Language and Thinking in Psychosis

Is There an Input Abnormality?

William M. Grove, PhD, Nancy C. Andreasen, MD, PhD

- We studied "formal thought disorder" in schizophrenics, schizoaffecitives, and manic by examining syntax processing and perception of meaning, using the "embedded click" and "memory for gist tasks," two paradigms that were developed by psycholinguists. To control for generalized performance deficits, a matched-task design was used. Contrary to expectation, patients did worse on a matched memory for digits task than on sentence processing. At a six-month follow-up examination, schizophrenics' performance did not improve while other patients' did. We concluded that psychotic patients have no specific language perception deficit but do have a short-term memory deficit. This deficit tends to remit for manic and schizoaffecitives, but not for schizophrenics.

(Arch Gen Psychiatry 1985;42:25-32)

The nature of psychotic "formal thought disorder" has been an elusive puzzle for many years. Attempting to understand it is analogous to attempting to defeat Russia in a land war. Everyone who attempts it seems to fail, although some have penetrated the territory deeply and maintained a prolonged occupation and siege.1-6

The present study is part of another set of forays into the topic. These studies derive from the belief that the subject of formal thought disorder will be best understood if approached systematically and clinically and the studies examine the functional brain system through which we most often observe disordered cognition: the language system. Aphasiologists have carefully studied and mapped this system, and their insights provide models that assist in understanding thought disorder. As the study of aphasia has made clear, one must separate input from output, or comprehension of language from production of language. As the study of psycholinguistics suggests, language functions include such various components as syntax (grammatical organization), semantics (units of meaning), and discourse (the organization of sentences to form topics). Current investigations also attempt to answer a series of important clinical questions about formal thought disorder. What is its relationship to diagnosis? What is its longitudinal course? What causes it?

The present report focuses on one aspect of this system-atic mapping of language function in psychosis: language input (perception and processing). It asks whether psychotic patients are competent in their understanding of spoken language. It examines two separate aspects of language: syntax and semantics. If patients' understanding is abnormal, is it because they do not decode the grammar of what we say or because they do not understand the meaning? Based on clinical experience, we have predicted that the language comprehension of psychotic patients is intact and that their primary abnormality is in language production. This pair of studies tested the hypothesis of intact comprehension, while subsequent ones will explore aspects of language production.

In these experiments we have attempted to correct a series of problems that have plagued investigations of formal thought disorder in the past. One is a failure to study a broad range of psychotic patients to see whether the disorder is specific to schizophrenia.7,8,25,7,8 Another is failure to study disorders over time so that their prognostic significance and underlying mechanisms can be better understood.8 A final difficulty is a failure to search for a specific deficit that is central to the disorders observed, rather than one due to such epiphenomena as poor motivation or distractibility.9 One would like to find a deficit that can be understood in terms of psycholinguists' mapping of normal language functions and aphasiologists' concepts of abnormal language.7,8

To these ends we have chosen to study the behavior of schizophrenics, schizoaffecitives, and manic on two language-processing tasks. These tasks have come from research by psycholinguists on normal persons and are tied to a fairly strong theoretical and empirical analysis of speech processing. A primary goal is to map functional abnormalities of language perception in diverse patient groups and to compare them to normal subjects and to each other. Furthermore, we have investigated the longitudinal course of language behavior in these psychoses to see whether the cross-sectional view of language obtained by studying acutely psychotic patients may be refined and supplemented by follow-up data. We have also used a special design to pick out specific rather than general cognitive deficits.

STUDY DESIGN
Click and Gist Paradigms

Two paradigms from psycholinguistics were used in this investigation, the "embedded click" paradigm and the "memory for gist" paradigm. The click paradigm is considered to measure the degree to which subjects process syntactically and analyze sentences into three major clausal constituents. The gist paradigm

Accepted for publication Dec 29, 1983.
From the Department of Psychology, University of Minnesota, Minneapolis (Dr Grove); and the Department of Psychiatry, University of Iowa, Iowa City (Dr Andreasen).
Reprint requests to Department of Psychiatry, University of Iowa, 500 Newton Rd, Iowa City, IA 52242 (Dr Andreasen).
stresses processing of semantic information in sentences. The distinction between syntax and semantics is somewhat arbitrary but indispensable.

The click paradigm rests on the following observation: listeners tend to displace the subjective location of a brief noise heard during a spoken sentence from its true location toward the major constituent boundary of the sentence. In the click paradigm, as originally reported by Fodor and Bever, the subject hears a sentence in one ear and a click in the other ear; after hearing a sentence such as “... good click get something nice for me.” The click is placed near or in a major constituent boundary during the recording. Subjects tend to report it as lying nearer to (or in) the boundary if it was really a few syllables away. This phenomenon is not due to acoustic clues such as intonations, or to the fact that clause boundaries are generally bordered by pairs of words low in probability of serial occurrence.

