Psychological Monographs: General and Applied

Combining the *Applied Psychology Monographs* and the *Archives of Psychology* with the *Psychological Monographs*

NORMAN L. MUNN, Editor
Department of Psychology, Bowdoin College
Brunswick, Maine

Manuscripts and correspondence on editorial matters should be sent to the Editor. *Psychological Monographs* publishes comprehensive experimental investigations and other psychological studies which do not lend themselves to adequate presentation as journal articles. Major space is given to the author's original contribution; introductory and bibliographic materials, as well as statistical tables and graphs, must be kept within reasonable bounds. Tables, graphs, and appendix materials which deal with detail not essential to adequate presentation of the findings may be made available through the American Documentation Institute—for details of this procedure, see the APA *Publication Manual*. Preparation of manuscripts for publication as monographs should follow the procedure given in the APA *Publication Manual*. Publication in *Psychological Monographs* is free of cost to the author, except in cases where early publication is requested or author's alterations are made in galley proofs.

Manuscripts must be accompanied by an abstract of 100–120 words typed on a separate sheet of paper. The abstract should conform to the style of *Psychological Abstracts*. Detailed instructions for preparation of the abstracts may be obtained from the Editor or from the APA Central Office.

HELEN ORR
Managing Editor

JOYCE J. CARTER
Editorial Assistant

Correspondence on business matters should be addressed to the American Psychological Association, Inc., 1333 Sixteenth St., N.W., Washington 6, D. C. Address changes must arrive by the tenth of the month to take effect the following month. Undelivered copies resulting from address changes will not be replaced; subscribers should notify the post office that they will guarantee third-class forwarding postage.

Vol. 76, No. 16 Whole No. 535, 1962

Psychological Monographs: General and Applied

INTELLECT AFTER LOBOTOMY IN SCHIZOPHRENIA: A FACTOR-ANALYTIC STUDY[1]

RICHARD DE MILLE
University of Southern California

In the early years of psychosurgery (Freeman & Watts, 1942), attempts were made to assess the intellectual consequences of brain modifications such as prefrontal lobotomy. Now that a large literature has accumulated (Klebanoff, 1945; Klebanoff, Singer, & Wilensky, 1954; Lewis, Landis, & King, 1956; Mettler, 1949; Pribram, 1960; Scherer, Klett, & Winne, 1957) it has become obvious that, whatever subtle intellectual changes occur after psychosurgery, tests of "general intelligence," as well as numerous specialized tests, have failed to analyze them and usually have failed even to detect them. Psychosurgery appears to be a landmark now passing into medical history, and yet the reluctance of investigators to relinquish the problem of measuring its effects has been expressed more than once.

Prefrontal lobotomy was chosen for study in this investigation not because of a need to assess psychosurgery as a procedure but mainly because of theoretical and methodological considerations. Twenty years ago the observation was made by Lashley (1941) that functions which might be lost as a result of localized brain injury did not seem to correspond to the behavior variables of classical psychology. He suggested that they might be found to correspond to unitary parameters of behavior developing out of factor analysis. The possibilities of a factor-analytic approach were explored at length by Halstead (1947), but no direct test of Lashley's suggestion was carried out prior to the present investigation.

The essential design of this study was proposed by Harvey (1950), who contended that a factorially complex test might be performed at the same level of competence by different people using different arrangements of the necessary factorially defined abilities. Failures to detect differences in abilities might be the result of not taking account of differences in the factor composition of a test for the different groups under study.

In the present investigation a distinction was made between *factor-defined* tests of in-

[1] While taking responsibility for shortcomings in the present report, the investigator wishes to express his indebtedness and gratitude to Herman Harvey, J. P. Guilford, and P. R. Merrifield for meticulous guidance and generous consultation throughout the dissertation project (de Mille, 1961) on which this report is based.

Sincere thanks are due the managers, directors, and staffs of the Veterans Administration Center, Los Angeles, and of Brentwood, Palo Alto, and Sepulveda Veterans Administration Hospitals; in particular, F. Harold Giedt, Thomas W. Kennelly, Joseph Rubinstein, and John R. Schlosser; and Frank J. Kirkner, Long Beach Veterans Administration Hospital. Their cooperation in the job of finding, selecting, and testing subjects was indispensable. Special thanks are due Harry M. Grayson, Brentwood Veterans Administration Hospital, and Paul McReynolds, Palo Alto Veterans Administration Hospital, who greatly facilitated the two major testing programs.

The investigator is grateful to Leon Epstein of the California Department of Mental Hygiene for his encouragement at the outset of the study, and to the superintendents and staffs of Agnews, Atascadero, Camarillo, DeWitt, Mendocino, Metropolitan, Modesto, Napa, Patton, and Stockton State Hospitals for their cooperation and hospitality during the data gathering period. In particular, thanks are due to Trent E. Bessent, William A. Cook, Jr., Sarah Counts, Hershel Fogelson, Walter Freeman, Gerald Green, Kristian Johnsen, Bernard Meer, Margaret Noszlopi, Gordon L. Riley, Frances Sheridan, Benjamin Siegel, Wilson M. Van Dusen, and Seymour L. Zelen.

