BEHAVIORAL STUDIES OF APHASIA:
METHODS OF INVESTIGATION AND ANALYSIS*

M. Sidman, L. T. Stoddard, J. P. Mohr and J. Leicester
Behavior Laboratory, Neurology Service, Joseph P. Kennedy, Jr. Memorial Laboratories
Massachusetts General Hospital, Boston, Massachusetts 02114, U.S.A.
(Received 26 October 1970)

Abstract—Methods for testing and analyzing behavioral deficits in aphasia were derived from
a functional analysis of stimulus-response relations (the stimulus control of behavior). Deficits
were classified as breakdowns of controlling stimulus-response relations, and profiles of
deficient and intact performances in a severely aphasic patient illustrated the fruitfulness of
the methods. A distinction between identity and nonidentity tasks emerged from the data,
cutting across the different responses, stimulus modalities, and stimulus materials. Separate
examination of these two types of tasks made possible a finer classification of the patient's
deficits into input, output, or relational categories, and each of these into specific types. The
data also demonstrated that repeated testing over a prolonged time may reveal changes in deficit
classification that would have been unsuspected on the basis of early tests only.

INTRODUCTION

Aphasics, despite its clinical prominence and vast literature, has achieved only rudimentary
agreement on methods for examining patients and classifying their deficits. One source of
confusion, stressed by Jackson [1], has been the uncritical mixing of psychologic and phsio-
logic or anatomic concepts. This distinction has, unfortunately, led some workers to advo-
cate the primacy of one or the other method of classification. Brain [2] has provided an
excellent history of the controversy.

Although the psychology-physiology distinction is valid, the problem is not to determine
which is primary. Observed aphasic deficits are behavioral, but they result from anatomical
lesions and deranged physiology. Correlations of lesions with behavioral deficits, however,
will be exercises in futility if either set of observations is improperly controlled or has no
consistent classificatory scheme of its own. The behavioral examination and classification
must be at least as rigorous as the physiologic and anatomic if the correlation is to be
meaningful.

This paper describes behavioral techniques for examining, analyzing, and classifying
certain aphasic deficits, illustrating the methods through a relatively comprehensive exa-
mination of one patient. Two basic behavioral concepts, reinforcement and stimulus
control, underlie the methodology.

Reinforcement

A powerful empirical principle is that consequences govern behavior. Behavior is likely
to recur if it produces reinforcing consequences (for example, food); behavior that con-
nositently produces no reinforcement will not change (no learning), or will decline. Typically, a response is controlled by several consequences, each with its own modulating factor (deprivation, lesions, etc.).

Human behavior often produces intermediate consequences like money. The intermediate consequence is a common element in behavioral pathways to many different reinforcers. A person may be satiated with food, drink, and sex, yet money will remain a controlling consequence because it can still lead to various comforts, amusements, etc. Because their reinforcing effectiveness does not depend completely on any single modulating factor, behavioral consequences like money are called generalized reinforcers [3, 4].

One may reasonably question the validity of tests that fail to provide reinforcers for the patient. Personal satisfaction or the examiner's approval may suffice, but many patients merely go through the motions or even react to tests with disdain. To avoid this problem, we provided immediate generalized reinforcement, money, for each of the patient's correct responses.

Geschwind [5] related disturbances of higher function in animals to lesions which destroy the controlling relation between stimuli and reinforcement. He noted, however, that, "... the types of sensory-rhinencephalic linkages seen in subhuman primates are less common in man ..." ([5], p. 269). A reason may be the pervasiveness of generalized reinforcers in the control of human behavior. If a lesion destroys the relation between behavior and a particular reinforcer, generalized reinforcers will remain effective and behavior may seem undisturbed. The breakdown of relations between behavior and generalized reinforcers is a separate category of deficit [6], and although not involved in the present study must have a place in any comprehensive classification.

**Stimulus control**

A response usually produces reinforcement only when a particular stimulus is present. Receptor orientation and the less clearly defined processes of selective attention determine which elements of complex stimuli exert behavioral control. The role of selective attention in behavioral deficit will be clarified in this report.

The patient's experience with the stimulus-response relations being studied, physical and social features of the test environment, and the instructions he receives will influence his performance throughout the examination. Such factors modulate stimulus control [7]. Automated test procedures and immediate reinforcement of the patient's correct responses give him reliable nonverbal instruction about the adequacy of his performance, and help circumvent his particular social and educational background.

Since a modulating factor controls many stimulus–response relations, a disease that changes the influence of such factors will cause widespread deficits in stimulus control. Similarly, if reinforcement is no longer effective, deficits will be widespread. To distinguish a general stimulus-control deficit from a breakdown of the response–reinforcement relation, the examiner must demonstrate that the reinforcer he is using is still effective. Also, the general stimulus-control deficit must be classified separately from breakdowns confined to restricted stimulus–response relations.

Stimulus–response relations have another critical characteristic: a single stimulus controls many responses; a single response is controlled by many stimuli. Disease need not break down all relations in which a particular stimulus or response participates. For example, S may be a printed word that controls oral reading (S–R₁), copying (S–R₂), and pointing to a picture (S–R₃). A cerebral lesion that destroys S–R₁ but leaves the other
S–R relations intact is often said to leave the patient able to read the word without knowing its significance. If the lesion leaves S–R3 intact but destroys the other S–R relations, we have Broca’s aphasia. The deficit may involve broader classes of stimuli and responses: the patient may respond normally to printed words but fail when he hears them, or his only deficiency may be oral speech in response to all stimuli.

Not all classifications make the same distinctions, but they inevitably define aphasias and agnosias by sense modalities, stimuli, and responses; the basic observations are deficient stimulus–response relations.

Different stimuli, responses, and stimulus–response relations may be affected by different variables, and in quantitatively different ways by the same variables. To clarify the factors responsible for such variability, tests must hold modulating variables constant, maintain stimulus constancy while varying the responses required of the patient, and response constancy while varying stimuli. To compare stimulus–response relations that differ in their stimuli and responses is contrary to elementary scientific good sense. It is meaningful only under special circumstances to compare S1–R1 with S2–R2; rather, the comparison should be S1–R1 vs. S1–R2 or S2–R1. This principle will be elaborated.