If psychotic patients fail to displace clicks in the usual manner, a specific language deficit in syntactic processing might be isolated. Two studies have examined the click displacement of schizophrenic listeners. Rochester, Harris, and Seeman found that schizophrenics were inaccurate in placing clicks, but they were as likely to displace clicks into major constituent boundaries as were normal subjects. Carpenter reported that the percentage of clicks objectively located in the constituent boundary that were accurately reported was as high for schizophrenics as for nonschizophrenics. However, a difference between schizophrenic subtypes was found. Compared with schizophrenics with histories of good premorbid functioning, good premorbid patients tended to report clicks earlier in the sentence than they actually were. This points to heterogeneity in schizophrenia, and suggests that subtyping schizophrenics may be important in psycholinguistic research as it has been elsewhere.

To assess perception of meaning, we used a psycholinguistic task called the memory for gist paradigm. The gist effect depends on the often-observed superiority of our ability to remember the information content, or gist, of what we have heard over our ability to remember the exact words we have heard. The task materials consisted of four sets of sentences, all intermixed. Each set of sentences concerns a different set of four related ideas. A typical idea set, expressed in one sentence, is “The rock which rolled down the mountain crushed the tiny hut at the edge of the woods.” The four ideas embedded in this sentence are: the rock rolled down the mountain; the rock crushed the hut; the hut was at the edge of the woods; and the hut was tiny.

The subject listens to one-, two-, and three-idea sentences, but never hears any four-idea sentences. The subject is encouraged to attend to meaning. After another interfering task that prevents rehearsal of the sentences, the subject is unexpectedly given a recognition test. The subject must indicate whether or not he thinks he has heard exactly a given sentence before, and how confident he is of his judgment.

Note that the linear gradient of confidence ratings, from a high degree of belief that they have heard the four-idea sentence (which in fact they have not heard), through steadily lesser degrees of belief that they have heard the three-, then two-idea sentences, to a negative confidence rating on the new one-idea sentences. On a few one-idea sentences, which in fact the subject had heard before, the subjects tested by Bransford and Franks indicated no greater confidence than in new, unheard one-idea sentences. Studies have confirmed that the confidence-in-recognition gradient is indeed semantic in nature and is not an artifact of the instructions given.

Knight and Sims-Knight examined schizophrenic and control task performance. Their subjects were hospitalized male veterans who scored 4 or higher on the New Haven Schizophrenia Index. Schizophrenics were divided into those with a history of good premorbid functioning and those with histories of poor premorbid functioning. Nonpsychotic patients and college students were also tested. Among schizophrenics, only patients with good premorbid functioning who were not known to have a long history of hallucinations were not always able to use context in memory as well as normals.

Our studies have tested hypotheses similar to those of Carpenter and Knight and Sims-Knight. We predicted that the major deficit would be in the area of production rather than comprehension. Syntactic processing should be intact, as reported previously. However, we predicted that no group would show a deficit in semantic processing on the gist task, unlike Knight and Sims-Knight. We used the Schedule for Affective Disorders and Schizophrenia (SADS) and the Research Diagnostic Criteria (RDC) for choosing study subjects. Recent studies show that RDC diagnoses have substantially lower reliability and error rates than patients diagnosed by the New Haven Schizophrenia Index. We wished to examine the most clinically homogeneous group with the severest degree of illness possible. It is these patients that one may expect will show the greatest cognitive defects.

**Matched-Task Design**

Because we expected that a general performance impairment among psychotic patients might decrease performance on the click and gist tasks, we endeavored to control for this by using a matched-task or “deficit” design. Chapman and Chapman have argued that a deficit is not presumptively crucial to an understanding of cognitive pathology unless the deficit “stands out” from a background of impaired performance. The matched-task design controls for inattentive and poor motivation by matching experimental tasks rather than patients and control subjects. A differential deficit appears when subjects show lower performance on one task than on another. To avoid artifact, tasks must be matched on several parameters describing psychometric discriminations: discriminability, distribution of item difficulty parameters and intertask consistency. Although the matched-task design is hard to execute, it promises to enlighten researchers more than merely finding a deficit would. It has, to our knowledge, never been applied to the perception of sentences by psychotic patients. Rattan and Chapman and Kagan and Olmman have applied this strategy to the study of single-word comprehension and speech. However, people do seem to use sentence structure as well as isolated words.

