National Institute of Child Health and Human Development, and by Grant NS03535 from the National Institute of Neurological Diseases and Stroke.

but instead, to stimulus positions, stimulus sequences, reinforcement patterns, and irrelevant aspects of compound stimuli. This has been the factor underlying the many reports of "error factors", "error patterns", and "hypotheses" (Cumming and Berryman, 1965; Harlow, 1959; Levine, 1965; Sidman and Stoddard, 1967). The extraneous sources of stimulus control may actually be quite complex, sometimes more so than the relevant stimuli, even with severely handicapped subjects. It is necessary to determine by direct testing that the subject is in fact responding to the stimuli that have been specified as relevant by the experimenter.

It is equally necessary in clinical testing to ensure that the patient is responding to the stimuli that the tester presumes to be his target. The present report illustrates the necessity of maintaining the same rigorous specification of stimulus control in applied testing that has been shown to be necessary in laboratory experiments. In this demonstration, we mimicked the type of laboratory situation in which the subject is incapable of responding to the stimuli specified by the experimenter. We presented our subjects with a set of multiple-choice questions that had been designed to determine whether they understood the content of several paragraphs of text. They were given the questions, however, without having had the opportunity to read the relevant paragraphs. They were incapable of responding to the actual paragraphs simply because they had never seen them. Correct answers to the multiple-choice questions, therefore, would have to be based on stimuli other than the relevant reading material.

The demonstration will, perhaps, gain additional relevance from the fact that the multiple-choice questions were taken directly from a widely used aphasia test.

METHOD

Subjects

Five females and 10 males, ranging in age from 18 to 59 yr, comprised the group of 15

subjects. All were coworkers or acquaintances of one of the experimenters. At the time of testing, 14 subjects had matriculated at a college or a university for a period of at least 1 yr, and one had completed high school. None of the subjects had any diagnosis of neurological disorder.

Procedure

The subjects, tested individually, were given a five-page mimeographed extract of the silent reading comprehension questions contained in Eisenson's *Examining for Aphasia* (1954). Each subject was told to see how many of the questions he or she could answer without first having read the stories to which the questions pertained. The subjects were also given the following written instructions: "Read the questions, decide which answer you think is correct, and underline your choice". Although each subject was given the opportunity to ask questions, none did so. Each was allowed as much time as was necessary to complete the task. At no time was the questionnaire introduced or specified as a measure of the subject's intelligence, nor was it indicated from what text the questions were obtained. All subjects volunteered and none was paid.

The five sets of multiple-choice questions, each set taken from a different paragraph of text, appear in Table 1 (the subjects never viewed the actual paragraphs).

RESULTS

Figure 1 shows the number of subjects who selected each multiple-choice alternative in each question. The solid black bars indicate the "correct" choices.

In all items except Question a of Paragraph 4, more subjects selected the correct choice than any of the others. Of the 20 questions, 17 were answered correctly by at least two-thirds of the subjects, in spite of their never having seen the paragraphs upon which the questions were based.

Table 1

Multiple-Choice Questions Designed to Measure Reading Comprehension
(Adapted from Eisenson, 1954)

PARAGRAPH 1

a. Who was very tired?	1. Bill	2. Ted	3. Both	4. No one
b. What did Bill want to play?	1. Cards	2. Piano	3. Ball	4. Nothing
c. What did Ted do in the morning?	1. Work	2. Play	3. Swim	4. Hunt

PARAGRAPH 2

a. Where did Jack go?	1. To his grandfather's house	2. Fishing with his father
	3. Into the water	4. To the ocean to fish
b. What did the fish do?	1. Helped	2. Pulled back
	3. Ate	4. Got away
c. What did Jack try to do?	1. Go swimming	2. Put fish in the brook
	3. Pull his father back	4. Land his own fish
d. Jack enjoyed his trip because	1. The fish got away	2. He caught a fine fish
	3. The brook was near his grandfather's	4. He went to a stream

PARAGRAPH 3

a. What is the most valuable thing that St. Bernard dogs do?	1. They climb over snow drifts	2. They are good mountain climbers
	3. They rescue lost travelers	4. They are large and strong
b. The St. Bernard dog is	1. Cowardly	2. Speedy
	3. Courageous	4. Rough
c. How does the dog assist worn-out travelers?	1. By digging large holes in the snow	3. By bringing first-aid
	2. By climbing the drifts to the travelers	4. By his great strength
d. What do the dogs do for the travelers they cannot help?	1. Stand the man on his feet	2. Give him food and medicine
	3. Return to the village for aid	4. Carry him down the mountain

PARAGRAPH 4

a. Where is irrigation carried on?	1. Where crops are grown in the summer	3. Near cities
	2. Where mountain streams make it possible	4. Where there are crop failures
b. Where irrigation is possible, crop failures are	1. Increased	2. Reduced
	3. Impossible	4. Permitted
c. The best title for this story would be	1. Damming up Streams	2. Supplying Water for Irrigation
	3. Crops in Dry Weather	4. Building Huge Dams
d. Land that is irrigated yields better harvests because	1. In summer it gets little rain	2. It is high up in the mountains
	3. Water may be applied as needed	4. The water can be turned off

PARAGRAPH 5

a. In what part of the country were few manufactured articles produced?	1. North	2. East	3. South	4. West
b. From whom did American manufacturers want protection?	1. Farmers	2. American labor		
	3. Foreign manufacturers	4. Chinese labor		
c. The people from what part of the country were most anxious to get manufactured goods as cheap as possible?	1. North	2. East	3. South	4. West
d. Competition with foreign manufacturers lowers the price of what kind of goods?	1. Farm products	2. Raw materials		
	3. Manufactured products	4. Luxury items		
e. The best title for this paragraph is	1. Taxes	2. Manufacturing	3. Protective Tariff	4. Commerce

DISCUSSION

It is of some interest to speculate on the cues for correct responses that might have been present within the questions and answers themselves. In Paragraph 1, Question a, answers 3 and 4 are likely to be ruled out because Questions b and c concern Bill and Ted. Of the two, Ted was the more likely to be very tired, because Question b indicated that Bill wanted to play. In Question b, the assumption that Bill and Ted were boys would have made answer 3 the most likely. The reply to Question c could have been based on the answer that had been given for Question a; if Ted was very tired, then it is likely that his morning activity had been work.