METHODS

The patient sat in a sound-resistant room and faced a square matrix of nine translucent windows, each 2 inches square. He performed four types of tasks: simultaneous matching-to-sample, delayed matching, oral naming, and written (or printed) naming. The stimuli to be matched, named orally, or named in writing were presented through vision, hearing, or touch.

**Simultaneous matching-to-sample**

Visual–visual matching (visual sample and visual choices). A sample stimulus was projected onto the center window of the matrix (Fig. 1-A). When the patient pressed the center window, stimuli (choices) appeared on the outer windows (Fig. 1-B). He was then to press the outer window that corresponded to the sample. If he chose correctly, chimes rang and a nickel was automatically delivered. Incorrect responses produced neither chimes nor nickel. Immediately after each choice the windows were darkened; 1.5 sec later a new sample appeared.

Auditory–visual matching (auditory sample and visual choices). Auditory samples were words: the spoken equivalents of the pictures, colors, numbers, objects, letters, and words that were used as visual stimuli. A master tape read the word onto a tape loop which repeated it approximately once every 2 seconds. The center window, although illuminated, contained no visual sample. The patient’s task was: press the blank center window to produce visual choices; then, press the outer window that had a visual stimulus corresponding to the auditory sample. Correct choices produced chimes and nickel; incorrect choices neither. After a choice, repetition of the auditory sample ceased, windows darkened, and 1.5 sec later a new auditory sample began the next trial.

Tactile–visual matching (tactile sample and visual choices). Tactile samples were small common objects, plastic letters (lower-case, 1.25 in. high; upper-case, 2 in. high), or plastic digits (2 in. high). By placing his hand into a box below the window matrix the patient could feel the samples without being able to see them. The center window was dark, but visual choices were exposed while the patient palpated the sample. After he pressed a visual choice, producing chimes and nickel if correct, the sample was withdrawn and a new one inserted.

Delayed matching-to-sample

The sample disappeared before choices appeared; the patient had to remember the sample (Figs. 1-C, D, E). If the sample was visual, the center window darkened when the patient pressed it. Auditory samples had to occur at least once and then ceased when the patient pressed the blank center window. Tactile samples were removed as soon as the patient completed his tactile exploration and pressed the blank center window.

Choices at first appeared immediately upon disappearance of the sample (zero-delay). Each time the patient chose correctly, a longer delay was interposed between removal of the next sample and appearance of choices. If he made an error, a shorter delay intervened on the next trial. The patient’s accuracy adjusted
the time over which he had to remember the sample, the delay increasing or decreasing in 4-sec steps (adjusting delay).

No stimuli were tested in delayed matching unless the patient achieved at least 75 per cent correct on simultaneous matching. A delay test ended when he chose correctly after a 40-sec delay or completed the test set without passing 40 seconds.

Oral naming

The patient had simply to say aloud the name of the sample. Free to say as many names as he wished, he had to indicate when he was finished. The final name was taken as his response. If the patient named the sample correctly, the examiner pressed a button which sounded the chimes and delivered a nickel. After incorrect responses, the only consequence was presentation of the next sample.

Written naming

The patient had to write or print the name of the sample. He had a new piece of paper for each trial; previously-written names were never visible to him. If he wrote more than one, the first was recorded as his response unless he crossed it out. When the patient handed the paper to the examiner, reinforcement and sample changing proceeded as in oral naming.

Stimulus materials

Sample and choice stimuli appear in Appendix 1. With repeated tests, a systematic rotation through four to six equivalent sets of each type prevented the patient from learning display configurations, irrelevant display features, and sequences of choices or window positions, all of which would have permitted him to achieve high scores without observing the samples.

Most sets had 20 trials; some had 18. To conserve time, a test ended after ten trials if more than seven or less than three were correct. Number and sequence of tests per session, examiner, and time between sessions varied. The final task of a session was usually one which the patient normally did well, providing a control for fatigue, boredom, discouragement, or satiation. A poor score was reported only if a good performance was obtained on a different task later in the session, or if the poor performance recurred during an adequately controlled session.

Summary of procedures

In Fig. 2, items in boxes are observable, controllable, and measurable. Unboxed items are intervening events, to be considered later. Samples, at the left, provided initial input to the patient in any of three modalities: vision, hearing, or touch; he processed samples according to the demands of each task. In simultaneous matching (top line), pressing the sample window exposed a second set of input stimuli, the choices; the final response in the sequence was to press a choice window. Delayed matching (second line) had the same initial input and final output as simultaneous matching, but the delay intervened between sample press and appearance of choices; choices appeared without the sample. Naming and writing (lines 3 and 4) had
the same initial input as matching; the name spoken or written by the patient produced auditory or visual feedback.

Because of our patient’s right hemiparesis, he wrote, pressed windows, and felt all tactile stimuli with his left hand.

![Diagram](image)

**Fig. 2.** Summary of test procedures. See text.

**RESULTS AND ANALYTIC METHODS**

The patient’s clinical description is in Appendix 2. Test scores from a 31-week series will be examined first with responses constant and sample stimuli varying. The simplicity of the tasks (normal middle-class children perform nearly perfectly after the first grade), and probabilities based on the number of choice stimuli available on each matching-to-sample trial led to the following conservative scale for quick evaluation of the scores:

- **Less than 30 per cent correct**: Definitely deficient. In the range of chance performance.
- **30 to 70 per cent correct**: Definitely deficient, but better-than-chance performance.
- **75 to 100 per cent correct**: Satisfactory performance.
- **Delayed matching**: Failure to pass 40-sec is a deficit.

*Sample input varied; response output constant*

**Initial deficits and subsequent time course.** Different recovery courses will demonstrate the need for caution in comparing a patient’s performances on different materials. Even with stimuli held constant, temporal change may invalidate comparisons of tests given at different times. Therefore, we evaluated the test results in terms of relative rates of change.

Figure 3 is a prototype of those which follow. Each frame on the left side shows the patient’s scores for a given response: matching, naming, or writing single letters. Tasks are identified within each frame, modality of sample-letter presentation on each curve. Scores for simultaneous matching, naming, and writing are per cent correct; for delayed matching, the longest delay the patient passed correctly. No test was given every week; a modality label for a single unconnected point means the test was only given once.