The choice of a control task for comparison with the psycholinguistic tasks is difficult. One wants a control task that shares all of the features of the experimental task, except for that feature that measures the performance component that may be deficient. Obviously this can only partly be done. Many skills are required for even the simplest of cognitive performances. We used a digit-span control task for several reasons. First, its instructions and procedures are simple, so that one does not produce impaired performance just by using an overcomplicated task. Second, it is unaffected by phonelizacion administration. Third, it uses the auditory input of commonly heard materials, like the experimental digit and gist tasks. Fourth, digit span requires brief remembering of a sequenced input, like the experimental tasks. Fifth, and very importantly, we have the data to show that random digit strings lack syntax and meaning. We reasoned, then, that the differences between a linguistic task and a digit-span task would be a good measure of differential deficit in language comprehension.

Based on previous click-paradigm studies, we predicted that psychotic patients would all use syntax normally so that there would be no deficit in click-task performance in comparison with a matched task. We further predicted that if on follow-up, manic patients would show the greatest improvement on language task performance and the schizophrenics would show the least, to the extent that any performance deficit was observed at the index evaluation. We also predicted that if any differences were observed between paranoid and hebephrenic schizophrenics, they would favor paranoid patients.

We predicted that there would be no differential deficit on the gist task. We also made a contingent prediction about gist deficits: if manics were impaired at index evaluation, they would be normal at follow-up. If schizophrenics were initially impaired, they would still show some impairment at follow-up. We generalized from clinical findings that schizoaffective patients would be heterogeneous but would most closely resemble manics in performance.

Both studies reported here used identical subjects, assessments, and task-matching procedures. The two studies are described together except for specifics of psycholinguistic tasks.
Table 1.—Subject Characteristics

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, yr; $\bar{x}$ ± SD</td>
<td>28.92 ± 9.13</td>
<td>35.04 ± 14.01</td>
<td>28.16 ± 9.50</td>
<td>26.68 ± 7.41</td>
<td>27.75 ± 6.21</td>
</tr>
<tr>
<td>Education, yr; $\bar{x}$ ± SD</td>
<td>14.10 ± 3.42</td>
<td>13.38 ± 2.24</td>
<td>13.50 ± 1.89</td>
<td>12.56 ± 1.66</td>
<td>13.74 ± 2.77</td>
</tr>
<tr>
<td>Shipley IQ score, $\bar{x}$ ± SD</td>
<td>115.48 ± 7.71</td>
<td>102.88 ± 11.98</td>
<td>108.00 ± 10.09</td>
<td>95.26 ± 12.88</td>
<td>104.00 ± 9.80</td>
</tr>
<tr>
<td>No. (%) female</td>
<td>26 (52)</td>
<td>13 (52)</td>
<td>14 (56)</td>
<td>12 (48)</td>
<td>8 (32)</td>
</tr>
<tr>
<td>No. (%) acute</td>
<td>6 (26)</td>
<td>1 (4)</td>
<td>1 (4)</td>
<td>(\ldots)</td>
<td>(\ldots)</td>
</tr>
<tr>
<td>Duration of episode, wk; $\bar{x}$ ± SD</td>
<td>9.88 ± 9.63</td>
<td>64.48 ± 164.44</td>
<td>246.96 ± 239.55</td>
<td>256.58 ± 262.24</td>
<td>(\ldots)</td>
</tr>
<tr>
<td>Total time in hospital, wk; $\bar{x}$ ± SD</td>
<td>28.46 ± 30.84</td>
<td>41.72 ± 76.27</td>
<td>66.71 ± 160.86</td>
<td>79.71 ± 139.76</td>
<td>(\ldots)</td>
</tr>
<tr>
<td>No. (%) taking lithium carbonate</td>
<td>21 (84)</td>
<td>15 (60)</td>
<td>2 (8)</td>
<td>1 (4)</td>
<td>(\ldots)</td>
</tr>
<tr>
<td>No. (%) taking major tranquilizer</td>
<td>17 (68)</td>
<td>23 (92)</td>
<td>19 (76)</td>
<td>20 (80)</td>
<td>(\ldots)</td>
</tr>
</tbody>
</table>

Table 2.—Standardized Scores for Click and Control Tasks in Manics, Schizoaffective, and Schizophrenics

<table>
<thead>
<tr>
<th>Group, No. Patients</th>
<th>Digit Span</th>
<th>Control, $\bar{x}$ ± SD</th>
<th>Difference, $\bar{x}$ ± SD</th>
<th>t</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manics, 25</td>
<td>-0.31 ± 1.14</td>
<td>-0.37 ± 0.70</td>
<td>0.06 ± 1.59</td>
<td>0.16</td>
<td>&gt;0.10</td>
</tr>
<tr>
<td>Schizoaffectives, 25</td>
<td>0.32 ± 1.84</td>
<td>-0.54 ± 0.86*</td>
<td>0.97 ± 2.07</td>
<td>2.34</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Hebephrenics, 25</td>
<td>-0.18 ± 1.57</td>
<td>-0.70 ± 0.61*</td>
<td>0.51 ± 1.66</td>
<td>1.63</td>
<td>&gt;0.05</td>
</tr>
<tr>
<td>Paranoids, 25</td>
<td>0.19 ± 1.73</td>
<td>-0.54 ± 0.61*</td>
<td>0.66 ± 1.26</td>
<td>3.41</td>
<td>&lt;0.01</td>
</tr>
</tbody>
</table>

*The difference from normal subjects was $P<0.05$.†The difference from normal subjects was $P<0.01$.

