This dissertation has been microfilmed exactly as received

LLOYD, Dee Norman. REALITY TESTING IN THE THOUGHT PROCESSES OF SCHIZOPHRENICS.

The Ohio State University, Ph.D., 1961
Psychology, clinical

University Microfilms, Inc., Ann Arbor, Michigan
REALITY TESTING IN THE THOUGHT PROCESSES
OF SCHIZOPHRENICS

DISSERTATION
Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of the Ohio State University

By
Dee Norman Lloyd, B. A., M. A.

The Ohio State University
1961

Approved by
Harold B. Pepinsky
Adviser
Department of Psychology
ACKNOWLEDGMENTS

The author is especially indebted for the contributions to this study made by Mr. Thomas O. Karst and Mr. Stanley Pavey, who aided with the reliability scoring of the judgmental measures; and to Mrs. Helen Lloyd, who assisted in the preparation of the manuscript.

The study was carried out under the capable chairmanship of Dr. Harold B. Pepinsky, for whose counsel and assistance throughout the study the author is deeply appreciative.
# CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION AND BACKGROUND</td>
<td>1</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Background</td>
<td>2</td>
</tr>
<tr>
<td>II. STATEMENT OF THE PROBLEM</td>
<td>15</td>
</tr>
<tr>
<td>III. METHOD AND PROCEDURE</td>
<td>20</td>
</tr>
<tr>
<td>Subjects</td>
<td>20</td>
</tr>
<tr>
<td>Pretesting Procedures</td>
<td>22</td>
</tr>
<tr>
<td>Matching</td>
<td>29</td>
</tr>
<tr>
<td>Experimental Task</td>
<td>31</td>
</tr>
<tr>
<td>Scoring of the Experimental Story</td>
<td>34</td>
</tr>
<tr>
<td>Analysis of the Data and Experimental Hypotheses</td>
<td>38</td>
</tr>
<tr>
<td>IV. RESULTS</td>
<td>41</td>
</tr>
<tr>
<td>Evaluation of Matching</td>
<td>41</td>
</tr>
<tr>
<td>The Central Hypotheses</td>
<td>46</td>
</tr>
<tr>
<td>Secondary Analyses</td>
<td>54</td>
</tr>
<tr>
<td>V. DISCUSSION</td>
<td>62</td>
</tr>
<tr>
<td>Discussion of the Results</td>
<td>62</td>
</tr>
<tr>
<td>Theoretical Implications</td>
<td>65</td>
</tr>
<tr>
<td>Critique of the Experiment</td>
<td>68</td>
</tr>
<tr>
<td>Implications for Further Research</td>
<td>69</td>
</tr>
<tr>
<td>VI. SUMMARY</td>
<td>71</td>
</tr>
</tbody>
</table>
## Bibliography

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIBLIOGRAPHY</td>
<td>74</td>
</tr>
<tr>
<td>APPENDIX A</td>
<td>79</td>
</tr>
<tr>
<td>APPENDIX B</td>
<td>83</td>
</tr>
<tr>
<td>AUTOBIOGRAPHY</td>
<td>88</td>
</tr>
</tbody>
</table>
TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Means, Standard Deviations, and t Tests for Mean Differences for the Experimental and Control Groups on the Matched Variables</td>
<td>42</td>
</tr>
<tr>
<td>II. Mean Differences and t Tests between the Normal and Patient Groups</td>
<td>43</td>
</tr>
<tr>
<td>III. Correlations of the Matched Pretest Variables with the Experimental Criterion for Rate of Learning</td>
<td>45</td>
</tr>
<tr>
<td>IV. F Test for Difference between the Normal Experimental and Control Groups in Trials to Reach the 24th Criterion with Covariance Adjustment for Differences on the First Trial</td>
<td>49</td>
</tr>
</tbody>
</table>
FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>TITLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Design of the Experiment</td>
<td>32</td>
</tr>
<tr>
<td>2.</td>
<td>Average Learning Curves for the Normal Experimental and Control Groups</td>
<td>47</td>
</tr>
<tr>
<td>3.</td>
<td>Average Learning Curves for the Patient Experimental and Control Groups</td>
<td>51</td>
</tr>
<tr>
<td>4.</td>
<td>Average Learning Curves for the Experimental and Control Group Patients Learning beyond the 22nd Criterion of the Experiment</td>
<td>53</td>
</tr>
<tr>
<td>5.</td>
<td>Mean Usage of Synonyms over Trials for the Combined Patient and Combined Normal Groups</td>
<td>58</td>
</tr>
<tr>
<td>6.</td>
<td>Mean Number of Added Components over Trials for the Combined Patient and Combined Normal Groups</td>
<td>59</td>
</tr>
<tr>
<td>7.</td>
<td>Mean Number of Correct Out-of-Order Components for the Combined Patient and Combined Normal Groups</td>
<td>60</td>
</tr>
<tr>
<td>8.</td>
<td>Mean Number of Out-of-Order Components Changing the Meaning for the Combined Patient and Combined Normal Groups</td>
<td>60</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION AND BACKGROUND

Cushioned on dunes
The skull smiles at the parching sky.
Laughingly the wind stuffs sand in its mouth.

Sartora

I. INTRODUCTION

The fragmented thought, the bizarre and incongruous juxtaposition of ideas, the illogical and often incomprehensible use of words, and the mixed expression of reality and fantasy that characterize schizophrenic thought and language disturbance have been of interest to students of psychopathology for at least as far back as the clear and insightful description of them by Bleuler in 1911. Investigators have turned their attention to the understanding of the content of schizophrenic speech, to designing tests to determine the presence of schizophrenic thinking, and to formulating theories to account for this unusual phenomenon. Thought and language disturbance, as described above, is a major diagnostic indicator of schizophrenia, and the understanding of it is vital to the understanding of schizophrenic behavior in general. Although many theories have been presented to explain these aberrant thought processes, most remain too incomplete or oversimplified for the researcher of today. In recent years, there has been more attention given to the formal aspects of thought disturbance, and more empirical
research has been attempted. There are still almost as many theoretical positions to take, however, as there are experimental studies.

The following investigation attempts to test experimentally certain hypotheses about the formal aspects of schizophrenic thinking processes proposed by Schilder. Taking Schilder's theory as a basis, the author has attempted (1) to define clearly the constructs concerning thought processes of schizophrenics, (2) to determine the behavioral consequences that follow from these constructs, (3) to design an experimental model that is congruent with those aspects of the thought processes under investigation, and (4) to control as many as possible of the irrelevant variables that might affect the experimental results, through sampling and matching techniques. Finally, the findings of the experiment are compared with the results of related investigations.

II. BACKGROUND

Important theoretical positions concerning schizophrenic thought have been reviewed below. This review is selective and presents only portions of theories that are relevant to the present study. An attempt has been made to show similarities and differences between these theories, and to present some of the questions that have important implications for future research in this area.

Throughout the present study, the terms "thought disorganization," "thought disorder," and "scattering" are used
interchangeably to refer both to the manifestations in the language and test behavior of schizophrenics and to the aberrant thinking inferred from such manifestations. Although investigators differ in how they account for schizophrenic thought, there is considerable agreement on the group of schizophrenics to which these terms refer. Kasanin (1944) states that these cases are

...so clear-cut and so sharply differentiated [from schizophrenics grouped under other clinical symptoms] that I take the liberty of calling them cases of "primary thought disorder."

One of the earliest investigations into the thought processes of schizophrenics centered on disturbance of associations and is exemplified by the work of Jung and Bleuler.

Bleuler (1950) felt that disturbance of the associative processes was one of the basic symptoms of schizophrenia, along with disturbance of affectivity and ambivalence in affect, intention, and cognition. He considered the last two, however, as resulting from the basic disturbance of association processes. It was Bleuler's belief that the association disturbance was in turn the result of organic changes in the brain. He described the thinking processes of schizophrenics as follows:

Of the thousands of associative threads which guide our thinking this disease seems to interrupt, quite haphazardly, sometimes such single threads, sometimes a whole group, and sometimes even large segments of them. In this way, thinking becomes illogical and often bizarre. Furthermore, the associations tend to proceed along new lines, of which so far the following are known to us: two ideas, fortuitously encountered, are combined into one thought, the logical form being determined by incidental circumstances. Clang-associations receive unusual significance, as do indirect associations. Two or more
ideas are condensed into a single one. The tendency
to stereotype produces the inclination to cling to
one idea to which the patient then returns again and
again. (1950, p. 14)

Although Bleuler's contribution remains monumental, his inclusive
construct to account for thought disturbance tends to over­
simplify many of the relationships pointed out by subsequent
investigators.

Jung (1944), drawing on the research of Stransky and
Pelletier, attempted to analyze the laws of association that lead
to thought disorganization. From the results of Stransky's
experiment (interruption of the continuous speech of normal
subjects with distracting stimuli), Jung concludes:

"Whenever there is a disturbance between the points
of contact of one's emotional life and ideation,
producing thereby in normal thought the lack of
orientation by a controlling idea, there must result
a stream of thought resembling the flight of ideas
that is found in schizophrenia." (1944, p. 17)

Jung inferred the laws of association that a stream of thought
follows to be those of similarity, co-existence, motor speech
connections, and combinations according to sounds, perseverations,
and repetitions. He further proposed that some associations come
about through a mediated predicate to the stimulus word. Thus,
for example, the response, "dead cousin" to the stimulus word,
"wood" is interpreted: wood...wooden coffin...dead cousin
(the example is from Bleuler).

Another approach to the schizophrenic thinking processes
has been through the study of concept formation. One of the
earliest investigators in this area was Vigotsky (1934). From
his studies with the Ach blocks (best known in this country as the Hanfmann-Kasanin Test), Vigotsky described two developmental levels of thinking. The first, characteristic of young children, was termed "the complex." It was defined as, "a whole constituted of organically united parts," but also as a union that is "concrete and mechanical." The second and higher level, "the concept," was defined as, "a general and abstract principle on the basis of which the conceptual association is formed." The term "pseudoconcept" was used by Vigotsky to denote the situation when the complex coincides with a concept. That is, although one object may be referred to by an adult and by a child, the child may classify or think about the object in terms of concrete relationships rather than as part of the concept by which the adult classifies it.

Vigotsky concluded that the schizophrenic resembles the child in the use of complexes and pseudoconcepts, and has regressed from the normal adult level of mature conceptual thinking to a level intermediate between that of the adult and that of the child.