Early deficits improved gradually but unevenly, as in simultaneous tactile-visual and auditory-visual matching, naming visual and tactile letters (visual-naming and tactile-naming), and delayed visual-visual and tactile-visual matching. Auditory-naming (repeating the auditory sample) and all writing, first tested in weeks 6 and 7, were nearly perfect.

On the right side of Fig. 3 are the patient’s scores for tests with trigrams. Initial deficits, similar to single letters, were more severe and longer lasting.
"Identity" and "nonidentity" tasks. Simultaneous visual-visual matching of letters is an "identity" task; sample letter and correct choice are exactly the same. The patient can match a letter without recognizing it as a letter, even without having seen it before. The same considerations apply to numbers, words, colors, and objects.

So long as the patient can discriminate and match any physical stimulus aspect, for example, shape, area, or angularity, he need have no behavior uniquely common to a particular sample and choice. He cannot, however, match a letter seen to a letter heard without having learned names or other responses that are uniquely common to a particular visual letter and its auditory counterpart; auditory-visual matching is "nonidentity."

Tactile-visual matching may be a second type of identity task. Having learned that tactile and visual stimuli, although physically different, may be equivalent, the patient can match a tactile to a visual letter without having experienced the letter before in either modality.

Certain naming and writing tasks constitute additional identity classes. For example, auditory-naming may involve only repetition. A patient who can imitate sounds will be able to name a dictated letter he has neither heard nor spoken before. He may also copy
a visual letter without having seen or written it before. Tactile-writing may also be an identity task; a patient who can draw tactile samples may copy a tactile letter he has never felt or written before.

By contrast, the patient could not do nonidentity tasks by imitation, copying, or form equivalence, but only by virtue of learned mediating responses. Nonidentity tasks in this study were: all auditory-visual matching, visual-naming, tactile-naming, and auditory-writing; visual matching of words with pictures, words with colors, words with numbers, and numbers with dots.

We believe that delayed matching, even of identity materials, is a nonidentity task. In delayed matching, sample and choice are never available for simultaneous comparison; the patient must respond to the sample with some behavior that permits him to bridge the delay. Two observations would be consistent with the interpretation of delayed matching as a nonidentity task: a breakdown in the patient's performance when an identity task is changed from simultaneous to delayed matching; different relative rates of improvement in simultaneous and delayed matching.

We shall now document the validity of the identity-nonidentity distinction by showing that:

A. Within each response category identity tasks improved more rapidly than, or along with nonidentity tasks, but never more slowly. All materials but one demonstrated this sequence.

B. The patient's performance often broke down upon a change from simultaneous to delayed matching, in identity as well as nonidentity tasks.

C. Delayed-matching tasks did not always improve in the same temporal sequence as the corresponding simultaneous identity-matching tasks.

In Fig. 3, identity tasks have larger data points and solid lines. Within each response category, identity performances improved most rapidly, although all single-letter writing tasks were approximately the same when first tested. In simultaneous matching, the trends of improvement were: first, visual-visual identity; second, tactile-visual identity; third, auditory-visual nonidentity. In naming, auditory identity naming (repeating) came first, then tactile and visual nonidentity. Differential writing trends appeared only in the trigram data; the order was like simultaneous matching.

Although the patient scored 75 per cent or better on simultaneous matching, delays caused all his good performances to break down, except auditory-visual matching of single letters. The order of improvement in delayed single-letter matching was the reverse of simultaneous matching. Delayed visual-visual trigram matching also lagged behind delayed tactile-visual matching; delayed auditory-visual matching was not tested, since simultaneous matching was never good.

The left side of Fig. 4 shows the trends when the patient matched, named, or wrote 3-letter words. There were two types of visual samples, 3-letter words and pictures, and two types of auditory samples, 3-letter words spelled to the patient (aud. sp.), or pronounced (aud. pr.).

Again, identity tasks improved first. In matching and writing, the patient improved most rapidly in response to visual and tactile word samples; next, to pronounced or spelled samples; and last, to pictures. The identity naming task, repeating pronounced words, was ahead of all nonidentity naming except pronouncing spelled words (tested relatively late). All matching broke down during the early delay tests. Delayed matching did not clearly preserve the identity-on-identity recovery sequence.
Performances in the right side of Fig. 4 involved color names in the responses. Colors or color names provided the input. The only identity tasks administered, auditory-naming and visual-writing of color-name samples, improved earlier than or along with nonidentity naming and writing. The matching tasks, both nonidentity, broke down when delay was introduced.

On the left side of Fig. 5 are trends for word-picture and picture-picture matching. Samples were the same pictures and visual, tactile, spelled, or pronounced 3-letter words as in Fig. 4. The only identity task, simultaneous picture-picture matching, was ahead of all the others, but did not break down in delayed matching.

Colors were choices in the matching tasks on the right side of Fig. 5. Samples were colors and visual or auditory color names. Simultaneous color-color identity matching, perfect in the first test, broke down when delay was introduced. Rates of improvement in delayed and simultaneous matching followed a similar sequence.

The patient's scores on saying and writing the names of pictures and colors, not shown in Fig. 5, were included in Fig. 4.

Figure 6 summarizes tests in which samples were visual or tactile digits, visual or auditory digit names, and visual or tactile dots. The dots varied in pattern but corresponded
Fig. 5. Test scores as a function of time; matching-to-sample with pictures and color choices.

Fig. 6. Test scores as a function of time; responses involving digit names, digits, and dots.
in quantity to the appropriate digit. In tests involving digit names as responses (left side of Fig. 6), the only identity task given was auditory-naming (repetition of dictated digit names). When tested, this was above or equal to nonidentity naming. Simultaneous matching tasks, all nonidentity, were performed well in their first tests and all broke down when delay was introduced.

In the center column, responses involved digits rather than digit names. All writing was performed well when first tested. Identity matching of visual digits, perfect in the first week, broke down in delay. Tactile-visual matching was considerably weaker than auditory-visual in the second week, the only instance in which an identity task fell behind a nonidentity task involving the same choice stimuli.