SUBJECTS AND METHODS

Subjects

The subjects were 50 normal persons used in constructing psychologically matched tasks, and 100 patients from the University of Iowa Psychiatric Hospital, Iowa City, and Mount Pleasant (IA) Mental Health Institute. The 100 patients were selected from consecutively admitted patients by chart review and an abbreviated interview based on the SADS. All but four patients were collected at the University of Iowa. If the subject was 18 years of age or older, spoke English as his native language, was white, and appeared to meet the RDC for hebephrenic or paranoid schizophrenia, schizoaffective disorder, manic or depressed subtype, or mania at a definite level, he was asked to participate. After obtaining informed consent, diagnoses were confirmed by administering the full SADS interview. Approximately 75% of the subjects who were asked agreed to participate, and just one failed to complete the experimental tasks at intake evaluation. Diagnoses were stratified to include 25 each of hebephrenic schizophrenics, paranoid schizophrenics, schizoaffectives (irrespective of subtype), and manics. In no case was the initial brief structured interview diagnosis in such error that a subject had to be excluded, but in a few cases the administration of full SADS resulted in reassigning a subject from one inclusion diagnosis to another.

The 50 normal subjects were collected by advertisements in community newspapers. No systematic matching was employed, but an attempt was made to attract nonstudents and older subjects to achieve rough comparability to the local patient population.

Assessments

In addition to the SADS interview and the RDC, each schizophrenic patient was measured on the Phillips Fremdorph History Rating Scale. All patients completed the Shipley Institute of Living Scale. All treatments administered just prior to and during hospitalization were also recorded.

Follow-up Evaluations.—Approximately six months after the index evaluation, subjects were sought out and reevaluated with the SADS. Any medications the subject had been recently taking were recorded, and the click, memory for gist, and digit-span tasks were repeated. Different click and gist materials were given, to avoid any possibility of memory for the earlier sentences. Procedures were otherwise identical to those of the intake evaluation.

Matching the Experimental and Digit-Span Tasks.—As discussed previously, items on the digit-span task were selected to match psychometric characteristics between digit-span and click tasks. This was done using data from the normal subjects. The parameters on which tasks were matched were coefficient $a$ reliability, mean and variance of item difficulty, and shape of distribution of item difficulty. This gave a set of six digit-span items to match click items, and eight to match gist items. A good match on psychometric parameters was obtained ($\pm 2\%$ on X and SD of item difficulty and coefficient $a$ reliability). All scores were standardized to a mean of 0 and an SD of 1 in normal subjects. This matched-task design ensures that even if psychiatric patients show a generalized performance deficit, the estimation of differential deficit will not be compromised. This holds because equal discriminating power of the two tasks will cause equal shifts in scores on the matched tasks in psychotic patients, if there is only a generalized deficit.

Click-Paradigm Materials.—The materials consisted of 20 nine-word, two-clause sentences recorded onto one track of 1/4-in. stereo cassette recorder. The embedded click, recorded on the other track of the tape, was a 25 ms capacitor-discharge click as loud as the loudest speech sound in the corresponding sentence. The click might be placed in the boundary, in the word before the boundary, or in the word after the boundary. Four sentences for each click location relative to the clause boundary were prepared, and eight other sentences were recorded as well. Four of these came first on the tape and were practice sentences, and four other sentences disguised the regular construction of the test sentences. Thus each subject heard 20 sentences. The materials were identical to those used by Carpenter.

In recorded instructions, the subject was told that sometimes a copy of the sentence would be given to the subject immediately after hearing it, so the subject would have only to mark the click location, but that sometimes he would have to write the sentence from memory, and then mark the click location. This was to encourage attention. Four sentences of the 20 were in fact written
Table 3.—Task Scores at Follow-up

<table>
<thead>
<tr>
<th>Task</th>
<th>Manic, 18</th>
<th>Schizoaffective, 19</th>
<th>Hebephrenic, 17</th>
<th>Paranoid, 15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Click, x ± SD</td>
<td>7.67 ± 6.31</td>
<td>10.36 ± 6.23</td>
<td>7.80 ± 5.91</td>
<td>9.92 ± 6.19</td>
</tr>
<tr>
<td>Digit-span control, x ± SD</td>
<td>1.94 ± 1.82</td>
<td>2.06 ± 1.47</td>
<td>1.00 ± 1.06</td>
<td>1.53 ± 1.30</td>
</tr>
</tbody>
</table>

*Comparison between groups: F(3,65) = 0.91, P > .1.
†Comparison between groups: F(3,65) = 1.89, P > .1.