The Project for Studies of Aptitudes of High-Level Personnel, University of Southern California, gave permission to reproduce a number of experimental tests. Correlations and analytical factor rotations were computed by the Western Data Processing Center, University of California at Los Angeles. The investigation was carried out during tenure of a predoctoral fellowship from the National Institute of Mental Health, United States Public Health Service.

tellect, such as those developed by Thurstone (Thurstone & Thurstone, 1947) or Guilford (1959, 1961), and *task-defined* tests, such as the subtests of the Wechsler-Bellevue scale (Wechsler, 1944, 1947). A comparison of these two kinds of tests, or approaches to the measurement of intellect, was a central purpose of the study.

In recent years animal work has taken great strides toward the specification of unitary functional brain systems (Pribram, 1960). If research in physiological psychology is to continue to be fruitful, specification of psychological variables must keep up with specification of physiological variables. The unitary behavioral parameters developed by factor analysis appear indispensable to the maintenance of this parity. A goal of the present study was to facilitate increased application of multifactorial methods to problems in physiological, abnormal, and clinical psychology. It was thought that if progress could be made, at this late date, in the measurement of effects of psychosurgery, investigators would be encouraged to apply multifactorial methods to a number of more current problems dealing with intellectual, physiological, psychopharmacological, and psychopathological variables.

PROBLEM

The study was designed to demonstrate factor-analytic measurement of behaviors in a physiologico-behavioral problem, to compare the factorial composition of tests for two different groups of subjects, and to assess the relative effectiveness of task-defined and factor-defined tests in differentiating the groups under study.

The independent physiological variable was the presence or absence of a prefrontal lobotomy. The dependent behavioral variables were the scores on tests of intellectual ability. A structure of intellectual factors derived from a group of lobotomized subjects was compared with a structure based on the same variables derived from a matched group of unlobotomized subjects.

The task-defined test variables included the subtest and IQ scores of the Wechsler-Bellevue Intelligence Scale, Form I (Wechsler, 1944, 1947). The factor-defined test variables were 16 scores from a battery of relatively univocal tests relating to Guilford's structure-of-intellect model (Guilford & Merrifield, 1960). Group mean score comparisons were made and statistically evaluated, and the comparative ability of the two kinds of tests to differentiate lobotomized from unlobotomized subjects was observed.

The majority of investigators who have studied intellectual consequences of lobotomy in schizophrenia have reported neither loss of ability nor inferiority of lobotomized to unlobotomized patients. Some have depreciated the role of the frontal lobes in intellect. Some have suggested that there may be gains after lobotomy or superiority of lobotomized to unlobotomized patients. In keeping with this background, the following hypothesis was tested: the unlobotomized control group have mean scores on intellectual indicators that are not different from or are lower than those of the lobotomized experimental group.

Three stages were observed in testing the hypothesis. In the first stage, the groups were matched on nontest variables. In the second stage, matched subgroups were selected so that two nontest sources of heterogeneity were removed. In the last stage, the previous nontest matching was maintained, while experimental and control subgroups were selected to match on a test variable, the Wechsler-Bellevue IQ.

The purpose of the third stage was to remove intergroup differences on task-defined test variables and observe whether intergroup differences on factor-defined test variables would remain.

PROCEDURE

Description of the Tests

The 13 tests of the main battery are listed below. Three of them provide two scores, to make 16 test variables in all. Unpublished tests bear their code numbers from the Studies of Aptitudes of High-Level Personnel (Guilford & Merrifield, 1960).

For each test, a brief description of the task and scoring procedure, a sample item (for some tests), the number of parts and items per part, the working time, and the structure-of-intellect nomenclature of the expected leading factor loading (Guilford & Merrifield, 1960) are given. When two working times appear, the one in brackets is the original time limit, before modification in the present study.

All tests are pencil-and-paper tests except the last two, in which the subject responds orally. In previous studies, Ship Destination and Social Situations were used with machine-scoring answer sheets. To decrease their difficulty and factorial complexity for abnormal subjects, a different format was adopted for them and, similarly, for the Letter Series test.

1. *Numerical Operations*—Guilford-Zimmerman Aptitude Survey, Part III, Form A (Guilford & Zimmerman, 1947). Choose one of six numbers that is the correct sum, difference, or product for a given numerical problem.

One part, 132 items, 8 minutes working time.

Score: Number right plus 1/6 the number of omitted items.

Factor: NSI, Numerical Facility.