The third column shows trends for nonidentity matching of digit samples to the appropriate numbers of visual dots.

Samples in Figs. 7 and 8 were common objects, seen or palpated, or their spoken names. The only identity task in Fig. 7, auditory-naming (repeating), ranked ahead of nonidentity naming.

![Graph](image)

Fig. 7. Test scores as a function of time; responses involving object names.

In Fig. 8, visual and tactile object and spoken object-name samples were matched to pictures of the actual objects, projected in color and shadowed to give 3-dimensional effects. Simultaneous visual-visual and tactile-visual identity matching improved before auditory-visual nonidentity matching. Only the nonidentity task broke down in delay, and relative trends within delayed and simultaneous matching were similar.

*Additional findings consistent with the identity-nonidentity distinction.* Single-letter matching in Fig. 3 had lower-case samples and choices. Tactile-visual tests in which the patient had to match upper-case sample letters to lower-case choices returned to normal 4 to 6 weeks later than the corresponding performances of Fig. 3.

Although many upper-case letters have essentially the same shape as their lower-case counterparts, some pairs are quite dissimilar, the most obvious being: A-a, B-b, D-d, E-e, G-g, H-h, N-n, Q-q, and R-r. Table 1 combines all tests in which the patient made at least one error: when upper-case samples were members of dissimilar pairs the error rate
Fig. 8. Test scores as a function of time; matching-to-sample with objects (pictures) as choices.

(errors per opportunity) was 40 per cent, compared with 21 per cent when upper-case samples were members of similar pairs. Error rates showed no such difference between the two sets of letters when samples and choices both were lower case. Hence, within tactile-visual matching of single letters, identity tasks were performed more accurately than non-identity.

Table 1. Error rates in matching Upper-case and Lower-case sample letters with Lower-case choices

<table>
<thead>
<tr>
<th>Samples</th>
<th>Upper-case (%)</th>
<th>Lower-case (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dissimilar pairs</td>
<td>40</td>
<td>27</td>
</tr>
<tr>
<td>Similar pairs</td>
<td>21</td>
<td>24</td>
</tr>
</tbody>
</table>

Sample letters for writing were lower case, but the patient was reinforced for correct writing or printing in either upper- or lower-case styles. Lower-case printing predominated in his responses to visual or tactile 3-letter words. He preferred to print upper-case letters, however, when samples were auditory and their form unspecified. The difference supports the view that this patient’s letter-writing from visual or tactile samples was an identity, or copying, task. Writing trigrams, color names, and digit names were similarly consistent.

Posner and Mitchell [8], using upper- and lower-case letters as well as other stimuli, found that normal adults matched physically identical stimuli more rapidly than physically different stimuli which had the same class names. Their analysis suggested strongly that different levels of stimulus processing were involved in the identity and non-identity tasks.

Deficits classified as input, output, or relational

One can rule out input deficit as an explanation for a subnormal performance by showing that the patient can, under some circumstance, respond normally to the same stimuli. For example, even while our patient had trouble naming visual letters, he wrote and matched those same letters (Fig. 3). Therefore, his problems with visual letters could not be classified as input deficits. In the early weeks, however, no test with tactile letter samples
was normal; tactile input deficit for single letters could not be ruled out until later.

To exclude output deficit, one must show that the patient can execute the response satisfactorily under some circumstance. For example, even while our patient had trouble naming visual and tactile letters (week 6 and later) he was able to name auditory letters; his subnormal letter-naming was not an output deficit. Output deficit could not be ruled out, however, before intact repetition had been demonstrated.

If input and output are intact, a subnormal performance that involves the intact elements must, by exclusion, fall into the relational category. It indicates a deficient stimulus-response relation.

This patient’s early deficits were difficult to classify because he did not have all the necessary tests, but starting with weeks 6–7, a complete analysis was possible. Like single letters, visual trigram samples (Fig. 3, weeks 6 and 7) led to satisfactory simultaneous matching and writing; auditory samples were named adequately; tactile samples produced satisfactory simultaneous matching. Each modality was involved in at least one satisfactory performance, demonstrating that each trigram input was intact.

On the output side, identity matching, naming, and writing were satisfactory in weeks 6–7, demonstrating that each output was intact. Subnormal performances with trigrams were neither input nor output deficits.

Similar analysis of Figs. 4–8 reveals at least one satisfactory performance with each input, and at least one instance in which each form of response was good. With input and output deficits ruled out, all deficits after week 5 must be classified as relational.

An examination that fails to combine each output with each input may lead to mistaken identification of the deficit; in weeks 1–5, before identity naming had been tested, one might have concluded that the patient suffered a naming output deficit. Also, if the patient is tested early, before various input-output combinations have revealed different recovery courses, one may fail to observe input or output deficits resolving into relational deficits.

Categories of relational deficit

The relational category of deficit is not a unitary entity. Different nonidentity performance trends among stimulus materials, input modalities, and output responses illustrate its multifaceted nature.

Modality-specific relational deficits. Separation of the two types of stimulus-response relation, identity and nonidentity, even when they share a common output, is required before one can accurately delineate modality-specific deficits.

For example, if the sequence of improvement in writing or simultaneous matching were examined without separating identity from nonidentity tasks, it would appear that the patient’s responses improved first to visual, then to tactile, and finally to auditory samples. Would a classification of deficits according to input modality be preferable to the identity–nonidentity distinction? In support of the greater generality of the identity–nonidentity distinction, we offer the following:

First, in naming, the identity task improved most rapidly even though its input was always auditory. Thus, the identity–nonidentity distinction generalized from matching and writing to naming; input modalities did not.

Second, delayed matching to auditory samples sometimes improved before delayed matching to visual or tactile samples. The patient’s improvement in these nonidentity tasks did not consistently follow the same sequence of input modalities as simultaneous matching.
Third, identity matching and writing of words in response to visual-word samples were superior to nonidentity matching and writing of the same words in response to pictures (Fig. 4), even though these tasks shared the same input modality. Similar distinctions held between different visual inputs for color-name writing (for example, color vs. color-name samples, Fig. 4), and matching of color or picture choices (Fig. 5). Digit names, digits, and dots (Fig. 6) did not clearly differentiate various visual inputs.