The later use of the concept formation test by Hanfmann and Kasanin (1942) and others, with the investigations in this area by Goldstein, has made concept formation one of the most widespread methods for investigating schizophrenic thought. Vigotsky's theoretical position that schizophrenia is a regressive disorder, however, has not been accepted by all the investigators using this test. The principle of regression is incorporated into several of the theories in this review and deserves separate consideration.
Two forms of regression can be distinguished, although both are usually accepted by theorists who accept one. The first can be described as ontogenetic, and refers to the return to the use of thinking processes that occurred earlier in the development of the individual. The second, which is usually designated as phylogenetic, refers to regression of thought processes to forms of thinking prevalent in "primitive" cultures and in genetically lower forms of life. Few, if any, regression theorists have contended that schizophrenic thinking or behavior is identical with that of children or primitives. As stated by Arieti (1955), regressed behavior is essentially different from the behavior of an organism which is integrated at a lower level. In regression to more "archaic mental mechanisms," there is not an integration at a lower level, but a disequilibrium with the functions of different levels of adjustment interfering with each other.

Examples of theories in which regression is central to describing schizophrenic thought are those of Arieti and Werner. Arieti (1950, 1955) described the regressed thought of schizophrenics by a principle of "paleologic reasoning," proposed by Von Domarus (1944). This principle of logic states that the identity of objects is established on the basis of predicates rather than on subjects, as it is in Aristotelian logic. Thus, "Christ was born on Christmas; I was born on Christmas; therefore, I am Christ" is an acceptable statement to the paleologist. The Von Domarus principle was used by Arieti as a formal statement of what others have described as widening of the concept basis, metonymic distortion (see below), and distant
associations. One recent attempt to test the Von Domarus principle experimentally (Gottesman and Chapman, 1960) failed to find preference in schizophrenic subjects for paleologic solutions to syllogisms.

The comparative developmental theory of Werner (1948) represents a general theory in which schizophrenic thought is regarded as regressive. In general, Werner stated that thinking develops from the general to the specialized, from the diffuse to the articulated, and from conceptual undifferentiation to hierarchic differentiation. These dichotomous terms have been applied to both the phylogenetic and ontogenetic development of thinking and perception. Thus, the regressed thinking of schizophrenics in these terms is characterized by a lack of centrality and subordination of concepts, more attention directed to the perceptual aspects of a situation, and syncretic combination of images and ideas in place of one central governing idea. In support of his theory, Werner has amassed many examples of similarities in the thinking of children, schizophrenics, and primitives, and his theory has been the stimulus for considerable research on schizophrenic thought.

Goldstein (1944), whose major work has been the investigation of language and thinking in aphasic patients, noted similarities in the findings of such men as Vigotsky with his own and applied his own theory to the problem of thinking in schizophrenia. As with cortically damaged patients, Goldstein found that much of the behavior of schizophrenics can be described as concrete or
lacking the "abstract attitude." Several investigators have
taken the concrete-abstract dichotomy as a basis for describing
schizophrenic behavior, and some, such as Hanfmann and Kasanin
(1942), Rapaport (1945), and McGaughran and Moran (1956), have
gone beyond Goldstein's use of the single construct, loss of the
abstract attitude, and have attempted to distinguish different
levels of concrete and abstract behavior.

While Rapaport (1951) agreed that much of schizophrenic
verbal behavior reflects the concreteness Goldstein described,
he felt that Goldstein's theory neglected the aberrant behavior
in abstracting that has been noticed by Benjamin (1944) and
Bleuler (1950). A recent study by Chapman (1960), comparing
errors of literal and figurative misinterpretation, suggested
that schizophrenics make both types of errors, but a greater
number of literal errors.

Several recent studies have focused on whether schizophrenic
thought can be described best in terms of concrete-abstract
behavior or in terms of disturbance of the interpersonal relation-
ships of schizophrenics. The work of Cameron and Sullivan, the
major exponents of this latter view, are considered before
reviewing these studies.

Cameron and Sullivan both emphasized the importance of
social or interpersonal development of thought and language
processes. Cameron (1947, 1951) stated that although language
functions both for interpersonal communication and for private
thinking, "thinking must be made to correspond with operations
in the shared social field of interbehavior if it is to remain socially valid." The schizophrenic, however, becomes isolated from the common social environment (social disarticulation). He no longer shares the attitudes of those around him, and he is no longer able to assume the roles of others. Without returning to the social community to verify his thinking, the schizophrenic, in his thinking, becomes increasingly "fragmentary, discontinuous, individualistic, and approximate" (Cameron, 1951).

As Sullivan (1944) stated, any expression in language that communicates to others relies upon "consensual validation." That is, in order to insure communication, an individual audits his thinking so that it follows accepted conventions for expression. In the inadequately socialized schizophrenic, the processes of consensual validation fail, and verbal expressions do not communicate to others.

Cameron (1938, 1944a) described schizophrenic language and thought as characterized by (1) metonymic distortion, which is the substitution of approximate, or related terms, in place of precisely definitive terms, (2) asyndetic thinking, which is the juxtaposition of terms and ideas without adequate linkage or integration, and (3) interpenetration, which is the intrusion of preoccupations and personal material into the reasoning and communication of the schizophrenic.

Several recent studies have given support to the thesis of Cameron and Sullivan that the thought disturbance of schizophrenics is related to their withdrawal from social relationships.
and lack of communication skills. McGaughran and Moran (1956) found, on the same test, that scoring for consensual conceptual areas showed significant differences between schizophrenics and nonschizophrenics; whereas scoring for levels of concrete-abstract concepts did not. Whiteman (1954) concluded that schizophrenics, when equated with normals on performance in determining formal (impersonal) concepts, were lower than normal subjects when determining social concepts. Flavell (1956) found a correlation between social behavior ratings and ability to select word meanings that reflect essential abstract characteristics of given words. Affleck (1954) found that on tasks previously judged for degree of interpersonal relationships involved, schizophrenics were poorer than normals in time to solution and number of errors.

The psychoanalytic position in regard to schizophrenic thought centers on the concepts of the "primary process" and the "secondary process" of Freud's theory.

Freud (1951) described the secondary process as conscious goal directed thinking, operating to satisfy needs in the context of external reality. With the development of the secondary processes, primary processes in the normal adult are subordinated and have expression only in dreams, or other atypical situations. Freud held that with any relaxing of the control of the reality oriented secondary processes, as in sleep in normals and in withdrawal from reality in schizophrenics, the primary process mechanisms come into play, operating to secure the
shortest path to wish fulfillment. In schizophrenic thought and language, characteristics of the primary process intrude and interfere with the operation of the secondary process. Dominant fantasies, hallucinatory images, condensations, approximations, and pars pro toto substitutes do not remain subordinated and occur along with logical, reality oriented thought.

One of the strongest aspects of psychoanalytic theory is that it enables both normal and abnormal thought to be explained within the same conceptual framework, and it also provides for a description of both the development of thinking processes and their dynamic operation at any one time.

The last theoretical approach to be reviewed is that which describes thought in terms of an assumed sequence of stages or phases that occur in the development of a single thought. Flavell and Draguns (1957) have given the name "microgenesis of thought" to this approach. In the theory of Schilder (1951), the thought processes in schizophrenia are described within this framework.

According to Schilder, three major phases occur in the development of all thoughts. In the first, or "preparatory phase" of thought development, mental contents consist of a series of diffuse images that arise from the life experience of the individual. Although these image-like formations have been initiated by some stimulus and increase in clarity to a final thought, this goal is vague and diffuse in this initial phase. Through the process of association these mental contents multiply
and are subject to symbolic change or combination according to affective similarity. Many of the associations, or "presentations," that occur are relevant to the thought, others are irrelevant.

In the "transitory phase," as the thought process continues, both a differentiation and integration of these presentations occur, and the irrelevant associations are suppressed in favor of the reality relevant elements.

In the last phase of the thought process, there is a unification of the reality elements into the final, completed thought, and the affective images have been subordinated to the background. According to Schilder, thinking gains its reality orientation only in the final phase of thought development. At this final phase, in the normal individual, there is an awareness of the truth or falsity of the completed thought.

Flavell (1955), comparing and integrating the formulations of Schilder and Mead (1934), expanded the statements of Schilder to emphasize the relationship of social reality to this phase of thought development. The reality oriented thought is a socially oriented thought. This also provides a link between Schilder's theory and Sullivan's description of "consensual validation" (see above).

Flavell further proposed a fourth phase of thought development, the importance of which has been alluded to by Whorf (1956) and Sullivan (1944), but which has been largely neglected by students of psychopathology. Flavell termed this final phase
the "externalization" phase. Here, the reality orientation of thought is further refined and translated into the language of the external world. Flavell proposed, as many feel from introspection, that inner thought, even reality oriented thought, does not have the same form as the externally communicated expression of it in language.

The disorganized thought that is found in schizophrenia, Schilder attributed to the closing of the thought processes prematurely, so that products of earlier phases of thought development are expressed in place of completed thoughts.

Summary. Although the number of experimental studies in this area has increased in the last several years, not enough is known about schizophrenic thought to relate the many aspects of the phenomenon that have been described. Most recent theorists agree on the important effect of social development on schizophrenic thought disorder, however, and all seem to agree on the observable characteristics that thought disorganization manifests.

Several theorists have stressed a single concept to account for thought disorganization. Jung and Bleuler felt the chief disturbance was in the associative processes. Vigotsky and Goldstein felt that loss of conceptualizing or abstracting ability was the most important aspect. In the theories of Werner, Von Domarus, and Arieti, the regression principle was the primary explanatory construct, and their theories emphasized the similarities between schizophrenic thought and the thought
characteristics of children and primitives. The social and interpersonal context in which thinking occurs was of prime concern to Sullivan and Cameron. Development of the intrapsychic processes and their dynamic functioning have been emphasized by the psychoanalytic writers. Finally, Schilder's description of schizophrenic thought in terms of the developmental phases that he assumed to occur in the production of a thought was presented.

Very briefly, these are some of the major differences in the theories presented. When one regards these theories in their complete form, however, many points of agreement can be seen. For example, Freud had a great influence on many of the theorists mentioned, especially Jung, Bleuler, Schilder, and Sullivan; and many of his concepts are incorporated into their theories of schizophrenic thought. If, however, all the major statements of the theorists reviewed are relevant to the understanding of schizophrenic thought, it becomes evident that the complexities of this phenomenon are greater than any one theory has proposed. Much more experimental evidence concerning the relationships between the various aspects of thinking, especially between those aspects that result in deviant thinking, is needed before an integration of these theories can be made.
In Chapter I, the description of thought as a series of developmental phases was introduced. This view of the thought process, as described by Schilder (1951), is analyzed further in this chapter to ascertain the measurable behavioral referents of some of these theoretical statements. At the end of the chapter, hypotheses to be tested in this study are stated.

According to Schilder, in normal thought development reality relevant associations are integrated into a completed thought in the final phase of development. At this point, in normals, there is an awareness of the truth or falsity of the completed thought.

The thought of schizophrenics is characterized by the premature closure of the thought processes, with a lack of what Schilder termed "correction processes" to subordinate irrelevant associations and integrate the reality oriented thought. At the final phase of development, premature unification of the thought results in a lack of awareness of the truth or falsity of the thought.