In Paragraph 2, Question a, the choices could easily be narrowed down to 2 and 4 because of the many references to fishing in the later questions; answer 4 could be eliminated because of the references to brook and stream in Questions c and d, thereby eliminating the ocean as a sensible choice. In Question b, 2 and 4 were the only sensible alternatives, and in the absence of any cue to differentiate between these, the subjects distributed their choices almost evenly between them. Given the many references to fishing, the only answer that made sense in Question c was 4. Similar reasoning would have led to the overwhelming selection of answer 2 for Question d.

The most consistently correct answers were

Fig. 1. Number of subjects who selected each alternative answer to all multiple-choice questions about each paragraph. Letters designate the questions pertaining to each paragraph, and numbers designate the four responses to each question, as listed in Table 1. Solid bars indicate "correct" choices.

given to the questions about Paragraph 3. These answers would be obvious to anyone who knew anything about St. Bernard dogs, even though they had not read the paragraph.

In Paragraph 4, Question a, the second and third answers could easily be eliminated as the least sensible. In the absence of the relevant paragraph, however, most of the subjects selected alternative 4. Once the first question had indicated the topic of the missing paragraph, however, there remained little doubt about the correct responses to the remaining three questions.

Again, in the questions about Paragraph 5, the questions themselves indicated the topic of the missing paragraph, and most of the answers then became obvious.

Clearly, correct answers to most of the 20 questions need not have been under the control of the presumably relevant paragraphs of text. Cues to the correct answers were present in the questions and answers themselves.

It might be argued that the superfluity of the reading paragraphs does not invalidate this test as an instrument for evaluating reading comprehension. The subject had to be capable of relatively sophisticated comprehension ability in order to deduce the correct choices without having read the paragraphs. It is certainly true that a person with impaired reading comprehension would not have done nearly as well as the normal subjects of the present experiment, and their low scores would have pointed to difficulty in this area.

Nevertheless, it is difficult to justify the present format of the test, which is explicitly stated to make use of the paragraph texts. Furthermore, the paragraphs are considered to be graded in difficulty and thereby to give at least a semi-quantitative measurement of impairment. Although the paragraphs themselves do appear, on an intuitive basis, to be graded in difficulty, the data from our subjects indicate that such a progression does not exist with respect to the extraneous cues. It does not appear possible to justify either the assertion that the questions

measure a subject's ability to read the paragraphs or that the items are graded in such a way as to produce a quantitative estimate of comprehension deficit.

Although aphasic subjects who were allowed to read the paragraphs might have used them, or attempted to do so, in answering the questions, the presence of cues extraneous to the paragraphs precludes accurate identification of the controlling stimuli. The rule that stimulus control requires rigorous specification holds true not only in experimental studies, where it may lead to incorrect conclusions about basic behavioral processes, but also in tests of human capabilities, where it may lead to incorrect diagnosis, classification, and even treatment.

Murray Sidman
*Northeastern University and
 The Eunice K. Shriver Center
 for Mental Retardation
 Walter E. Fernald State School
 2000 Trapelo Road
 Waltham, Massachusetts 02154
 Martha Willson-Morris
 Department of Psychology
 Northeastern University
 Boston, Massachusetts 02115*

REFERENCES

- Cumming, W. W. and Berryman, R. The complex discriminated operant: studies of matching-to-sample and related problems. In D. I. Mostofsky (Ed.), *Stimulus generalization*. Stanford: Stanford Univ. Press, 1965. Pp. 284-330.
- Eisenson, J. *Examining for aphasia*. (Rev. ed.) New York: The Psychological Corporation, 1954.
- Harlow, H. F. Learning set and error factor theory. In S. Koch (Ed.), *Psychology: a study of a science*, Vol. 2. New York: McGraw-Hill, 1959. 492-537.
- Levine, M. Hypothesis behavior. In A. M. Schrier, H. F. Harlow, and F. Stolz (Eds.), *Behavior of non-human primates*, Vol. 1. New York: Academic Press, 1965. 97-127.
- Ray, B. A. Selective attention: the effects of combining stimuli which control incompatible behavior. *Journal of the Experimental Analysis of Behavior*, 1969, 13, 539-550.
- Ray, B. A. and Sidman, M. Reinforcement schedules and stimulus control. In W. N. Schoenfeld (Ed.), *The theory of reinforcement schedules*. New York: Appleton-Century-Crofts, 1970. 181-214.
- Sidman, M. The behavioral analysis of aphasia.

- Journal of Psychiatric Research*, 1971, 8, 413-422.
- Sidman, M. and Stoddard, L. T. The effectiveness of fading in programming a simultaneous form discrimination for retarded children. *Journal of the Experimental Analysis of Behavior*, 1967, 10, 3-15.
- Sidman, M., Stoddard, L. T., Mohr, J. P., and Leicester, J. Behavioral studies of aphasia: methods of investigation and analysis. *Neuropsychologia*, 1971, 9, 119-140.
- Terrace, H. S. Stimulus control. In W. K. Honig (Ed.), *Operant behavior: areas of research and application*. New York: Appleton-Century-Crofts, 1966. 271-344.
- Received 14 September 1973.
(Revised 18 December 1973.)
(Final acceptance 23 January 1974.)