When nonidentity tasks were considered separately a different type of orderliness emerged according to input modalities. Whenever differences existed within nonidentity writing, naming, or simultaneous matching, responses to auditory samples improved first. Trends distinguishing vision and touch for the same samples were not evident except for digits. The analysis of identity tasks was incomplete, for tactile-tactile and auditory-auditory matching were not tested. Delayed matching, inconsistent with respect to input modalities, will be examined separately.

Combined modality- and material-specific relational deficits. Although visual samples generally produced more severe nonidentity deficits than auditory samples, we have noted variations among tasks that differed only in the type of visual input. For example, oral picture naming was less impaired than saying the same names in response to visual-words, writing or matching the same names in response to pictures (Fig. 4), and matching picture choices to visual-words (Fig. 5). Critical to these differences was the involvement of visual words as initial sample input (visual word naming and word-picture matching), choice input (picture-word matching, or response-produced input (picture-writing). Only oral picture naming involved no visual words. Also, the patient's naming of pictures and spelling of pronounced word samples were the only instances in which nonidentity naming was superior to comparable matching tasks with visual word choices.

Visual-digit naming was also less impaired than saying the names in response to visual words or writing the names in response to visual digits (Fig. 6). Matching digit-name choices to visual-digit samples, however, showed no deficit.

Visual color names did not share the special status of digit names and picture names in producing relational deficits. The patient's naming of visual colors was impaired and was not superior to naming visual color words, or to writing and matching color names to visual color samples (Fig. 4). The combined modality-material specificity of visual-word deficits did not extend to color-name deficits.

Output-specific relational deficits

Up to now we have considered the effects of different input stimuli while keeping outputs constant. This type of analysis exposed the need to separate identity from nonidentity tasks, led to the classification of this patient's deficits as relational rather than input or output, and permitted the identification of modality- and material-specific relational deficits. Now, in order to identify output-specific relational deficits, we shall reverse the method of analysis and examine different responses while keeping the input constant.

Naming vs. writing. The left side of Fig. 9 shows the patient's nonidentity naming and writing scores in response to visual samples; the right side, to auditory and tactile samples. Naming was superior to writing in response to visual pictures, digits, and objects; spelled and pronounced words; tactile digits and objects. Only with color samples were naming and writing not easily distinguishable.

Our analysis differs from the customary evaluation of dyslexia and dysgraphia in two important ways. First, the most often reported comparison, writing-from-dictation (auditory-writing) with reading-from-text (visual-naming), violates the principle that different
responses are comparable only when the input remains constant. A consequence of failure to observe this restriction is illustrated by Fig. 10, in which a comparison of our patient's auditory-writing with visual-naming fails to reveal the difference between naming and writing.

A second feature is our avoidance of the term “reading” as a response classification. Reading is a term appropriate both to saying and writing the names of visual (or tactile) stimuli, but not auditory stimuli. The operational response classifications, (oral) naming and writing, encompass reading as expressed in speech and writing, yet permit separate examination of both responses to visual and auditory stimuli, whether or not the performance can be labelled “reading”.

Combined modality-material specificity and naming vs. writing. The role of visual words in this patient’s deficits helps explain the unusual superiority of naming over writing (Fig. 9). Naming and writing were compared only as responses to the same samples, and visual word samples were eliminated because they permitted identity writing, or copying. Because of these constraints visual words were never inputs in naming tasks that could legitimately be compared with writing. All writing, however, had visual words as response-produced input from the writing itself (Fig. 2).
As noted before, visual color names did not produce the same unique relational deficit as visual picture or digit names. Consistent with this, writing color names was the sole exception to the rule that writing was inferior to oral naming (Fig. 9).

This patient's relational writing deficit was at least in part secondary to the more general visual-word relational deficit.

Matching vs. naming and writing. Nonidentity matching of names to various input stimuli usually improved more rapidly than or along with saying or writing the names. However, this generalization was weakened by certain exceptions which reflected the visual-word relational deficits described above; the patient's data are not conclusive on this point.

Delayed matching-to-sample. Did the delayed-matching deficits fall into the same category as other relational deficits, or is a memory category required? Some evidence suggests that the delay deficits were a secondary consequence of relational deficits other than memory.

The change from simultaneous to delayed matching introduced two new factors. First, the sample was no longer available when the patient made his choice. If sample disappearance caused the delayed-matching deficit, length of the delay would be immaterial.

Delay duration was the second factor. If the time interval were critical, mere disappearance of the sample would cause the patient little difficulty, but longer delays would produce more errors.

The adjusting-delay procedure does not easily lend itself to the differentiation of sample disappearance and delay duration. The longest delay the patient can attain in 20 trials may be influenced by the extent to which absence of the sample, by itself, causes errors. The following considerations suggest that this patient's delayed-matching deficits were a function of sample disappearance per se.

First, even in the early weeks the patient did not show clinical signs of the global amnesia that would be expected if his severe delayed-matching deficits reflected a memory loss.
Second, more than 50 per cent of the patient’s responses had to be correct if he was to advance to the 40-sec delay. If sample disappearance reduced his accuracy to just above 50 per cent, he might fail to pass 40 sec within 20 trials but might succeed if he had more. This was, indeed, observed on a few occasions when the patient was given more than 20 trials, and indicates that errors caused by sample disappearance, rather than delay duration, limited his performance.

Third, a few tests with fixed delays gave similar error scores at short and long delays. This was to be expected if sample disappearance rather than delay duration were critical.