2. *Verbal Comprehension*—Guilford-Zimmerman Aptitude Survey, Part I, Form A (Guilford & Zimmerman, 1947). Select one of five words that is similar in meaning to a given word.

One part, 72 items, [25] 12 minutes working time.

Score: Number right plus 1/5 the number of omitted items.

Factor: CMU, Verbal Comprehension.

3. *Ship Destination Test* (Christensen & Guilford, 1955). Find the distance from ship to port, considering the influence of several variables, such as wind velocity and direction. Circle 1, 2, 3, 4, or 5, as the correct answer, on the answer sheet.

One part, 48 items, 15 minutes working time.

Score: Number right plus 1/5 the number of omitted items.

Factor: CMS, General Reasoning.

4-5. *Letter Series*—An adaptation of Thurstone's PMA Reasoning subtest (Thurstone & Thurstone, 1947). Circle one of six letters as the next one to follow a given series.

Sample item: a b a b a b a b   a b c d e f

Answer: a

One part, 30 items, [6] 12 minutes working time.

Score: Number right plus 1/6 the number of omitted items. Variable 4, odd items; Variable 5, even items.

Factor: CSS, Symbolic Patterns.

6. *Unusual Details*—EMS 01B. Write down two incongruities appearing in sketches of common scenes.

Sample answers: Lamp cannot give light when cord is not plugged in.

Bottom half of clock dial has numbers reversed.

Two parts, 16 items per part, [4] 8 minutes per part.

Score: Number right.

Factor: EMS, Experiential Evaluation.

7. *Social Situations*—EPO3A. For a given situation, select the action leading to the most desirable consequences. Circle the right letter: A, B, C, or D.

Sample item:

You are on a weekend trip with a group of friends. Most of them would prefer spending the day hunting, but you would prefer to go fishing. You should:

A. Go hunting with them.
B. Tell them to go hunting while you go fishing.
C. Try to convince them that they will have a better time fishing.
D. Offer to toss a coin to decide whether the whole group goes hunting or fishing.

Answer: A

Two parts, 15 items per part, 5 minutes per part.

Score: Number right plus 1/4 the number of omitted items.

Factor: EMS, Experiential Evaluation.

8-9. *Seeing Problems*—CSO6A. Write as many as five problems arising from the presence of a given object.

Sample object: Candle.

Answers:

1. How to light it.
2. Keeping it from falling over.
3. Keeping it from flickering.
4. How long will it burn.
5. What to do with the drippings.

Two parts, 6 items per part, [4] 8 minutes per part.

Score: Number of appropriate problems listed. Variable 8, Part I; Variable 9, Part II.

Factor: EMI, Sensitivity to Problems

10-11. *Ideational Fluency* (Christensen & Guilford, 1957). List objects that belong to specified classes.

Sample item: Name FLUIDS that will BURN.

Answers: gasoline kerosene hydrogen alcohol

Four parts, 1 item per part, 3 minutes per part.

Score: Number of acceptable responses. Variable 10, Parts I and II; Variable 11, Parts III and IV.

Factor: DMU, Ideational Fluency.

12. *Brick Uses (shifts)*—CFO4A. List different uses for a brick.

One part, one item, 10 minutes working time.

Score: Number of times the subject shifts from one kind of use to another.

Factor: DMC, Semantic Spontaneous Flexibility.

13. *Alternative Uses*—Experimental version of Alternate Uses (Guilford, Christensen, Wilson, & Merrifield, 1959). List different unusual uses for common objects.

Sample item:

Given a newspaper (used for reading), you might think of the following other uses for a newspaper:

1. To start a fire
2. To wrap garbage in
3. To swat flies
4. Stuffing to pack boxes
5. To line drawers or shelves
6. To make up a kidnap note

Two parts, 4 items per part, 6 minutes per part.

Score: Number of different acceptable uses.

Factor: DMC, Semantic Spontaneous Flexibility.

14. *Digit Symbol*—Wechsler-Bellevue subtest (Wechsler, 1947). Write symbols corresponding to digits, according to a key.

One part, 67 items, 90 seconds working time.

Score: Number right plus ½ the number reversed.

Factor: NSI, Numerical Facility (Davis, 1956).

15. *Vocabulary*—Wechsler-Bellevue subtest (Wechsler, 1947). Orally define a list of words orally presented.

One part, 42 items.

Score: Sum of unit and half credits, according to standards in the manual (Wechsler, 1944).

Factor: CMU, Verbal Comprehension (Davis, 1956).

16. *Arithmetic*—Wechsler-Bellevue subtest (Wechsler, 1947). Without a pencil, solve orally or visually presented problems of the type: How many hours will it take a man to walk 32 miles at the rate of 4 miles an hour?

One part, 10 items, 15–120 seconds per item.

Score: Number right plus time credits on Items 9 and 10, according to the manual (Wechsler, 1944).

Factor: CMS, General Reasoning (Davis, 1956).