Recalling a short story was one of the methods Schilder found useful in studying thought processes, and many of the statements he has proposed have been in connection with the
results of his studies in recall. Describing the behavior of subjects in recalling stories, he stated that the patients accept whatever comes to mind as true. "Truth awareness is lacking in intensity and does not initiate correction" (1951, p. 561).

Schilder's use of the term correction processes and their relationship to awareness of the truth or falsity of the completed thought is somewhat ambiguous. He has implied that correction processes employed to bring about reality oriented thought lead to truth awareness, and that truth awareness in turn initiates corrections.

Schilder's concept of correction processes and their relationship to awareness can be restated parsimoniously as (1) correction_1 which occurs with the association process and brings about the integration of reality oriented thought, and (2) correction_2 which occurs in the anticipation set (apperception or registration process). In pathological thought, lack of correction_1 results in lack of awareness of the truth or falsity of a thought because reality integration does not occur. Lack of awareness, in turn, results in lack of correction_2. In the reproductions of Schilder's patients, there was distorted recall of the story from the lack of correction_1 and slow improvement after several repetitions because of the additional lack of correction_2.

In comparing Schilder's concepts of correction processes and awareness of the truth and falsity of a thought with concepts of other writers, it is seen that he was describing what others
have referred to in general terms as reality testing and reality feeling, respectively. Correction\textsubscript{1} and correction\textsubscript{2} as reality testing in the thinking process depend upon utilization of memory and past experience to modify anticipations of productions in view of external evidence. This can be compared with the
definition of reality testing given by Freud (1952):

It is now no longer a question of whether something perceived (a thing) shall be taken into the ego or not, but of whether something which is present in the ego as an image can also be re-discovered in perception (that is, in reality). Once more, it will be seen, the question is one of external and internal. What is not real, what is merely imagined or subjective, is only internal; while, on the other hand, what is real is also present externally. When this stage is reached, the pleasure principle is no longer taken into account.

Fenichel (Blum, 1953) emphasized the anticipative aspect of reality testing in his definition:

> Reality testing is the ability to anticipate the future in the imagination by trying actively in small doses what might happen passively and in unknown doses in the real world.

A more specifically defined statement of the correction processes and awareness of the truth and falsity of an idea has been stated by Bartlett (1958), in connection with thinking in problem solving. Bartlett treats thinking as a succession of steps to fill in gaps that exist in the given evidence. In order to be able to select and arrange evidence (correction\textsubscript{2} and correction\textsubscript{1}), it is necessary that the thinker be able to detect points of agreement and difference in the evidence (awareness).
Ability to detect agreement and difference in evidence (reality awareness) and ability to select and arrange evidence (correction processes) are the dependent variables in the present study.

The following rationale underlies the present experiment:

1. Schilder stated that the schizophrenic does not employ sufficient correction processes in his thinking and does not arrive at an awareness of the truth or falsity of the completed thought.

2. If this occurs in the development of one thought production, it has been assumed that it also will occur, by the circular effect described, over successive recall of the same material. That is, there will be minimal correction processes employed and minimal truth awareness; hence, recall will improve slowly.

3. If the opportunity is presented to an experimental group of schizophrenics to compare the story to be recalled with their own reproductions of it, they are artificially being given more opportunity to detect agreement and differences in the evidence than a control group of schizophrenics not presented with this condition.

4. If the experimental group does not profit from this comparison by improving recall over successive trials at a greater rate than the control group, there is indication that the ability to detect the agreement and difference in the evidence does not initiate correction in successive recall.
In other terms that have been used, there is indication that schizophrenics are lacking in their awareness of the truth or falsity of their reproductions, or, at least, this awareness does not initiate corrections.

**Hypotheses.** In the present experiment, four groups were used, a matched experimental and control group of normal subjects, and a matched experimental and control group of schizophrenic subjects.

From the foregoing rationale, it is predicted that:

1. The difference between the normal experimental and control groups in the mean number of trials to reach successive criteria of recall will increase, reaching a significant difference. The experimental group will have the lower mean.

2. There will be no significant difference between the schizophrenic experimental and control groups in the mean number of trials to reach successive criteria of recall.
CHAPTER III

METHOD AND PROCEDURE

I. SUBJECTS

The setting for the experiment was the Veterans Administration Neuropsychiatric Hospital, Chillicothe, Ohio. Subjects were 40 male, chronic schizophrenic patients and 30 male nursing aides at the hospital.

Patients. Schizophrenic patients who were homogeneous in that they showed manifestations of thought disorganization were selected by pretesting (see below). In addition to thought disorganization, the sample was selected to be homogeneous in the following ways:

1. All patients were male and of the Caucasian race.
2. All patients carried the diagnosis of schizophrenia.
3. All patients had a chronic pathological condition, determined by the length of their hospitalization or by repeated hospitalizations.
4. No patient had any known organic brain damage. This criterion excluded lobotomized patients from the study.
5. All patients were testable in that they completed the pretests.
6. All patients were receiving a tranquilizing drug at the time of testing. The majority of the patient subjects had been
on a stable maintenance dosage of a drug for over one year.

Kraepelinian subdiagnoses were not considered in selecting
the patients.

None of the patients selected was undergoing somatic
therapy during the research; however, 24 were known to have
received insulin or electro-shock therapy at this hospital
sometime in the past. It was suspected that others had received
these somatic treatments while residing in other hospitals.

The majority of the patient sample had delusions of some
variety, many of which were still evident at the time of testing.
Records indicated that almost all patients had had hallucinations
and violent or confused periods sometime in the course of their
hospitalizations.

Patients were excluded from the study if there were indi­
cations in their levels of education, vocabulary scores, or
retention scores that they did not have the potential ability
to learn the experimental story.

Normals. The normal subjects were psychiatric nursing
aides with assignments throughout the hospital. All subjects
were volunteers, having been asked by their supervisors or the
experimenter if they would participate in a psychological study.

No criteria were used in the selection of normal subjects
except that they be under age 55 and not have had more than one
year of education beyond high school.
II. PRETESTING PROCEDURES

The pretest session consisted of the administration of four short tests, which provided data for the subsequent matching of subjects for the experiment. At the beginning of the test session, the patients were told:

I am a psychologist on the station, Mr. __________. We are doing research on some of the factors influencing memory, that is, what things people remember and how they remember things, and we would like to have you help out if you would. This will mean spending about a half-hour today, and another half-hour some other day this week. Would you be willing to do that?

If the subject said "No," or showed some hesitancy with expressions such as, "Well, my memory isn't very good," the experimenter replied:

We have been testing many of the people in this building, and we would also like to have you participate, if you would.

The experimenter answered other questions, or if hesitancy continued, explained a little more about the research. No information was given, however, about what would be done at the second session. This approach was successful with all the patients, except the most resistive; and with few exceptions testing was begun with an expression of the patient's consent. The exceptions were those cases where there was no expression of agreement or disagreement to the question, "Would you be willing to do that?" The procedure here was to assume agreement and move on to the next instructions. It was soon determined whether the subject would be passively cooperative or mute.
With the normals, these initial comments were varied somewhat since these subjects already knew when they came to the test session that they were volunteering to be subjects for research. They were asked:

Have you ever participated in any of the research that has been conducted at the hospital? (then, regardless of the answer) Well, from time to time the hospital conducts research to find out various things, and sometimes it is necessary to ask some of the hospital personnel to volunteer as subjects. As you know, this in no way has anything to do with your job, and none of the things you do here are used for evaluation of your work.

This research has to do with some of the things that influence memory, that is, what things people remember and how they remember things, and we would like to have you help out if you would. This will mean spending about a half-hour today, and another half-hour some other day this week. Would you be willing to do that?

After this initial introduction, the examiner stated:

Today I would like to ask you some questions. I use this tape recorder to record the answers so that I won't have to write them down, and we can finish sooner. Have you ever heard your voice on a tape recorder?

If the subject answered "No," he was told that some of the record would be played back to him. The subject was then asked his age, education, and the place where he was born and raised. Somewhere in this initial questioning the experimenter interrupted, supposedly to see if the recorder was working properly. Some of the record was then played for the subject to hear. The purpose of this was to accustom the subject to the tape recorder and to minimize the distraction that might occur during the
experiment as a result of the subject's hearing his voice from a tape recorder for the first time.

Instructions to the first test were then given. The tests are discussed below in the sequence in which they were administered. Each is described in terms of instructions, administration, and scoring.

**Wechsler Adult Intelligence Scale: vocabulary.** Instructions and administration of the WAIS vocabulary subtest were standard, following the manual for the test (Wechsler, 1955). Standard scoring was followed with one exception, a response was scored correct when a correct response was followed by blatant contamination. It was reasoned that this procedure would give a better estimate of the verbal abilities of the patient subjects.

**Wechsler Adult Intelligence Scale: comprehension.** Instructions, administration, and scoring of the WAIS comprehension subtest were standard (Wechsler, 1955).

This test was included in the battery to obtain an additional verbal sample from subjects that could be scored for disorganization of thought. Standard scoring was used only to determine the number of failures for discontinuing the test. (See thought disorganization score, below.)

**Benjamin Proverbs.** Nine proverbs which had discriminated between normals and schizophrenics in a previous investigation
(Allison, 1960) were selected from the 14 originally used by Benjamin (1944). They were the following:

When the cat's away, the mice will play.
Don't cry over spilt milk.
It never rains but it pours.
The burnt child dreads the fire.
Don't cross your bridges until you come to them.
A rolling stone gathers no moss.
Don't count your chickens before they are hatched.
The proof of the pudding is in the eating.
New brooms sweep clean.

The instructions were these:

I would like you to interpret some proverbs for me now. Do you know what a proverb is? (then, without waiting for an answer) A proverb is a saying that people use, that has some meaning. In each case I would like you to interpret the proverb for me, that is, to tell me what it means. Here is the first one...

If the response to the first proverb was concrete, or a concrete generalization, the examiner asked:

Could that proverb mean anything else? Could it be used to refer to people, for example? (or, could that proverb be applied to people in any way?)

Before each of the successive proverbs, the examiner began:

"What does this proverb mean?"

The proverbs were scored for thought disorganization.

Babcock paragraph. This paragraph is one of the subtests of the Revised Examination for the Measurement of Efficiency of Mental Functioning, by Babcock and Levy (1940). The paragraph is reproduced below.

December 6th, last week, a river overflowed in a small town ten miles from Albany. Water covered the streets and entered the houses. Fourteen persons were drowned and 600 persons caught cold because of the dampness and cold weather. In saving a boy who was caught under a bridge, a man cut his hand.
Instructions, administration, and scoring of the test were the same as those used by Babcock and Levy (1940).

Thought disorganization score. Although patients with thought disturbances are easily distinguished clinically, a measure of this variable has the advantage of enabling the experimenter to describe and match subjects with more objectivity. In other words, the score is an operational definition for this important subject variable.

The score consisted of all responses reflecting thought disorganization that were given on the vocabulary, comprehension, and Benjamin Proverbs tests, described as a per cent of the total responses given on these three tests.