Table 4.—Gist and Matched Control Task Scores in Manics, Schizoaffectives, and Hebephrenics

<table>
<thead>
<tr>
<th>Group, No. Patients</th>
<th>Gist, x ± SD</th>
<th>Control, x ± SD</th>
<th>Difference, x ± SD</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manic, 25</td>
<td>0.84 ± 1.18</td>
<td>-0.62 ± 0.85</td>
<td>1.46 ± 1.47</td>
<td>4.97†</td>
</tr>
<tr>
<td>Schizoaffective, 25</td>
<td>0.41 ± 1.23</td>
<td>-0.57 ± 1.05</td>
<td>0.99 ± 2.07</td>
<td>2.34§</td>
</tr>
<tr>
<td>Hebephrenic, 25</td>
<td>0.20 ± 1.28</td>
<td>-1.07 ± 0.73</td>
<td>1.26 ± 1.63</td>
<td>3.79†</td>
</tr>
<tr>
<td>Paranoid, 25</td>
<td>0.08 ± 1.13</td>
<td>-0.80 ± 0.87</td>
<td>0.73 ± 1.89</td>
<td>1.89</td>
</tr>
</tbody>
</table>

*The difference from normal subjects was P < .01 (higher than normal).
†The difference from normal subjects was P < .01.
§The difference from normal subjects was P < .001.
††The difference from normal subjects was P = .05.

RESULTS

Subject characteristics are shown in Table 1. It is clear that the normal subjects used in standardizing tasks were a little brighter and better educated than the patient groups, especially the hebephrenics (Table 1). This difference is probably not large enough to induce large group differences on the study tasks. Moreover, if samples are to represent their respective populations, then the normal subjects should be a bit higher on intelligence and educational attainment than the psychotics, with the possible exception of manics. Table 1 also reveals that the sex distribution of the patient sample was fairly evenly split. The conditions of the schizoaffectives were almost entirely subchronic or chronic by the RDC diagnosis, with very few remitting schizoaffectives. Overall, the hebephrenics had a relatively early onset of illness, a high rate of premorbid asociality, chronic courses, a low likelihood of asymptomatic status between exacerbations, but relatively little time spent in hospital. The schizoaffectives lay between the manics and schizoaffectives on age, the duration of episode at index evaluation, and the total time ever spent in the hospital. Thirteen hebephrenics were manic in the index episode, seven were depressed, and five were cycled. Sixteen were of the manic subtype when tested, and the balance were depressed. The sampling of schizophrenic patients seemed representative of typical schizoaffectives.

Performance on Click and Control Tasks at Index Evaluation

Table 2 gives the results for several tasks for the four patient groups. Since normal subjects have by definition a mean of zero and an SD of unity for both tasks, their scores are not shown in Table 2. Table 2 also shows the results of significance tests for the difference between normal subjects and each patient group. No group differed significantly from normal subjects on the click displacement score, yet all but manics differed from normal subjects on the digit-span test. This indicates that psychotic patients make effective use of syntax in processing sentences, as such use is measured by the click paradigm.

This suggests a differential deficit in digit-span performance. Only manic patients showed no evidence of this, and this is because they did poorly on the click task. Inspection of the manics' click test answer sheets disclosed that they were sometimes careless in marking click locations; they made more "wild" errors. Thus their failure to show the same pattern of results for the click-digit-span task comparison may simply have been due to inattentiveness. In any case, no psychotic group showed any tendency toward a differential deficit in syntax comprehension.

One can examine the pattern of differences between psychotic groups on task scores. Such an analysis tests whether the groups differ in their patterns of performance. The analysis (not tabulated) indicated that the manics' pattern did not differ significantly from that of the other patients. No other comparisons approached significance.

Another aspect of this study involved follow-up evaluation. One might expect improvements in digit-span performance in the follow-up interval for manics but little improvement for schizoaffectives. For hebephrenics, a further decrease in performance could be expected.

Follow-up efforts succeeded in the case of 18 manics, 19 schizoaffectives, 17 hebephrenics, and 15 paranoids; this diagnostic distribution did not differ from the original one. Those patients followed up did not differ from those not located on the initial
Global Assessment Scale (GAS) score, a 100-point scale that combines severity of symptoms and impairment in role functioning.

All patient groups had improved significantly (P < .02 or better) on the GAS score at follow up, but the groups differed in GAS-score improvement from index to follow-up evaluation. The scores of manics and schizoaffective patients improved significantly more than schizophrenic patients, with hebephrenics showing almost no improvement. The groups differed on their GAS scores at follow up (P < .001), with the GAS scores of manics and schizoaffective patients almost identical, and those of the hebephrenics markedly worse, with paranoiacs less ill than hebephrenics.