With sample disappearance implicated as a cause of the break downs in delayed matching, we may speculate on connections between delay deficits and relational deficits other than memory. The input must generate some process that “carries” the sample after it disappears. Naming is the usual bridge between sample and delayed choice. A patient who is unable to name samples, choices, or both is likely to perform poorly when first exposed to delayed matching. Although this patient’s initial breakdowns in delayed matching were accompanied by deficits in naming the samples or choices, there were two exceptions. Delayed visual-visual matching of digits was poor (Fig. 6, center) even though the patient could name visual digits; delayed matching of visual digit samples with digit-name choices was poor even after the patient had recovered his ability to name visual digits and digit names (Fig. 6, left). These exceptions, along with the failure to observe any instances in which good naming accompanied good initial delayed matching, cast some doubt on the relation between naming and delayed matching deficits. There is also the problem of the patient’s improvement in some delayed-matching tasks even while naming was still deficient. Although naming is the normal way to bridge delays, however, the patient need not apply names normally. He may invent names or name only certain stimulus aspects: long, short, large, small, bright, dull, straight, crooked. Then, sample disappearance will no longer block delayed matching. As in identity tasks, improvement in delayed matching may reflect attentional shifts rather than recovery from the relational deficit.

**DISCUSSION**

**Analytic methods**

In the analysis of behavioral deficit we have chosen the traditional stimulus materials and responses of the aphasia battery. Clearly they are complex, varying along more dimensions than we have delineated. Just as clearly, however, these very stimuli and responses have traditionally revealed the rich and varied phenomena of aphasia and agnosia. Rather than refine them, and risk refining their associated disorders out of existence, we have taken them as they are, adding only the laboratory refinements of procedural automation and consistency while observing the elementary principle of controlled stimulus and response variation.

Because different factors influence responses to the various stimulus materials, we have avoided direct comparisons of the patient’s scores even on similar tests done with different materials. For example, we have not attempted to delineate separate letter and number “centers” by comparing auditory-visual matching of single letters with single digits; there are 26 letters, but only 10 digits. We have not tried to compare word and trigram writing on the basis of “meaningfulness”; three-letter words usually have one of five vowels as the middle letter, but consonant trigrams have one of 21 consonants in the middle. We have not attempted to delineate separate agnosias by comparing scores for picture, color, and number naming; 3-letter picture names contain fewer letters than most color and number names.
Instead, we have compared the patient's performances across materials only with reference to the test profiles within each material. For example, naming pictures was superior to saying the same names in response to printed text, but naming colors was not superior to naming printed color words. These and other profile differences within color-naming and picture-naming tasks, not the absolute test scores, distinguished the two materials.

Comparisons of absolute test scores across materials are justified only when the more deficient performance is also the simpler. For example, words are longer than letters, contain several letters as elements, and come from a larger stimulus population. If a patient's auditory-visual matching of words is intact while he is deficient with letters [9], the difference is meaningful, particularly if other patients show the opposite.

Perhaps more important than revealing differences, comparison of test profiles permitted the demonstration of uniformities across materials when absolute scores could only have been confusing. For example, trigrams produced more profound deficits than single letters or 3-letter words, but the trend relations for corresponding tasks within each of these materials were similar. In spite of wide variations in absolute test scores across materials and over time, we were able to delineate large areas of orderliness.

Comparisons of input or output sufficiency across materials are also of limited value, and may even be misleading. Although a good performance with, for example, visual letter samples tells us that the patient is not blind, he may still have an input deficit for colors or even visual words. Similarly one cannot rule out a priori the possibility that a patient may be able to say words but not letter names or even color names. Although all these are speech, they are not the same output.

By evaluating different responses to the same stimulus, and the same response to different stimuli, one can identify a subnormal performance as an input, output, or relational deficit. Standard terminology does not easily adapt itself to these distinctions. A reasonably good aphasia examination might wrongly classify the present patient as "mixed expressive-receptive". His widespread deficits in naming and writing, however, were not "expressive," or "motor." All were relational deficits. His difficulties with certain types of visual words were also relational, not "receptive," "sensory," or "agnosic." Inputs were deficient only when related to certain types of output, and outputs only when related to certain inputs. Only by rigorously applying the methodological practice of testing all combinations of each available input and output can relational deficits be identified and analyzed.

The existence of orderly but different recovery courses for various stimulus-response relations, while not surprising, raises certain problems relevant to clinical examination methodology. Trend patterns in our patient's test scores from one period of time, considered alone, would lead to different conclusions than the same tests from another period. Also, trend patterns for identity and nonidentity tasks were different. Biased selection or incomplete testing would permit one to bolster a number of contradictory conclusions about the functional significance of lesions, depending on one's theoretical predilections. The problems are complicated by variability in the recovery course even under relatively well-controlled laboratory conditions, and by the likelihood that some apparent recoveries merely reflect the patient's new approach to a task.

**Deficit categories**

If response-reinforcement relations are deficient, the patient may be behaviorless; if stimulus control is generally deficient, he will be unable to respond appropriately to the test stimuli; if the sample, changing on each trial, does not "instruct" the patient, he will be unable to do the tests. Such deficits set an upper limit on the severity of the disorders to
which our methods are applicable. At the other extreme, patients with subtler aphasic disorders will do all our tests perfectly. Therefore, the methods we have presented are applicable to patients whose deficits fall into a middle range of severity. Within this range some patients may show only a few deficits or even only one, restricted perhaps to a single stimulus–response relation, while others may show widespread problems of several types. The present patient fits into the latter category. His pattern of deficiency and normal performance provides a basis for comparison with other classification systems and patients, and for testing certain theories of higher nervous function.

**Deficient nonidentity and intact identity performances.** Consistently intact identity performances automatically rule out input and output deficits. Head [10] has also noted the importance of intact identity performance as a control for the patient’s understanding of the task. In our patient, the identity–nonidentity distinction took precedence over stimulus modalities, materials, and responses; only after classifying the identity tasks separately was it possible to analyze the relational deficits into specific modalities, materials, and responses.

The identity classification recognizes that patients may do some matching, naming, and writing tasks as identities, but normal adult subjects are likely to do them as nonidentity language tasks, via naming. It is of interest to note the finding of Birch and Bornter [11] that first-grade and older children are more likely to match objects on the basis of class or functional attributes (nonidentity) than common stimulus properties (identity), while younger children do the opposite. A patient’s treatment of letters or words as identities suggests that the controlling stimuli are letters as shapes, not letters as language; merely sounds, not sounds of a word. Because he cannot process letters and words into language, he learns to attend to aspects of those stimuli which would not normally exert exclusive control over his responses. Our patient’s early improvements in identity tasks probably reflect his learning substitute behavior rather than his recovery from deficit. Since the stimulus aspects to which the patient reacts may differ from those the examiner specifies, comparisons of identity with nonidentity tasks in tests of language, of specific agnosias, or of inter- vs. intra-modality integration, are likely to be invalid. Teuber [12] has provided an excellent discussion of this and related matters.