As a guide to identifying deviant responses, descriptions of schizophrenic thought and its manifestations were taken from relevant published literature and placed into six major categories on the basis of the theoretical processes proposed to underlie the descriptions. This guide is included in Appendix A. The major sources for the compilation were Bleuler (1950), Cameron (1938, 1944b), and Benjamin (1944).

It became apparent during this compilation that categories used in the literature are often restricted to descriptions of responses to the particular test that served to elicit the responses. Thus, some categories are better for the description of disorganized responses in word association, others for concept formation, etc. In many instances, categories overlap or appear to describe reflections of the same underlying process. The
attempt was made to make the categories as mutually exclusive as possible. This aim, however, was not entirely successful. There is so little known, or even postulated, about the relationship between extent of thought disorganization and types of disorganized responses given by patients that it was decided not to weight responses, and to combine the scores for the separate categories into a total score. (See Hunt and Arnoff, 1955, for an attempt to scale responses.)

Below are some specific procedures followed in scoring:

1. Multiple scoring for the response to one test item was permitted if the categories assigned to the response did not overlap by definition.

2. With the proverbs, one of the categories of abstraction or literal interpretation was assigned first. If other of the categories applied, they were also scored if they could be judged as separate from the previously assigned category.

3. In order to keep the verbal sample as standard as possible for all subjects, spontaneous comments made before testing was begun, between tests, and after testing was completed were not scored.

4. To calculate a score, the total number of responses was arbitrarily determined to be the sum of all the scored responses on each test, including failures and responses of "don't know." On the vocabulary and comprehension tests, this number equaled the responses up to and including the five and four
consecutive failures that are the criteria for discontinuing these tests (Wechsler, 1955).

All scoring was done by the experimenter. The major control over possible experimenter bias (scoring in favor of his hypothesis, or scoring against his hypothesis as an over-precaution) was that the thought disorganization scores were obtained before subjects were assigned to groups.

A random selection of ten protocols was scored by an independent judge to give indication of the reliability of the score. Out of 541 judgments made, the experimenter and the independent judge agreed on 433, or 80 per cent, as representing disorganized responses. There was 72 per cent agreement on the 541 judgments for the 14 separate scoring categories. Although higher agreement is desirable, the results compare favorably with reliability scoring of word association response categories (Flavell et al., 1958), where the shorter response should permit more definitive categories and lead to less error in judgment. All differences between the percentage scores assigned by the independent judges were within the range in which 75 per cent of the subjects were matched. Agreement on the separate categories is high enough to warrant further investigation of their relationships to other variables, especially with refinement of the scale. This, however, is beyond the scope of the present study.
III. MATCHING

In general, the variables on which the groups were matched were chosen to control for possible effects on (correlation with) rate of learning and, in the patient group, extent of pathology. Because of the large number of variables that it was desirous to control, a technique of approximate matching was employed. This technique is analogous to the multiple cut-off method described by Guilford (1950, p. 466). Subjects were matched as closely as possible on all variables, but close matching on several variables was allowed to compensate for greater differences on a few. The variables that were thought to have the greatest correlation with rate of learning and those variables known to have high reliability were given most consideration in pairing the subjects.

The patient experimental group (EP) and control group (CP) were matched on a man-to-man basis on age, education level, vocabulary raw score, immediate recall score, thought disorganization score, and length of hospitalization. The normal experimental group (EN) and control group (CN) were matched on a man-to-man basis on age, education level, vocabulary raw score, and immediate recall score. By the selection of this particular normal sample, it was also hoped that there would be matching across the patient and normal groups on age, education level, and vocabulary. That is, there would not be significant differences between the means of the two experimental groups, nor between the means of the two control groups.
The vocabulary score, immediate recall score (from the Babcock paragraph), and the disorganization score were discussed in the previous section of this chapter. The method for determining the other matching variables is described below.

**Age.** Age was determined to the nearest year. Since the sample was fairly homogeneous on this variable, between 30 and 50 years of age, it was not considered an important matching variable.

**Education level.** Level of education was determined to be the last year of formal schooling completed. Any part of a year completed was entered as one-half. Although the number of years in school is a gross indication of the education level of an individual, this variable was considered an important matching variable.

**Hospitalization.** This score was taken from the clinical records of the patients. Because records of previous hospitalizations were often incomplete, the score is only a rough estimate of the chronicity of the patients' pathology. The score equaled the number of years of continuous hospitalization, or of repeated hospitalizations, to the present. To include the years back to the beginning of previous hospitalizations that had been separated from the present hospitalization by an interval of over three years, there had to be evidence in the clinical records that the patient had had no work history or had shown no other signs of social adjustment in his interval out of the hospital.
IV. EXPERIMENTAL TASK

The experimental task was the successive recall of a short story modified from one used by Schilder (1951). The story was presented to the subjects through earphones from a tape recorder. Immediately after each recall, the experimental groups heard the recording of their recalls, while the control group had a 15 second waiting period (the independent variable). A short standard time interval of 45 seconds separated each trial. Trials were continued until the experimenter was sure that the subject had reproduced a preset number of components of the story in correct order. It was anticipated that some of the patient subjects might be unable to reach the predetermined criterion of 22 components correct-in-order that was set for the normal subjects. With cases where this did occur, testing was discontinued after the 15th trial. The design of the experiment is summarized in Figure 1.

**Apparatus.** Two tape recorders were used in the experiment. The first contained the story originally recorded by the experimenter and then dubbed on the tape at 45 second intervals. This provided a standard presentation of the story on each trial, without the necessity of rewinding the tape, and an automatic timing of the interval between trials. The second recorder was used to record (and for the experimental group subjects, to play back) the recalls of the story.
FIGURE 1

Design of the Experiment

Patients

<table>
<thead>
<tr>
<th>Group</th>
<th>Hear recall</th>
<th>Feedback of recall</th>
<th>Time hear recall</th>
<th>Etc. story</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>hear story</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>hear story</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Normals

<table>
<thead>
<tr>
<th>Group</th>
<th>Hear recall</th>
<th>Feedback of recall</th>
<th>Time hear recall</th>
<th>Etc. story</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>hear story</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control</td>
<td>hear story</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Each subject wore a headset with earphones and a microphone attached. For the experimental group subject, the line to the earphones was moved from the first recorder to the second after each recall so that he also heard the playback of his recall of the story through the earphones. The microphone, attached to the headset close to the subject's mouth, gave a clear recording and minimized variations in volume caused by the subject's shifting in his seat or moving his head.

Instructions. The following instructions were given at the beginning of the experiment. The parts enclosed by parentheses were given to the experimental group subjects only.

You recall that the other day I read you a story and asked you to tell me all you could remember of it? Well, today I am going to have you listen to another story through these earphones, but this time the procedure will be a little bit different. This time I would like you to **learn** a story.
This is what we will do. First, listen carefully to the story, and as soon as it is over, repeat as much of the story as you can remember, like you did the other day. (Then, listen carefully again, and you will hear the story you have just given, played back to you from this tape recorder.)

Then, there will be about a minute pause, after which you will hear the story again. And, when the story is finished, again repeat back as much of it as you can remember. We will do this several times, until you have learned most of the story.

So, each time retell all you remember of the story. (And also each time you will hear what you have said played back to you.) You should not feel rushed or hurried. Just sit back and relax, and when the story comes on listen carefully and after it is over, try to repeat as much of it as you can. This story will be a little longer than the one I read you the other day, so, if you can't remember how it started or much about it the first time, just tell me anything about the story you do remember.

When you have the headset on, you will not be able to hear me, so, let's review the instructions again. Can you tell me what we are going to do?

The subject was then helped, if necessary, to repeat the essential parts of the instructions. Recorded on the tape before the original story were the following instructions: "Ready? Here is the story. Listen closely."

The subjects were not allowed to start over, once they began to tell the story. If they asked whether they could start again, or if they blocked, or paused for more than ten seconds, they were told, "Go ahead, and tell me anything else you remember," or, "Is there anything else you remember?"

The story. To ascertain the effect of the playback of recalls on rate of learning, it was necessary to use a story that would require several trials for most individuals to learn. In
order to use the same story with both normal and patient groups, however, the story needed to be short enough not to discourage the patient subjects, nor to introduce too much confusion into their reproductions. An estimation of the optimal length was arrived at by pilot studies where the length of the story was varied.

The story was divided into 45 "meaning components" for scoring. It is given below with slash marks separating the components:

On the morning / of December / the 8th, / the son / of a wealthy / Hungarian / landowner / was swallowed / by a shark. / He was the first / mate / on a Polish / freighter, / where, five / miles / west of / Borneo, / during a dreadful / storm, / he was washed / overboard. / In the blinding / rain / and tossing / waves, / rescue / was impossible, / and in front of / the eyes of / his terrified / shipmates, / the unfortunate / man / was grabbed / on the leg / by a huge / shark / which had been following / the steamer. / He was dragged / to the bottom / of the ocean / leaving behind him / a dark / streak / of blood. /

V. SCORING OF THE EXPERIMENTAL STORY

The number of correct components in correct order was used to test the central hypotheses of the study. The major concern in developing a scoring system for the story was to find criteria that were as objective as possible for determining correct components in the subjects' reproductions.

The following scores were used in secondary analyses of the data: synonyms, distantly related components, very distantly related components, and added components. Order was accounted for by the scoring categories: correct components out-of-order,
synonyms out-of-order, etc. Components that were out of order in such a way that the meaning of the component or adjoining components was changed were scored as correct component out-of-order and change-of-meaning, etc.

A scoring category was developed to account for all the components in the subjects' reproductions of the story. Not all categories, however, were analyzed in the study. The complete list of categories, with the symbol for each, is given below:

<table>
<thead>
<tr>
<th>Symbol for Category</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>Correct component in order</td>
</tr>
<tr>
<td>S</td>
<td>Synonym in order</td>
</tr>
<tr>
<td>D</td>
<td>Distantly related substitute in order</td>
</tr>
<tr>
<td>V</td>
<td>Very distantly related substitute in order</td>
</tr>
<tr>
<td>XO</td>
<td>Correct component out of order but not changing the meaning</td>
</tr>
<tr>
<td>SO</td>
<td>Synonym out of order but not changing the meaning</td>
</tr>
<tr>
<td>DO</td>
<td>Distantly related substitute out of order but not changing the meaning</td>
</tr>
<tr>
<td>VO</td>
<td>Very distantly related substitute out of order (by definition, changes the meaning)</td>
</tr>
<tr>
<td>XOM</td>
<td>Correct component out of order and changing the meaning</td>
</tr>
<tr>
<td>SOM</td>
<td>Synonym out of order and changing the meaning</td>
</tr>
<tr>
<td>DOM</td>
<td>Distantly related substitute out of order and changing the meaning</td>
</tr>
<tr>
<td>A</td>
<td>Added component (by definition, no order can be scored)</td>
</tr>
<tr>
<td>P</td>
<td>Parenthetical remark</td>
</tr>
<tr>
<td>R</td>
<td>Repetition of a component</td>
</tr>
</tbody>
</table>

All protocols were scored by the experimenter. In order to make the scoring as objective as possible, a list of examples (Appendix B) based on criteria for determining the separate
scoring categories was constructed prior to scoring the protocols. The criteria for the separate categories are given below.