Since the click task sentences administered at follow-up differed from those administered at index evaluation, the follow-up tasks were not psychometrically matched. It is not meaningful to compute differential deficit scores between click and digit-span tasks at follow-up. Therefore, results for follow-up evaluations are given in raw score terms.

Table 3 gives scores for the click and digit-span tasks at follow up (Table 3). There was no significant difference between groups. Click-task scores were quite similar to those found at index evaluation, even though the actual sentences heard on the two occasions were not identical. Digit-span scores for the schizophrenic groups were still significantly lower (P < .005) at follow up than for the normal subjects used for standardization. As expected, this was most pronounced for the hebephrenics, although the difference between paranoids and hebephrenics did not reach significance.

Although an exact differential deficit score was precluded, it appears that at follow up a differential deficit persisted for schizophrenics, with poorer performance on the digit-span task. On the other hand, the digit-span deficit shown by schizoaffective patients appears to have remitted with their episodes.

We also correlated the follow-up GAS score with follow-up click and digit-span scores. The correlation between the GAS and click scores was .36 (P < .003), and between the GAS and digit span it was .31 (P < .02) among all patients, indicating that clinical status and task performance did covary modestly. Less ill patients did better on these tasks at the follow-up test.

**Table 6.** Recognition Ratings for Old vs New Sentences

<table>
<thead>
<tr>
<th>Group, No. Patients</th>
<th>Sentence Type</th>
<th>Old, x ± SD</th>
<th>New, x ± SD</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal, 50</td>
<td></td>
<td>2.76 ± 2.28</td>
<td>0.51 ± 1.34</td>
<td>8.51†</td>
</tr>
<tr>
<td>Manic, 25</td>
<td></td>
<td>3.05 ± 1.84</td>
<td>1.64 ± 1.59</td>
<td>5.30†</td>
</tr>
<tr>
<td>Schizoaffective, 25</td>
<td></td>
<td>2.62 ± 2.99</td>
<td>1.06 ± 1.66</td>
<td>3.51†</td>
</tr>
<tr>
<td>Hebephrenic, 25</td>
<td></td>
<td>1.81 ± 2.59</td>
<td>0.78 ± 1.72</td>
<td>3.03†</td>
</tr>
<tr>
<td>Paranoid, 25</td>
<td></td>
<td>1.73 ± 2.67</td>
<td>0.56 ± 1.50</td>
<td>3.29†</td>
</tr>
</tbody>
</table>

*P < .001.
†P < .01.

**Table 5.** Gist Recognition Ratings by Number of Ideas in Sentences

<table>
<thead>
<tr>
<th>Group, No. Patients</th>
<th>1, x ± SD</th>
<th>2, x ± SD</th>
<th>3, x ± SD</th>
<th>F*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal, 50</td>
<td>0.88 ± 2.10</td>
<td>-1.58 ± 2.21</td>
<td>54.56</td>
<td></td>
</tr>
<tr>
<td>Manic, 25</td>
<td>2.55 ± 2.20</td>
<td>-0.21 ± 2.46</td>
<td>20.94</td>
<td></td>
</tr>
<tr>
<td>Schizoaffective, 25</td>
<td>1.73 ± 2.46</td>
<td>-0.48 ± 2.69</td>
<td>17.93</td>
<td></td>
</tr>
<tr>
<td>Hebephrenic, 25</td>
<td>1.36 ± 2.37</td>
<td>-0.91 ± 2.66</td>
<td>30.13</td>
<td></td>
</tr>
<tr>
<td>Paranoid, 25</td>
<td>1.49 ± 2.22</td>
<td>-1.66 ± 2.22</td>
<td>32.72</td>
<td></td>
</tr>
</tbody>
</table>

*P < .001.

Clearly, gist-task scores were not lower than those of normal subjects; the manics' scores were even significantly higher. By inspection of the distribution of manics' confidence ratings, it appears that this was due to a small but consistent tendency to overuse the 5 level of confidence rating (but not to stereotyped marking of 5 for nearly every item). All patient groups did more poorly on the digit-span task than did normal subjects. As expected, schizoaffective and hebephrenics, especially hebephrenics, did worst on the digit-span task.

These two relationships conjoin to make a significant differential deficit on the digit-span score for each group. As with the click task, we did not observe a differential deficit in semantic processing as measured by the gist task. Instead, we observed intact language processing, both absolutely and relative to digit-span performance. As with the click task, semantic processing of sentences was better preserved in these psychotic patients than was memory for digit strings. This "reverse" differential deficit was unexpected.