Failure to attend to stimuli as language may seem an input deficit. Language, however, unlike shape or color, is not a physical stimulus property, nor is it speech or any other single output. Language is a relational process that includes input and output. A general deficiency in this relation will, if a patient can discriminate physical stimulus dimensions, cause an attentional shift, reflected in good identity and concurrently poor nonidentity performances. Thus, the faster improvements in identity than nonidentity tasks were probably secondary to, and actually demonstrative of, continuing relational deficits.

The identity-nonidentity distinction assumes additional importance when one recognizes that not all patients show the same pattern. Patients with more restricted language deficiencies may show impairment predominantly for particular stimuli and responses in nonidentity-tasks, for example color naming [13], matching spoken letter names to printed letters [9], or nonrepetitive oral naming [14]. Patients with right-hemisphere lesions may even be more deficient in certain identity than nonidentity tasks, for example in tactile-visual vs. auditory-visual matching of letters [15].

**Relational naming deficits.** Our patient was able to repeat words when he could not say them in response to appropriate visual and tactile stimuli. He was able to match and write visual and tactile word samples that he could not name. His naming deficits, therefore,
were not simply an amalgam of receptive and expressive elements; they were a distinctly separate category. Relational naming deficits are crucial elements in Brain's [2] central aphasia (but not Goldstein's [16] central aphasia), in Nielsen's [17] description of Wernicke's aphasia, and in amnestic aphasia. Our operational definition of relational deficits may help clarify such concepts as Brain's "central word and meaning schemas," and Nielsen's "language formulation area." In the absence of pathologic confirmation we are not in a position to specify the anatomical correlates of such processes, and we venture no more than the behavioral description of a particular kind of breakdown in stimulus-response relations. That such behavioral description, independently elaborated, does make contact with anatomical observations may encourage attempts to classify more precisely the psychological disturbances in aphasia.

Geschwind [18] and Teuber [12] have stressed the importance of distinguishing failures of naming from failures of recognition. Our practice of comparing naming, writing, and matching, in both the identity and nonidentity modes, as responses to the same input stimuli, permits us to identify naming deficits which exist in isolation and those which, as in our patient, form only one element of a larger deficit complex. Differential classification is also possible within strictly naming deficits. Patients may range from severe output deficits with no oral speech, through more or less severe relational deficits in which only identity naming is intact, to more isolated deficits such as inaccurate naming of visual colors. Our examination techniques permit the isolation of these and other more subtle deficit categories.

Modality and intermodality deficit specificity. Although visual, tactile, and auditory inputs were associated with some deficient performances, our patient's nonidentity tasks were generally more deficient in response to visual and tactile than to auditory inputs. This difference probably reflects the combined modality-material specificity discussed below. At present, our examination does not permit complete evaluation of intermodality deficit specificity. In our matching-to-sample tests, all choices are visual stimuli. Therefore, a poor tactile-visual matching score, for example, lacks the necessary controls that would be provided by tactile-tactile, tactile-auditory and visual-tactile matching tests.

Combined modality and material deficit specificity. Nonidentity tasks that involved pictures and spoken picture names exclusively as stimuli and responses were somewhat impaired, but corresponding tasks that had the same names in the visual and tactile modalities were even more impaired. Performance profiles in number and number-name tasks showed a similar pattern, but colors and color names did not. The defining characteristics of words that produce this type of deficit have yet to be identified.

This patient's inability to respond correctly to visual words led to poor scores in auditory-visual word-word matching, and in visual-visual word-picture and picture-word matching. These deficits, in conjunction with his impaired auditory-visual matching of words with pictures, raise a puzzling theoretical issue. A previous review of the literature [9], particularly of the patients with corpus callosum sections reported by Gazzaniga et al. [19-21], suggested that word-word and word-picture matching could be accomplished by the nondominant hemisphere. The present patient's deficits suggest that either: (a) some part of the dominant hemisphere is required for undisturbed word-picture matching, even in patients with corpus callosum sections, and that the present patient's left-hemisphere lesions interfered with this function; or (b) undetected lesions were present in this patient's right hemisphere also.

The problem is further complicated by this patient's difficulties in auditory-visual
matching of letter names with letters. Other patients tested in the same way have had no or lesser problems with picture names and pictures but were markedly impaired in auditory-visual matching of letters [9, 14]. Resolution of these problems must await pathological findings, but it is clear that the method of behavioral classification yields data relevant to differential neurologic classification.

REFERENCES


APPENDIX I. STIMULUS MATERIALS

<table>
<thead>
<tr>
<th>Materials</th>
<th>Sample</th>
<th>Choices</th>
</tr>
</thead>
<tbody>
<tr>
<td>b, c, d, f, h, j, k, l, m, n, p, q, r, s, t, v, w, x, y, z</td>
<td>Letter or Letter Name: b</td>
<td>b, j, x, z, l, m</td>
</tr>
<tr>
<td>pfx, kvh, npt, rjt, mbd, sax, rqs, thk, kfr, vlm, jtg, wsg, hwp, xlb, dbk, wpj, bff, ctd, ekh, ngs</td>
<td>Trigram or Letter Names: pfx</td>
<td>pfx, pfx, xpf, pfx, pxp, xfr</td>
</tr>
</tbody>
</table>
Stimulus Materials (continued)

<table>
<thead>
<tr>
<th>Word or Picture:</th>
<th>Words or Pictures:</th>
</tr>
</thead>
<tbody>
<tr>
<td>cat</td>
<td>cat, cow, hat, pie, dog, car, bug, bee</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Color or Color Name:</th>
<th>Colors or Color Names:</th>
</tr>
</thead>
<tbody>
<tr>
<td>orange</td>
<td>orange, blue, green, red, white, yellow</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Digit, Dots, or Digit Name:</th>
<th>Digits, Dots, or Digit Names:</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>5, 9, 7, 6, 8, 2, 3, 4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Object or Object Name:</th>
<th>Object Pictures:</th>
</tr>
</thead>
<tbody>
<tr>
<td>cork</td>
<td>cork, ball, nail, marble, hairpin, screw, key, caster</td>
</tr>
</tbody>
</table>

APPENDIX 2. CLINICAL NOTES

Initial evaluation

The patient, J. L. P., a 62-year-old right-handed sheet metal worker, and sole source of information concerning his illness, claimed good health until 5 days prior to admission, when he noted moderate weakness of the right upper extremity and slight weakness of the right side of the face. The weakness worsened over the next two days to involve the right leg as well. On admission, he admitted to "some trouble" with speech but was uncertain of its duration.