**X score.** A component was scored correct when it was either an exact reproduction of the component in the original story, or was an accepted modification meeting one of the following criteria: (1) Change in the component's grammatical structure was the result of the absence of the preceding component; (2) the substituted component was a variant of the original component that was used by the subjects commonly and consistently enough to be classed as dialectical (e.g., mile substituted for miles in the phrase, "five miles west of Borned"); and (3) the change in wording could be attributed to transcription error (e.g., "a" and "the").

In order to obtain the highest reliability for the X score, the following procedure was followed with regard to order and corrections in the subjects' reproductions of the story: (1) the phrase containing the most X components was scored when subjects made corrections within a recall, and (2) the phrase containing the most X components was scored in order when phrases in the reproduction were out of order.

**S score.** A component of the subjects' reproductions was scored as a synonym when (1) the substitute wording was an acceptable modification of the original wording (Appendix B) or (2) a recognized synonym was substituted for the major word in the component, in the meaning that the original word carries in the context of the story.
**D score.** A component was scored as distantly related if it belonged to the meaning sphere of the original component (e.g., fish or whale as substitute components for "shark").

**V score.** A component was scored very distantly related if the substitute wording could be considered only vaguely related to the original component, but where the context indicated that the subject's component was a substitute for a component in the story.

**A score.** A component was scored as added when (1) it was unrelated to the components of the original story and (2) when it was not an off-task, parenthetical remark.

**Out-of-order scores.** The XO, SO, DO, and VO components were based on the same criteria as the X, S, D, and V scores, respectively. The assignment of components to categories was more difficult when the component was more distant from the correct component in meaning. When components were also out of order, surrounding components were the major guide to determine which category was applicable.

**Change-of-meaning scores.** Components were scored XOM, SOM, and DOM when they were displaced and combined with components from other parts of the story or with added components in such a way that the meaning of the component was changed. Since the criterion for V components implies a change of meaning, the possible category VOM was not used.
Interjudge reliability of the scoring. A random sample of 30 stories, taken from 10 subjects, was scored independently by a colleague of the experimenter to obtain an indication of the reliability of the scores. Over the 879 judgments made, the two scorers agreed on 737, or 84 per cent. To ascertain the percentage of agreement for the correct-in-order scores, the sum of the differences between the judges in assigning components this score was divided by the mean of the total number of components given this score by each judge. The percentage of agreement for the correct-in-order components was 94 per cent.

VI. ANALYSIS OF THE DATA AND EXPERIMENTAL HYPOTHESES

To test the central hypotheses of the study, average learning curves were computed for each of the groups by the method proposed by Melton (Woodworth and Schlosberg, 1954, p. 535). The curves represent the average number of correct components in correct order acquired over trials. The slopes of the curves were compared by inspection, and tests of the mean difference in the number of trials to reach the successive criteria (number of correct components) were used to determine significance levels of the differences.

The experimental hypotheses were as follows:

1. There will be a difference between the slopes of the curves for the normal experimental group and the normal control group, the curve for the normal experimental group being the steeper. The difference between the normal experimental and
normal control groups in the mean number of trials to reach successive criteria of recall will increase, reaching a significant difference at about the 22nd criterion (half the components of the story).

2. There will be no difference between the slopes of the curves for the patient experimental group and the patient control group, and the difference between the patient experimental and patient control groups in the mean number of trials to reach the 22nd criterion of recall will not be significant.

Two additional analyses of the data were made. The first consisted of a comparison between the combined patient groups and the combined normal groups in their usage of substitute, added, and out-of-order components. The hypotheses for this analysis were drawn from the findings of other investigators, and are presented in Chapter IV.

The second analysis was designed to ascertain what effects, if any, the independent variable (feedback of recall) had upon the subjects' usage of substitute, added, and out-of-order components. For the analysis, the two experimental groups and the two control groups were combined. Since both the normal and the patient groups were matched, the comparisons made between the groups combined in this manner were still comparisons of matched groups. Any differences that exist between the patient and normal samples on the compared categories should cancel out, leaving those differences that can be attributed to the independent variable. This procedure assumes there is no interaction
between the patient-normal condition and the experimental-control condition. This analysis was exploratory, and no predictions were made concerning the results.

In both the analyses described above, the procedure was to plot the mean usage of the separate categories over trials. Statistical tests were used only where inspection of the plots suggested a difference might be significant. The Kolmogoroff-Smirnov Test (Siegel, 1956) was used to test for significant differences.
CHAPTER IV

RESULTS

I. EVALUATION OF MATCHING

The experimental patient and control patient groups were matched on age, education level, vocabulary raw score, immediate recall score, thought disorganization score, and length of hospitalization. The experimental normal and control normal groups were matched on age, education level, vocabulary raw score, immediate recall score, and thought disorganization score. These groups were compared by use of mean difference tests for matched samples.

No significant differences were found on any of the variables on which the groups were matched (Table I).

Attempts were made to match the two experimental groups and the two control groups on age, education, and vocabulary raw score. It was expected that these groups (patient vs. normal) would differ on the thought disorganization score and the immediate recall score. These groups were compared by use of mean difference tests for independent samples (Table II).

No significant differences were found between the two experimental groups or between the two control groups on education level and vocabulary raw score, and neither of the mean differences on the immediate recall score was significant at the
<table>
<thead>
<tr>
<th>Variable</th>
<th>Experimental Mean</th>
<th>Standard Deviation</th>
<th>Control Mean</th>
<th>Standard Deviation</th>
<th>t</th>
<th>Level of Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>37.40</td>
<td>6.23</td>
<td>39.25</td>
<td>5.24</td>
<td>1.210</td>
<td>&gt;.20</td>
</tr>
<tr>
<td>Education</td>
<td>10.63</td>
<td>1.57</td>
<td>10.58</td>
<td>1.49</td>
<td>.129</td>
<td>&gt;.80</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>38.45</td>
<td>9.70</td>
<td>39.05</td>
<td>11.33</td>
<td>.764</td>
<td>&gt;.40</td>
</tr>
<tr>
<td>Recall</td>
<td>8.90</td>
<td>3.35</td>
<td>8.05</td>
<td>3.04</td>
<td>1.545</td>
<td>&gt;.10</td>
</tr>
<tr>
<td>Disorganiz.</td>
<td>45.35</td>
<td>16.16</td>
<td>40.20</td>
<td>14.85</td>
<td>.406</td>
<td>&gt;.10</td>
</tr>
<tr>
<td>Hospitaliz.</td>
<td>8.45</td>
<td>3.74</td>
<td>9.75</td>
<td>4.25</td>
<td>.992</td>
<td>&gt;.20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I. Patient Groups (N=20)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>30.80</td>
<td>6.64</td>
<td>29.67</td>
<td>4.89</td>
<td>.662</td>
<td>&gt;.50</td>
</tr>
<tr>
<td>Education</td>
<td>10.73</td>
<td>1.35</td>
<td>11.03</td>
<td>1.32</td>
<td>1.132</td>
<td>&gt;.20</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>33.53</td>
<td>8.67</td>
<td>34.40</td>
<td>9.21</td>
<td>.652</td>
<td>&gt;.50</td>
</tr>
<tr>
<td>Recall</td>
<td>9.80</td>
<td>2.69</td>
<td>9.80</td>
<td>2.59</td>
<td>.227</td>
<td>&gt;.70</td>
</tr>
<tr>
<td>Disorganiz.</td>
<td>14.40</td>
<td>6.31</td>
<td>14.53</td>
<td>6.07</td>
<td>.227</td>
<td>&gt;.70</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>II. Normal Groups (N=15)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

TABLE I

Means, Standard Deviations, and t Test for Mean Differences on the Matched Variables: (I) for Experimental Patient and Control Patient Groups, and (II) for Experimental Normal and Control Normal Groups
TABLE II

Mean Differences and t Tests for Unmatched Groups Between the Two Experimental Groups and the Two Control Groups on the Variables of Age, Education, Vocabulary, Immediate Recall, and Disorganization.*

<table>
<thead>
<tr>
<th>Variable</th>
<th>Experimental Groups</th>
<th>Control Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M_{diff}</td>
<td>t</td>
</tr>
<tr>
<td>Age</td>
<td>6.60</td>
<td>2.928</td>
</tr>
<tr>
<td>Education</td>
<td>-.11</td>
<td>.207</td>
</tr>
<tr>
<td>Vocabulary</td>
<td>4.92</td>
<td>1.508</td>
</tr>
<tr>
<td>Recall</td>
<td>-.90</td>
<td>.843</td>
</tr>
<tr>
<td>Disorganiz.</td>
<td>30.95</td>
<td>6.822</td>
</tr>
</tbody>
</table>

*A positive mean difference indicates that the patient group has the higher mean. In both the experimental and control groups N_1 = 20, N_2 = 15.

.05 level. The unexpected lack of significance between normals and patients on this last variable indicates that although the normal sample was not as representative of the normal population as some others might have been, it is more comparable to the schizophrenic sample on the matching variables chosen to control for differences in rate of learning.

The mean age difference was significant between the experimental groups and between the control groups, the normal groups being younger. The difference between these groups was also significant on the thought disorganization score, the ranges of the two samples barely overlapping.
Correlations of the matching variables with the experimental criterion. Pearson product moment correlations were computed between matching variables and the number of trials to reach the 16th criterion of the experiment for the combined normal and combined patient groups, to determine whether the matching procedure controlled relevant variables.

All correlations but two were in the expected direction, and all but three were significantly different from a zero correlation at the .05 level (Table III).

In matching the experimental and control groups, education level, vocabulary score, and immediate recall score were considered the most reliable and pertinent of the matching variables to control for rate of learning, and subjects were matched more closely on these variables than on the others. The correlations of these variables with the measure of rate of learning in the experiment supported this close matching procedure. In the normal groups, these variables were the three highest predictors of the rate of learning measure. The correlations for the patient groups were not as high as those for the normal groups. Correlations for three variables (vocabulary, immediate recall, and thought disorganization), however, were significantly greater than zero.

The correlations of age with the rate of learning measure were not significantly different from zero for either the patient groups or the normal groups. These low correlations suggest that the significant differences found on this variable
TABLE III

Correlations of Vocabulary Raw Score, Education, Immediate Recall, and Age with the Trials to Reach the 16th Criterion in the Experiment for the Combined Patient Groups (N=40) and the Combined Normal Groups (N=30).