In addition to examining differential deficit hypotheses, we analyzed patients' confidence-in-recognition ratings, comparing responses to four-, three-, two-, and one-idea sentences and to old vs new sentences. Table 5 gives the gist-recognition rating raw scores by the number of ideas per new sentence for each group tested. Note the very orderly and essentially linear decline in confidence ratings for all groups from the four- to the one-idea sentences. This trend was identical in rank order for each group. Moreover, our subjects almost all gave negative mean ratings for the one-idea sentences. Thus, all our patient groups performed very much like subjects in the experiment by Bransford and Franks, and in some ways unlike those of Knight and Sims-Knight.

The analysis of variance for these repeated-measures data (not tabulated) showed that there is no significant difference between groups in the relation of recognition ratings to the number of ideas per sentence.

Similar comparative results hold for the old sentences (those truly heard before) vs the new sentences. Table 6 gives recognition ratings for each type of sentence. In each group subjects rated the old sentences higher than the new ones. The group by the type of sentence interaction (old v new) was not significant, indicating that patients did not differ from normal subjects in how they rated old v new sentences.

The finding by Knight and Sims-Knight that schizoaffective, with poor premorbid histories integrated the semantic information in their gist sentences more poorly than did schizophrenics with good premorbid functioning, would lead one to expect that our premorbid history scale scores would correlate with gist-task performance. However, the scores did not. Abbreviated Phillips Scale scores at index evaluation correlated just .21 with the gist score, which is not significantly different from 0. We also examined the profile of four v three v two v one-idea sentence confidence ratings as a function of the Phillips Premorbid Scale score. At intake there was no tendency for patients with poor premorbid histories to show a profile difference from good premorbid, as measured by an F test of the Phillips score by the number of ideas interaction (F (3, 192) = .80; P > .1). Thus, we were unable to confirm the major finding of deficit on semantic integration reported by Knight and Sims-Knight.

Taken together, the gist-task results suggest that severely psychotic patients process semantic information from heard sentences in a normal manner. This would lead one to expect that their
Table 7.—Gist Task Scores at Follow-up

<table>
<thead>
<tr>
<th>Task Score</th>
<th>Manic, 18</th>
<th>Schizoaffective, 19</th>
<th>Hebephrenic, 17</th>
<th>Paranoid, 15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gist, X = SD</td>
<td>14.17 ± 42.87</td>
<td>10.73 ± 30.31</td>
<td>26.94 ± 35.60</td>
<td>20.47 ± 41.49</td>
</tr>
<tr>
<td>Digit-span control, X = SD</td>
<td>5.11 ± 5.97</td>
<td>5.00 ± 2.56</td>
<td>3.71 ± 2.80</td>
<td>4.33 ± 2.29</td>
</tr>
</tbody>
</table>

Performance on the gist task would be normal at follow-up evaluation as well. The performance deficit seen on digit span at index evaluation might remit for some patients. Since mania is typically a remitting illness, one would expect manic to recover on digit span at follow-up, while schizoaffectives might show a residual deficit. With schizoaffectives, a mean performance prediction may have less meaning (as well as certainty), since this group may be longitudinally as well as phenomenologically heterogeneous.

**Performance at Follow-up**

Since the gist sentences at the follow-up examination differed from the initial ones, we did not standardize gist and digit-span scores for comparison. Table 7 shows raw scores for the four groups. The analysis of variance shows that patient group differences on gist score were not significant (F[3, 65] = 0.64, P < .06). Nor were differences between patient groups on the digit-span score significant (F[3, 65] = 1.04; P < .38). The difference between normal subjects and patients was not significant for the gist task but was significant for the digit-span task (F[1, 114] = 7.08; P < .01), indicating that a dichotomous mixed group of patients was still not completely normal in digit-span performance at the follow-up examination. This difference was entirely due to the schizoaffectives, since manic and schizoaffectives did not differ significantly on digit-span scores from normal subjects at a follow-up examination. Schizoaffectives most closely resembled manics at the follow-up examination.

At the follow-up examination the patients' psychopathology was unrelated to gist performance. The GAS score correlation with the gist score was just - .09 (P > .1). As with the click-control task, GAS score positively correlated with the digit-span score. Although the correlation was not significant (r = .13; P > .1), it went in the expected direction; less severe illness is associated with better digit-span scores. This points to a weak, psychopathology-induced deficit in the short-term memory span. Gist performance showed no disturbance.

We also examined Phillips Premorbid Scale scores as predictors of gist task performance at the follow-up examination. We reasoned that even though the premorbid history was unrelated to gist performance at the index evaluation, those patients with poor premorbid histories might emerge as relatively poor processors of semantic information at the follow-up examination compared with patients with good premorbid histories. However, no such relationship was found using either the correlation between the gist and Phillips scores or a test for the Phillips score by the number of ideas interaction.