General Physical Examination showed a thin, unkempt, aphasic man with blood pressure 160/80, pulse 48 and regular. A grade 2/6 systolic ejection murmur was heard over the apex and radiated to the right carotid artery. Pulses in the carotid arteries were 2/4.

Neurologic Examination showed a dense right homonymous hemianopia. There was a moderate right hemiparesis, face and arm exceeding leg with increased tone and striking drift of the right arm, DTR's 3 plus in the right arm and leg. Vibration and pin appeared normally appreciated, but a penny placed in the right hand was misnamed. His speech was dysfluent, with frequent errors in word finding and rare paraphasias. Articulation was clear. He stated his name, approximate date, and place. At sight he named a watch, called a watchband "bracelet" and the stem "night hand." From dictated command he repeated short phrases, recited the number series and repeated 5 digits forward and 3 reverse; simple calculations were performed easily.

Laboratory Data: The white blood count was elevated to 10,500. Normal values were obtained for hematocrit, BUN, FBS, 2 hour post prandial sugar, Na, K, Cl, vanDenBerg, ESR. On the fifth hospital day a lumbar puncture showed clear, colorless fluid with protein of 81.

Hospital course

Second hospital day. A left common carotid arteriogram demonstrated complete occlusion of the left internal carotid artery, with faint filling of the carotid siphon, mainly from pharyngeal branches of the external carotid artery, slight filling through the ophthalmic artery, and minimal filling of branches of the middle cerebral artery in the 3-4 second film.

It was decided that no operative intervention would be undertaken, and the remainder of the patient's hospitalization consisted of evaluation and rehabilitation of his basic neurologic deficits.

Fourth hospital week. An Eisenson test for aphasia was carried out by the Language Clinic over several days time. At sight, he failed all tests of naming, but succeeded in matching 4/4 objects, 3/3 pictures, 5/5 simple forms, and 5/5 reduced size pictures. From dictated command he mimicked 3/3 sounds, repeated 10/10 digits, 6/6 words, 5/5 sentences, answered 10/10 questions requiring a single word answer, answered 4/5 questions requiring whole sentences, recited the digit series, the alphabet, weekdays, months, pointed to 5/6 body parts, obeyed 6/6 simple commands and 12/12 complex commands.

Sixth hospital week. He was tested at the bedside by two of us (J.P.M. and J.L.). At sight, reading aloud was tested with text from the Gray Oral Reading Test, the Wexler Intelligence Scale, the Stanford-
Binet, and the polylingual magazine *Quinto Lingo*. All these texts were read aloud as neologisms or with substituted short words. In the Stanford–Binet subtest, he accepted as correct sentences with scrambled word order. Although he refused to attempt to read aloud from simultaneously-translated text from *Quinto Lingo*, he separated the English text from that of other languages, and denied ever having seen before the characters in the Russian text. Although he misnamed with neologisms, he correctly pointed to the missing items and to the absurdities in the subtests of the Stanford–Binet. From *dictated command*, he mimicked 3/3 common sounds he failed to name. He defined 5/5 dictated words from the Wexler vocabulary subtest. He named 5/5 words spelled to him as letters (ks, sir, rise, spire, spiers).

*Seventeenth hospital week.* He was reexamined shortly prior to discharge. The dense right hemisynonymous hemianopia, dense right brachial weakness, slight right facial and leg weakness remained unchanged. The tongue was strong and articulation remained normal. He managed a feeble grip with the right hand and showed 7/10 grasp. The leg was lifted easily at the hip and toes dorsiflexed slightly against resistance. Joint position sense, vibration, and pin were appreciated equally in the right and left index finger. His conversation appeared relaxed and uninterrupted by dysfluency as on admission. At *sight* he performed no better with the same texts than he had at 6 weeks. He misnamed single letters presented in misoriented fashion, but succeeded in naming them after he re-oriented them himself. He said “*stop*” when shown a red card and asked what the color meant if on a street light. But he made errors pointing to colors representing grass, bananas, and sky. On *dictated command* he repeated long, unfamiliar words (i.e. nickelokshka), spelled hippopotamus, defined the Wexler vocabulary subset words repair, commerce, terminate, edifice, and impale, but not travesty or plagiarize. He stated “*30*” as the number of inches in 2 ft.

*After one year.* When last seen clinically, his examination showed little change. He has been living alone in a small hotel and finds his own way to and from the laboratory.

**Résumé**—A partir d'une analyse fonctionnelle des relations stimulus-réponse (le contrôle du comportement par le stimulus), on a dérivé des méthodes d'examen et d'analyse des déficits comportementaux dans l'aphasie. Les déficits étaient classés comme des effondrements des relations du contrôle stimulus-réponse; des profils de performances, déficitaires et intacts, chez un sujet atteint d'aphasie sévère illustrent l'intérêt des tests de contrôle. La distinction entre épreuves d'identité et épreuves de non-identité ressort clairement des données recueillies. Cette distinction passe à travers différentes réponses, les modalités et les matériaux de stimulus. En examinant séparément ces deux types d'épreuve, on obtient une classification plus précise des déficits du malade selon les catégories d'entrée, de sortie ou de relation, chacune de celle-ci à son tour pouvant être classée selon un type spécifique. Les données recueillies démontrent que des examens répétés sur une période prolongée de temps peuvent repérer des changements de classification du déficit, modifications qui n'auraient pas été notées en utilisant seulement les tests habituels.