<table>
<thead>
<tr>
<th>Correlation</th>
<th>Patient Groups</th>
<th>Normal Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>r</td>
<td>Level Signif.</td>
</tr>
<tr>
<td></td>
<td>Different from Zero</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-.09</td>
<td>Not Signif.</td>
</tr>
<tr>
<td>Education</td>
<td>.27</td>
<td>p &lt; .10</td>
</tr>
<tr>
<td>Vocabulary Recall</td>
<td>-.54</td>
<td>p &lt; .001</td>
</tr>
<tr>
<td>Disorganiz.</td>
<td>-.37</td>
<td>p &lt; .02</td>
</tr>
</tbody>
</table>

between the two experimental groups and the two control groups would not result in differences on the dependent variable of the experiment.

Several of the correlations are spuriously low because of the restricted ranges in the samples. Since the standard deviation of the population on the vocabulary test is known (Wechsler, 1955, p. 24), an estimate of this correlation with an uncurtailed sample could be made (McNemar, 1955, p. 150). The actual correlation for the patient groups was -.54, and was the highest predictor of the criterion for the patient sample. The estimated correlation with an uncurtailed sample, computed by
approximating the raw score equivalent of the scale score standard deviations, was of the magnitude of \(-.70\) to \(-.80\).

II. THE CENTRAL HYPOTHESES

**Normal groups.** Melton average curves for the experimental and control normal groups (Figure 2) support the major hypothesis that normals do profit from feedback of their recalls to initiate corrections in their reproductions.

The mean difference in the number of trials to reach the 19th, 22nd, and 24th criteria of acquisition were tested by the \(t\) test for matched groups. The probability level of the differences decreases, reaching statistical significance at the 24th criterion. The \(t\) tests and their significance levels for the three criteria were as follows (two-tail test):

- 19th \(t = 1.545\) \((.20 > p > .10)\)
- 22nd \(t = 1.932\) \((.10 > p > .05)\)
- 24th \(t = 2.180\) \((.05 > p > .02)\)

Although the two groups were matched on immediate recall score before the experiment, the mean number of correct components given on the initial trial of the experiment differs for the experimental and control groups. The difference is not significant \((.20 > p > .10)\); it suggests, however, that some caution should be taken in accepting the above results. On the one hand, close matching of the groups on the pretest recall would make it more likely, by chance alone, that they would differ on a second performance. On the other hand, the
FIGURE 2

Melton Average Curves for Normal Experimental and Control Groups
experimental situation was not the same as that of the pretest (subjects wore a headset, were told to learn a story, and the story was longer), and the difference may have been the effect of an uncontrolled variable.

Since the difference between the number of correct components recalled on the initial trial was presumably controlled, and the source of the variance could not be experimentally controlled after the matching, an analysis of variance with covariance adjustment for the initial recall was computed to test the mean difference in the number of trials to reach the 24th criterion (McNemar, 1955, p. 343). This covariance adjustment for the differences on the initial trial indicates whether the two groups would have differed in trials to reach the criterion if their initial recall scores had been identical. After this statistical control, the difference in the trials to reach the 24th criterion only approaches the .05 level of significance (Table IV).

**Patient groups.** Approximately half of the patient subjects (1) reached an extended plateau in the acquisition of correct-in-order components or (2) added to the number of correct components from trial to trial at such a slow rate that even after 15 trials they had not reached the criterion of 22 correct-in-order components. Most of these subjects gave the impression that they were actively trying to learn the story. Their extreme fluctuation in the ordering of components and in the different components given on each trial is not reflected in the scores they received. It
TABLE IV

F Test for Difference Between the Normal Experimental and Normal Control Groups in Trials to Reach the 24th Criterion with a Covariance Adjustment for Differences in the Number of Correct Components Recalled on the First Trial of the Experiment

<table>
<thead>
<tr>
<th>Sum of Squares</th>
<th>df</th>
<th>F</th>
<th>Level of Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>189.45</td>
<td>29</td>
<td>4.47</td>
</tr>
<tr>
<td>Within</td>
<td>163.34</td>
<td>28</td>
<td>4.47</td>
</tr>
<tr>
<td>Between</td>
<td>26.11</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

*In agreement with the previously computed t test

can be inferred from their response patterns, however, that these patients had extreme difficulty in organizing and integrating the story around a core concept of the story. The number of patients who showed marked inability to learn the story did not differ between the experimental and control groups. In a general sense, this finding supports the hypothesis that the experimental condition would not lead to a difference in the performance of the two groups. In order to compare the patient performance at the level that the curves of the normal groups separate, however,
the following two analyses were made: (1) to give a picture of the acquisition curve for the patient sample, the Melton average curves for the complete groups (20 pairs of subjects) were computed; (2) to compare the patient curves with those of the normal sample, the ten pairs of patients who learned the story past the 22nd criterion were treated as a separate sample.

Recomputation of the mean difference tests for this smaller sample on the matching variables showed no significant differences between the experimental and control groups. In fact, on all variables except education level, the mean differences between the groups were smaller than those between the complete experimental and control groups.

Average curves for the complete experimental and control groups are presented in Figure 3. A test for the significance of the mean difference in the number of trials to reach the 16th criterion supports the hypothesis for the patient group $t = 1.487; .20 > p > .10; \text{two-tail test}$. As with the normal groups, there was a difference between the patient experimental and control groups on the number of correct-in-order components recalled on the initial trial; this difference, however, was not significant ($20 > p > .10$). The control group had the higher mean. On the pretest immediate recall score, the mean for the experimental group was higher than the mean of the control group ($20 > p > .10$). This reversal of the means of the two groups on the two recall tasks suggests that the differences are chance fluctuations, especially since
SUCCESSIVE CRITERIA
NUMBER OF COMPONENTS CORRECT IN ORDER

FIGURE 3
they are within expected limits of chance occurrence. Inspection of the slopes of the curves for the two patient groups further suggests that the groups do not differ in their rate of acquisition of correct components.

In spite of this evidence in support of the hypothesis, however, the same cautions presented in regard to the normal groups must be made. Differences between the patient groups in recall of correct components on the initial trial may be the result of an uncontrolled variable, which also correlates with the dependent variable.

The difference between the Melton average curves for the ten experimental and ten control group patients who continued to learn beyond the 22nd criterion (Figure 4) is not significant at the 22nd criterion (.60 > p > .50). It was not possible to test the difference between these patient curves at the level where the normal curves approach a significant difference. Inspection of the two sets of curves shows that the curves for this higher ability group of patients accelerate in the same manner as the normal group curves until the 20th criterion, but then level off sharply. This change in slope, which is not present in the normal group curves, is similar to the leveling off seen in the curves of the total patient groups.
Melton Average Curves for Experimental and Control Group
Patients Learning to the 22nd Criterion of the Experiment

FIGURE 4

CONTROL (N=10)

EXPERIMENTAL (N=10)
III. SECONDARY ANALYSES

Comparison of the combined patient groups with the combined normal groups. The purpose of this analysis was to see whether the results of this experiment supported the findings of other investigators studying thought in schizophrenics with other methods.

Before presenting the results of this analysis, a more detailed description of the findings of two recent studies mentioned in Chapter I is presented.

Flavell (1955, 1956), asking normal and schizophrenic subjects to select words that were "closest in meaning" to the stimulus words, found significantly more abstract words in the choices made by the normal group, and significantly more homophone words in the choices made by the patient group. Although not significantly different, the means of the normals were higher on all the abstract categories, and the means of the patients were higher on all but one of the "nonessential meaning" categories. Of the abstract categories used by Flavell, the mean difference for "synonyms" was the largest.

Lipton (1954) divided schizophrenic patients into two groups by their high or low performance on the Goldstein-Scheerer Sorting Tests, and with a normal control group investigated the characteristics of immediate and delayed recall of a story. The low abstraction group of schizophrenics used by Lipton would be similar to many of the patients in the present study. His normal
control group also consisted of Veterans Administration Hospital employees. The general theoretical basis of the study was that of Werner (1948), and the author hypothesized that schizophrenics would show the same primitive characteristics in memory that are reflected in their productive thought processes.

Lipton's scoring categories were compared with the ones in this study, and several were found to be similar. In general, the differences in the scoring and procedure in the two studies were that Lipton used a longer story and employed larger components for scoring. Some of his categories appeared to be combinations of two in the present study and vice versa. Some of Lipton's results, pertinent to the analysis of the present data, are presented below.

The following categories were described by Lipton as representing a decreasing degree of articulation, or an increasing individual predilection for substitution: (1) synonym, which appears similar to the S score in this study, (2) analogous situation, which is closest to the D and V scores in this study, and (3) simple distortion and distortion, which are related to the A score in this study.

The findings for these scores were (1) that significantly more synonyms were used by normals (the category representing the least primitive change), (2) that there were no significant differences between the groups on the use of analogous situations, and (3) that simple distortion and distortion were
nonexistent in the normal groups, and more frequent in the low abstraction group than in the high.

The low abstraction group showed significantly more non-meaningful displacements and significantly more meaningful displacements of components than the control normals.

With the above comparisons in mind, the following specific predictions were made for the comparison of the normal and patient groups in the present study:

1. The combined normal groups will show more use of synonyms as substitute categories over successive recalls.

2. Although not significantly different, the normals will be higher in the use of distantly related substitutes than the patients (D score).

3. Although not significantly different, the patients will be higher in the use of very distantly related substitutes than the normals (V + VO score).

4. The patients will be significantly higher than the normals in the addition of components not related to the story (A score).

5. The patients will be significantly higher than the normals in the mean number of correct components out-of-order in their recalls (XO score).

6. The patients will be significantly higher than the normals in the mean number of components out-of-order and
changing the meaning in their recalls (sum of the XOM, SOM, and DOM scores).

To compare the data, the mean number of responses for each of the categories considered was computed for the first seven trials. These means were plotted for the combined normal groups (N = 30) and the combined schizophrenic groups (N = 40), so that the increase or decrease in the mean usage over trials could be inspected.

Inspection of the plots generally showed the predicted trend from higher use of the synonym category by the normal groups to the higher use of the added category by the patient groups. Not all of the predicted differences, however, were found. The separate hypotheses are treated specifically below.

1. The normal groups used increasingly more synonyms as substitute components over trials than did the patient groups (Figure 5). The Kolmogoroff-Smirnov Test for differences at the fifth trial showed a significant difference between the two groups (p < .01; one-tail test).

2. and 3. There was no difference between the two groups on the mean usage of either the distantly related substitutes or the very distantly related substitutes (plots not shown). The hypotheses in regard to these scores were not supported.
FIGURE 5

Mean Usage of Synonyms over Trials for the Combined Patient and Combined Normal Groups

4. The patient groups consistently used more added components than the normal groups (Figure 6); however, the difference between the groups at the fifth trial only approached significance ($10 > p > .05$, one-tail test). This hypothesis was only partially supported.

5. The patient groups consistently had more correct but out-of-order components in their recalls than did the normal
groups (Figure 7). This difference at the fifth trial also only approached significance (.10 > \( p > .05 \); one-tail test).