**COMMENT**

It seems clear that during no phase of their disorders do manics, schizoaffectives, hebephrenics, or paranoid schizophrenics evidence impaired syntactic or semantic processing of heard sentences. Their performance on the index and at follow-up examination was normal. On the other hand, all of the patients showed poor digit-span scores. This “reverse” differential deficit requires explanation.

We speculate that the experimental and control tasks differ in the amount of attention and type of output required to perform them, but that these differences did not matter for normal subjects. Patients were able to analyze the click sentences syntactically, and to remember information in gist sentences. The language tasks may require less active input organization. In attempting to recall digits, the patients' inability to organize came out. The digit-span task is highly taxing because it is disorganized. Only one encoding strategy, chunking, has been found to increase digit-span performance, and even this idea has been challenged. In contrast, the encoding and comprehension of speech is easily done by most normal subjects. This may be because it is highly practiced or because we have dedicated brain areas for it (or, as is likely, because of both). Probably under ordinary circumstances speech is easier to organize and recognize because attention is automatically directed to it. In that case, the difference in the attention, encoding, and recollection that speech (v or unorganized strings of items) requires would account for the differential deficit. The requirement that subjects recognize linguistic, but recall digit, materials may also be important, as numerous studies point to a schizophrenic superiority in recognition over recall.

Our results imply that no distinctively linguistic impairment is to be found in the way psychotic patients perceive and process sentences. Rather, a more general cognitive deficit is involved. This remits in affective disorder, but not in schizophrenia. A short-term or very short-term memory defect is implicated. This may be mediated by a primary, very general attentional deficit, in which the automatic disposition of concentration is blocked. This hypothesis, advanced by Holzman et al, is consistent with our data.

These results may be compared with studies of “contextual constraint” performed with schizophrenics. The procedure uses texts that vary from random word strings to connected sentences; the texts differ in the constraint that the context imposes. Subjects are asked to recall words from these strings. Normal subjects recall words much better if they are presented in connected strings than if they are presented in random strings. This is in part due to syntax processing, since even semantically meaningless strings are better remembered if they are grammatically correct.

The performance of chronic schizophrenics improves less than that of control subjects with the increasing orderliness of texts. However, the performance of all patients does improve. Depressives show even poorer response to contextual constraint than do schizophrenics. One study did not find a schizophrenic deficit in response to constraint. The one consistent finding is that the recall of schizophrenics improves with the increasing order of approximation to natural language. This is probably partly due to syntactic and semantic processing.

These findings suggest that the organization of input for the patient facilitates processing, compared with unorganized input. This conclusion is supported by work using other memory paradigms, including some studies using sentence materials. If this effect is relatively general, then the unorganized nature of input on the digit-span task, requiring a lot of active processing (eg, deliberate efforts to chunk digit strings), should present particular difficulties for patients. It is interesting that the group apparently most susceptible to this effect, however, is not schizophrenics but manics. This may be because both groups share similar defects in ability to organize input, or because of a
distractibility in manics or a more insidious, malignant degradation in schizophrenics' cognition. The follow-up data, showing that manics recover function, suggests that the deficits in the acute phase differ in origin.

Our gist results agree with previous research showing that various task features make high performance by psychotic subjects, especially schizophrenics, more likely. Auditory presentation of sentences produces more semantic organization among normal subjects than does visual presentation. The present task probably requires relatively little active processing; this improves schizophrenics' performance. The present task is a sort of recognition task at output, even though the effect depends on "misrecognition" of sentences; generally, schizophrenics find recognition easier than recall. In particular, the difference between visual presentation (used by Knight and Sims-Knight) and auditory presentation (used by us) may explain our failure to replicate their findings of deficit.

The results of the present study, along with those of earlier investigators, make it clear that manics, schizophrenics, and schizoaffective subjects have difficulty analyzing sentences syntactically into constituents. They can also recall the gist of what is said to them normally. What trouble they do have can be accounted for by a generalized performance deficit in the acute phase of illness. However, manics, hebephrenics, and paranoid schizophrenics show a differential deficit on matched digit span-control tasks, indicating that short-term memory for unrelated items is probably more impaired than sentence comprehension. On follow-up examination, only schizophrenics showed a continuing deficit in the performance of digit-span tasks, indicating that their short-term memory deficit has its roots in a general cognitive deterioration. This, we speculate, may be due to a defect in the automatic disposition of attention.

Apparently, natural speech comprehension under conditions of auditory presentation is not the source of most psychotic thought, language, and communication disorders. We think that the deficits accounting for most psychotic thought disorders are probably to be sought on the output side, rather than on the input side. This conclusion is supported by work with other paradigms showing that schizophrenic subjects have difficulty censoring deviant response tendencies. Subsequent articles will report on clinical and psycholinguistic studies of the speech output of the patients studied in the present investigations.

Richard Hurtig, PhD, provided assistance and suggested the use of the gist paradigm.

References