6. The components scored out-of-order and changing the meaning were used consistently more by the patient groups than by the normal groups (Figure 8). At the fifth trial, the Kolmogoroff-Smirnov Test for differences was significant \( (p < .01; \) one-tail test).

Comparison of the combined experimental and combined control groups. For this analysis, the patient and normal experimental groups and the patient and normal control groups were combined \((N = 35 \) in each group). These combined groups were compared on the same scoring categories as the combined groups in the analysis above. Plots of the mean usage of the
FIGURE 7

Mean Number of Correct Out-of-Order Components over Trials for the Combined Patient and Combined Normal Groups

FIGURE 8

Mean Number of Components Out-of-Order and Changing the Meaning for the Combined Patient and Combined Normal Groups

(Sum of the XOM, SOM, and DOM Scores)
categories over trials were also made to inspect for differences that could be attributed to the experimental condition (receiving feedback).

Inspection of the plots (not shown) indicated that there were no significant differences between the two groups on any of the categories or combined categories. Only two scores showed consistent differences between the groups. The mean use of synonyms was consistently higher for the experimental groups, and the mean use of added components was consistently higher for the control groups, except on the initial trial where, on both scores, the means of the groups were the same.

The mean use of synonyms, in both groups, increased from 1.5 synonyms on the first trial to approximately 4.0 synonyms on the seventh trial. This is the only score, other than the correct-in-order score, that showed increased usage over trials. The mean usage of the other categories, as in the comparison of the normal and patient groups, remained relatively constant over the first seven trials.
CHAPTER V

DISCUSSION

This chapter is devoted to the discussion of (1) the results of the experiment, (2) the theoretical implications of the results, (3) suggested improvements for future experiments, based on observations made during the course of the study, and (4) suggestions for further research.

I. DISCUSSION OF THE RESULTS

The findings that support the central hypotheses of the study are the differing slopes of the curves between the normal experimental and normal control groups, and the lack of this difference between the patient experimental and patient control groups. The results of the tests of the central hypotheses should be accepted with caution, however, because (1) the experimental and control group means of correct components recalled on the initial trial of the experiment tend to differ (.20 > p > .10), and (2) the patient groups did not learn the story to a level where their performances could be compared directly with those of the two normal groups.

The curves for the normal groups show an increasing difference over trials, with the experimental group improving more quickly. This difference approaches significance at the .05
level at the 24th criterion, when difference in recall on the initial trial is equated by statistical control. The patient experimental and control groups were also different on the recall of correct components on the initial trial. The curves for these groups, however, do not show the same increasing separation seen between the curves of the normal groups.

The general finding that some patients did not improve, or improved very little, over successive reproductions of the story supports the hypothesis that the schizophrenics would not profit from feedback to initiate corrections in successive recalls. Objective support for the hypothesis, however, is limited to those patients who did learn the story to the 22nd criterion of the experiment. The slopes of the curves for these ten pairs of patients do not differ, nor are there significant differences in the mean number of trials taken to reach the 22nd criterion (.60 > p > .50).

Differences in the initial recall of the groups suggest that further control in replication of the experiment would be necessary to test the hypotheses adequately. The failure to obtain a clearly significant difference between the normal groups suggests that further experimentation with the task material may be necessary to support the results found in this study.

Comparison of the combined normal and combined patient groups on the mean number of substitute, displaced, and added components in each recall supported the findings of previous investigators (1) that normals use more substitute components that
are close in meaning to the original component than do schizophrenics, and (2) that schizophrenics add, displace, and change the meaning of components more than do normal subjects. In comparing out-of-order components, only the mean usage of the correct out-of-order components was reported. Including the synonyms and distantly related components out of sequence in the comparison (sum of the XO, SO, and DO scores) would probably show a significant difference between the normal and patient groups.

Although the differences between the combined experimental groups and the combined control groups on the mean use of synonyms and added components are slight and within expected limits of chance, there are two characteristics of the data that suggest the differences may result from the experimental condition (feedback of recall). Of the six scoring categories considered, the largest mean difference between the groups on the initial trial was only .07 of a component. Differences between the groups occurred on the second trial, after the feedback had been introduced. On the two scores mentioned, where the means were different on the second trial, the difference remained in the same direction over all the trials compared.

If synonyms are thought of as substitute components given because the context is remembered enough to enable the subject to approximate the unremembered correct component, then the normal experimental group, acquiring more correct components as the result of feedback, should also be expected to acquire
more close substitute components. It could also be that the control groups, not receiving playback of their recalls, were more prone to add components in the attempt to reproduce the story; whereas, the experimental groups centered their recalls around their previously given reproductions.

The above interpretations are given only tentatively. The general conclusion is that the experimental and control groups did not differ, or differed only slightly, in the frequency with which they used the substitute components.

II. THEORETICAL IMPLICATIONS

The experiment, as it was designed, measured whether feedback of the subjects' recalls would enable the subjects to learn the correct components of a short story more quickly than subjects who did not receive this feedback. The results of the experiment, with consideration of the cautions given in accepting the results, suggest that this does occur with normal subjects and does not occur with schizophrenics exhibiting manifestations of thought disorganization. In Chapter II, it was proposed that the presence or absence of the feedback of recall provided a test of the theoretical construct, the ability to detect agreement and differences in evidence. This construct was equated with the constructs of reality feeling and awareness of the truth or falsity of a thought. The improvement in recalling correct components over successive trials in the experiment was presented as the behavioral consequence of selection and
The process enabling selection and arrangement of evidence was termed the correcting process or the reality testing process in thinking. By this logic, the results of the experiment give some support to Schilder's statement that truth awareness is lacking in the thought processes of schizophrenics and, therefore, corrections are not initiated.

The possibility remains that the schizophrenic may be aware of differences and agreements in his productions, but is still unable to correct them.

Flavell (1955) emphasized the importance of an additional stage in the microgenesis of thought, where inner thought is translated into communicative language (Chapter 1). He also contended that the role-taking behavior and socially oriented thinking necessary to carry the thought formation from completed thought to externalized communication is woefully inadequate in the schizophrenic. Assuming for the discussion that Schilder is right in his statement that thinking does not have conscious reality orientation until the final integrated and completed thought, it still remains possible that the schizophrenic does reach the point of truth awareness (reality feeling); however, he does not have the social motivation, or the intention necessary to formulate his thinking into external, consensual language. Schilder emphasized that disturbance in schizophrenic thought begins in the early phases of thought formation, and results in inadequate development of the thought through subsequent phases. There is evidence in the qualitative aspects of Schilder's data.
and the data of this experiment that the characteristics of the initial phase of thought (primary process) described by Schilder are found in the recall of stories. The lack of centrality in the organization of the reproduced story, the sequence alterations, the over-attention to irrelevants, the additions and bizarre distortion, etc., are all present. It is proposed here that disturbance in the externalization phase may occur in schizophrenic thought development over and above that disturbance occurring in earlier phases of thought development.

The concepts of consensual validation, role-taking, and reality testing can be conceived of as applying to both the final reality integrated stage of Schilder and the externalization stage added by Flavell. In the former, these reality or socially oriented correction processes are largely unconscious; whereas in the latter, the adaptation of the thought to consensual language, they are largely conscious. Evidence that schizophrenics may consciously employ what is termed thought disturbance comes from observations of several workers in the field that schizophrenics improve in communication and concept formation task performance when supportive transference has been established (Sher, 1957; Cameron, 1944b; Goldstein, 1959).

In the theoretical framework of the microgenesis of thought, the externalization phase of thought occurs after what has been called awareness. An indication of the importance of this phase as a locus of disturbance in the thought development of schizophrenics, might be determined with the methodology used in this
experiment if awareness could be observed separately from corrections in the recall of a story. Also, research directed at determining the characteristics of inner language and its differences from expressed language may lead to better understanding of this phase of thought development. Unfortunately, this area has had very little investigation.

III. CRITIQUE OF THE EXPERIMENT

The presence in this experiment of what has been termed by Underwood (1957) as an uncontrolled task variable must be considered along with the results of the experiment. In presenting feedback of each recall to the experimental groups, it was demanded that they spend more time in paying attention to the task. Although the effect of this variable may be minimal with the normal sample, it has been observed that schizophrenics often have a marked decrement in their ability to sustain attention on one task. Attempts were made to minimize the effect of this variable; however, no satisfactory way was seen to remove it from the experiment. If, for example, control subjects were given a task in place of the feedback received by experimental subjects, the effects of retroactive inhibition or diverted attention would be introduced. Without previous experimental results with this design, there was no baseline to rule out the effects of other introduced activities. It was decided to replace the feedback with passage of time in the control groups.
Although the instructions emphasized that the subjects tell a complete story on each recall, a few subjects gave only what they could reproduce verbatim. The experimenter's statement, "Tell me all you can remember," was successful in changing this response set on the initial trials; however, it is felt that experimental control would be improved if the rote memorization set were removed by making the instructions more explicit.

As was indicated above, the Babcock paragraph did not prove to be as high a predictor of the recall on the initial trial of the experiment, nor of the learning ability of the subjects, as was anticipated. Further control of learning ability was suggested as necessary to confirm the results of the present experiment.

In the experiment, the time encompassing a recall varied from subject to subject. Limiting the time allowed for recall to a standard interval was considered. This procedure was rejected because the standard interval might present different conditions to the normal and patient groups, handicapping the patients by introducing a speed factor. In research employing all normal or all patient subjects, this control would have fewer disadvantages.

IV. IMPLICATIONS FOR FURTHER RESEARCH

Some implications of the experiment for the theoretical constructs tested have been mentioned above. The experimental design has suggested other possibilities for research in the area of schizophrenic thought. One of the advantages of the design in studying memory, thought processes, or recall of meaningful
material, is that it provides a comparison of several reproductions to a standard stimulus.

One possible variant of the task instructions would be to tell the experimental subjects that the feedback they hear will help them learn the story. If the schizophrenic experimental group were told how to compare differences, would they show superior recall to a matched control group?

Variations in the experimental story could be used to study effects of different material on memory, distortion, or learning. For example, comparison could be made between emotional and objective material, or interpersonal and neutral narratives.

Qualitative aspects of the recalls were not investigated in the present study. This experiment, however, provides material that can be analyzed qualitatively, and also enables an analysis of changes that occur in a succession of recalls. In the present experiment, for example, the tendency to give the story a happy ending was observed in some subjects (both normal and patient). In the initial trials, parts of the story were changed, omitted, or added to give the story a desired ending until, in the face of repeated presentation of the original story, subjects were forced to alter their reproductions. Although following the changes that occur in each component over successive trials is a tedious task, this type of scoring would give a detailed analysis of qualitative changes in the reproductions of the story.

The reader is referred to Curran and Schilder (1937) and Lipton (1954) for other types of scoring categories that can be applied to qualitative analysis of recalls.
CHAPTER VI

SUMMARY

Schilder (1951) described thought in terms of phases of thought development, proceeding from affective images and associations in the "preparatory phase," to the unification of reality relevant elements into an integrated and complete thought in the "final phase." Schilder considered schizophrenic thought to be the result of the premature closure of thought, such that associations of the preparatory phase, which in normal thought are subordinated, are expressed as completed thought. Schizophrenics were further described by Schilder as lacking in the use of "correction processes" that integrate relevant elements and subordinate irrelevant associations, and as not attaining an "awareness of the truth or falsity of a completed thought."

Schilder's concepts of correction processes and truth awareness were related in this study to the concept of reality testing, and to Bartlett's concept (1958) of the ability to detect agreement and differences in evidence leading to the solution of a problem.

An experimental task, consisting of the recall of a short story over a succession of trials, was employed to investigate the hypothesis that in schizophrenics with thought disorganization, ability to detect points of agreement and difference in
evidence does not lead to correction of subsequent recalls of the story.

Subjects were 20 experimental and 20 control group Veterans Administration Hospital patients diagnosed as chronic schizophrenics, and 15 experimental and 15 control group hospital employees. Both patient groups and normal groups were matched on age, education level, vocabulary raw score, immediate recall score, and a measure of thought disorganization derived from the deviant responses in the pretest protocols. In addition, the two patient groups were matched on length of hospitalization. There were no significant differences on any of the variables on which these groups were matched. There were also no significant differences between the means of the two experimental groups and the two control groups on education level, vocabulary, and immediate recall. These groups, however, differed significantly on age and the thought disorganization score.

The experimental task was presented to the subjects by means of a tape recorder. Trials were discontinued after 22 correct components of the story were reproduced in correct order, or after 15 trials. Experimental group subjects received a playback of their reproductions on each trial in addition to the original story. For the control group subjects, a waiting period replaced the playback of recall.
The hypotheses were (1) that the difference between the normal experimental and control groups in the mean number of trials to reach successive criteria of recall would increase, reaching a significant difference, the experimental group having the lower means, and (2) that there would be no significant differences between the schizophrenic experimental and control groups in the mean number of trials to reach successive criteria of recall.

The results of the experiment supported the hypotheses. Because of differences between the groups in recall on the first trial and the failure of many of the schizophrenic subjects to learn the story to a level where their performances could be compared directly with those of the two normal groups, however, the results were interpreted with caution.

Two secondary analyses of the data were performed: (1) The combined patient and combined normal groups were compared on the mean number of substitute, added, and displaced components in their recalls. In general, this analysis supported the findings of Flavell (1956) and Lipton (1954) in the difference in the use of synonyms, number of added components, and number of displaced components found between normal and schizophrenic samples; (2) the two experimental and the two control groups were combined to investigate the effect of the independent variable (feedback of recall) on the mean number of substitute, added, and displaced components used by the subjects. No significant differences were found in this analysis.
BIBLIOGRAPHY


Cameron, N. Reasoning, regression, and communication in schizophrenics. Psychol. Monogr., 1938, 50, No. 1 (Whole No. 221).


Flavell, J. H. Abstract thinking and social behavior in schizophrenia. *J. abnorm. soc. Psychol.,* 1956, 52, 208-211.


Goldstein, K. The significance of psychological research in schizophrenia. *J. nerv. ment. Dis.,* 1943, 97, No. 3.


APPENDIX A

SCORING CATEGORIES FOR THE
DISORGANIZATION SCORE
I. PERSEVERATION

Responses which were a reiteration of a previous theme, or the repetition of a stimulus or response appearing earlier in the testing were classed in this category. Theoretically, these responses show a lack of shift from a previous set, or preoccupations and delusional ideation that take precedence over the appropriate response. Responses were scored for the following subcategories:

Stereotypy. A prolonged elaboration of related ideas to a stimulus, or a theme which is elaborated in subsequent responses without regard for the subsequent stimuli.

Perseveration. A response which is both (1) unrelated or distantly related to the stimulus and (2) identical with the response or stimulus in a previous part of the testing.

Clang. A response which is (1) related to the stimulus by alliterative or assonant characteristics and (2) unrelated or only distantly related to the stimulus in meaning.

Phrase completion. A response which added to the stimulus forms a cliche, or common phrase, and does not give a correct answer to the stimulus question.

II. METONYM

Responses were classed in this category when the phraseology of the response contained approximate, tangential, or distantly related words given in place of the more exact terms that would be judged as normal responses. The following subcategories were included under metonymic distortion:

Metonym. A response containing words approximately related to the stimulus concept.

Neologism. Words that are (1) corruptions of words that would apply to the stimulus question and (2) new words not found in the language.

Definition of a related word. A response on the vocabulary items that is a definition of a word that is related to the stimulus word (e.g., definition of "destruction" given to the stimulus word "obstruct").
III. ASYNDETIC

Responses were scored asyndetic when they reflected (1) a juxtaposition of ideas that were unrelated or not adequately linked together, (2) new ideas so expressed as to prove what preceded them, (3) ideas connected in a sentence that did not logically belong together, (4) associations from the current discussion (not the test stimuli) given as part of the response, or (5) changes in the subject at hand to divert the experimenter from the task.

IV. INTERPENETRATION

Responses were scored in this category when they reflected the interjection of dominant fantasies or personal preoccupations of the subject along with a logical response to the stimulus question. Responses were scored interpenetration, rather than asyndetic, when the meaningful relationship of the interjection to the stimulus was evident, or when the interjection was a comment not intended to be part of the answer. Interpenetration, rather than asyndetic, was scored when personal reference characterized the response.

V. FAULTY ABSTRACTION

This category was used only to score the proverbs. Responses that were abstractions (rather than concrete generalizations) were scored in this category when the abstraction ended at some point other than a general statement of the stimulus concept. The following subcategories were scored:

Overabstraction. An interpretation indicating that although the proverb was correctly understood by the subject, his attempt to put it in abstract terms led to a generalization beyond the meaning of the proverb (Benjamin, 1944).

Irrelevant abstraction. An interpretation abstracting from irrelevant or peripheral aspects of the stimulus proverb.

False desymbolization. An interpretation indicating that the symbols in the proverb have been translated, but incorrectly interpreted.
VI. LITERAL

Responses were scored in this category when there was a failure to abstract when the task required it. The category was used only to score the proverbs. The following subcategories were used:

**Literal or concrete.** An interpretation that (1) is a restatement of the literal meaning of a proverb, or (2) is a generalization from the literal meaning of a proverb.

**Semi-literal.** An interpretation of a proverb which meets both of the following criteria: (1) it is an abstraction to a specific example, and (2) it uses some of the wording of the proverb.
APPENDIX B

GUIDE FOR SCORING THE COMPONENTS OF THE EXPERIMENTAL STORY
<table>
<thead>
<tr>
<th>Key: X = correct component; S = accepted synonym; D = distantly related substitute; V = very distantly related substitute</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>X</strong></td>
</tr>
<tr>
<td>On the morning, in the morning</td>
</tr>
<tr>
<td>the morning</td>
</tr>
<tr>
<td>of December, on December, December</td>
</tr>
<tr>
<td>the 8th, 8th</td>
</tr>
<tr>
<td>the son, a son, 's son</td>
</tr>
<tr>
<td>of a wealthy, a wealthy</td>
</tr>
<tr>
<td>Hungarian, a Hungarian</td>
</tr>
<tr>
<td>landowner, a landowner</td>
</tr>
<tr>
<td>was swallowed, was devoured</td>
</tr>
<tr>
<td>by a shark, by a shark, a shark</td>
</tr>
<tr>
<td>He was the first, He was first, He was a first, was the/a first</td>
</tr>
<tr>
<td>mate, shipmate</td>
</tr>
<tr>
<td>on a Polish, of a Polish</td>
</tr>
<tr>
<td>X</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>freighter,</td>
</tr>
<tr>
<td>on a freighter</td>
</tr>
<tr>
<td>where, five</td>
</tr>
<tr>
<td>miles</td>
</tr>
<tr>
<td>west of</td>
</tr>
<tr>
<td>Borneo</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>during a dreadful,</td>
</tr>
<tr>
<td>in a dreadful</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>storm,</td>
</tr>
<tr>
<td>in a storm</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>he was washed</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>overboard</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>with the</td>
</tr>
<tr>
<td>blinding,</td>
</tr>
<tr>
<td>in the</td>
</tr>
<tr>
<td>blinding</td>
</tr>
<tr>
<td>rain</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>and tossing</td>
</tr>
</tbody>
</table>
waves
rescue
was impossible
and in front of
the eyes of
his terrified
shipmates
the unfortunate
man,
the man
was grabbed
on the leg,
by the leg
by a huge
shark
which had been
following,
that had been
following

seas (with tossing)
made rescue, made rescue work, saved (with could not), his rescue impossible, could not do in the eyes of his terrified mates
teaming mates, buddies, friends, companions

seaman, mate, son
grabbed him, was caught
by his leg
big, large
sharks
that had followed, which had followed, that/which trailed

waters, seas
anything about it
his rescue
impossible, unable to, could not be
in the eyes of his horrified mates

wind
<table>
<thead>
<tr>
<th>X</th>
<th>S</th>
<th>D</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>the steamer</td>
<td>the freighter, the ship</td>
<td>the boat, the liner</td>
<td>the streamer</td>
</tr>
<tr>
<td>He was dragged, was drugged</td>
<td>was pulled, was drug was drugged</td>
<td>was carried, was taken</td>
<td></td>
</tr>
<tr>
<td>to the bottom</td>
<td>on the bottom</td>
<td>under the water, across to the ocean floor</td>
<td></td>
</tr>
<tr>
<td>of the ocean</td>
<td>of the sea</td>
<td>of the lake</td>
<td></td>
</tr>
<tr>
<td>leaving behind him, leaving...behind him (following blood)</td>
<td>leaving behind, trailing behind, left behind him</td>
<td>flowing behind, left a, trailing a</td>
<td></td>
</tr>
<tr>
<td>a dark</td>
<td>black</td>
<td>large, huge, fresh</td>
<td></td>
</tr>
<tr>
<td>streak</td>
<td>stream, trail</td>
<td>pool, flowing</td>
<td></td>
</tr>
<tr>
<td>of blood</td>
<td>bloody</td>
<td>red</td>
<td></td>
</tr>
</tbody>
</table>
I, Dee Norman Lloyd, was born February 9, 1934, in Salt Lake City, Utah. I received my undergraduate training at the University of Utah, graduating with a Bachelor of Arts degree in psychology in 1956. The same year I entered graduate school at the University of Utah. In 1957, I was appointed to the Veterans Administration Psychology Training Program and served a part-time internship at the Fort Douglas Veterans Administration Hospital for one year. I received a Master of Arts degree in Counseling Psychology from the University of Utah in 1958. I was granted a Teaching Assistantship at The Ohio State University, which I held for two years while completing the requirements for a Doctor of Philosophy degree in the area of Counseling Psychology. From 1958 to the present, I have also continued internship work as a psychology trainee at the Veterans Administration Hospital, Chillicothe, Ohio.