Selection of the above tests was made for a number of reasons. It was thought desirable to include some much-studied factors, such as numerical facility, verbal comprehension, and general reasoning. A second desideratum was to have at least one factor from each of the four thinking divisions of the structure-of-intellect model: cognition (CMU, CMS, CSS), divergent production (DMU, DMC), convergent production (NSI), and evaluation (EMI, EMS). There was particular interest in productivity, since it has been suggested that productiveness and flexibility may be dependent upon the frontal lobes. A third requirement was to favor the more univocal tests, so that the interpretation of the resulting factors might be less ambiguous. A fourth requirement was to have tests of suitable length and difficulty for research with schizophrenic subjects.

Economy of testing time was achieved by making three Wechsler-Bellevue subtests do double duty, serving both in the factorial and Wechsler-Bellevue portions of the study. The Wechsler-Bellevue was preferred to the newer Wechsler Adult Intelligence Scale, because it had been more studied, had been adequately factor analyzed (Davis, 1956; Wechsler, 1958), was a source of early scores on the subjects for possible supplementary study, and was not so likely to have been recently administered to the subjects.

The scoring of tests in separate parts and the intercorrelation of those parts results either in the appearance of doublet factors or in the inflation of factor loadings by the confounding of common and specific factor variance. Some risk of restriction of interpretability of the factors was thought to be justified by the larger number of factors thus introduced. The same argument applies to the limitation of marker variables to two per factor, sacrificing the overdetermination of factors that is desirable.

Subjects

One hundred fifty lobotomized and 150 unlobotomized schizophrenic patients were selected from the rolls of 11 hospitals in California. More than half were veterans from Brentwood, Palo Alto, and Sepulveda Veterans Administration Hospitals. The rest were from Agnews, Atascadero, Camarillo, Mendocino, Metropolitan, Napa, Patton, and Stockton State Hospitals.

Experimental group. The 150 lobotomized subjects were chosen from a pool of over 800 available patients, to satisfy seven criteria. All had been given a bilateral prefrontal lobotomy, and were not known to have suffered any other brain damage. All had a diagnosis, before lobotomy, of paranoid, hebephrenic, catatonic, simple, undifferentiated, or mixed schizophrenia, with no other diagnosis to complicate the picture either before or after operation, except for a diagnosis, in 33 cases, of epilepsy following lobotomy. No subject was over 60 years of age. All spoke English as their most used language and had finished the eighth grade in school.

The postoperative interval, or remoteness of operation, varied from 5 to 15 years, with a mean of 8.7 years and a standard deviation of 2.2 years.

Control group. The 150 unlobotomized subjects were chosen to match the experimental group in age, highest school grade completed, duration of illness, sex, race, diagnostic subgroup, veteran status, whether or not on a tranquilizer at the time of testing, and hospital source.

Comparability of the groups and subgroups. Group matching on the continuous variables, age, schooling, and duration of illness is shown in Table 1. The differences in group means on all three variables are statistically nonsignificant and are very small. The fact that the control subjects are more variable in duration of illness than the lobotomized subjects is attributable to the fact that all lobotomies were performed within an 11-year period, tending to constrict the variability of the lobotomized subjects on chronological variables. This effect is increased in the veteran sample shown in Table 2. The interpretation of Tables 1 and 2 is that there are no meaningful differences in age, schooling, or duration of illness, either between the main groups or between subgroups.

TABLE 1

GROUP MATCHING ON CONTINUOUS VARIABLES

Variable	Experimental ($N = 150$)		Control ($N = 150$)		Ratios	
	M	SD	M	SD	F	t
Age	41.86	6.48	42.09	6.75	1.09	0.30
Schooling [a]	11.82	1.70	11.71	1.93	1.29	−0.52
Illness [b]	16.59	4.23	16.63	5.40	1.62**	0.07

[a] Highest grade completed.
[b] Duration of illness.
** Significant at the .01 level, two-tailed test.

Group matching on category variables is shown in Table 3. The first two columns compare the main groups; the last four columns compare subgroups. Matching in all categories is very close, with two minor exceptions. More lobotomized subjects are diagnosed as catatonic; more controls are diagnosed as undifferentiated. If any bias is introduced by this discrepancy, it probably tends to reduce test score differences between lobotomized and control subjects, for in the last two columns, where group mean Wechsler-Bellevue IQs have been equated, the discrepancy is slightly increased.

Matching for race is not shown in the tables.

TABLE 2

SUBGROUP MATCHING ON CONTINUOUS VARIABLES

Variable	Experimental ($N = 69$)		Control ($N = 72$)		Ratios	
	M	SD	M	SD	F	t
Age	40.03	5.35	40.19	6.53	1.49*	0.16
Schooling [a]	11.65	1.73	11.51	1.93	1.25	−0.05
Illness [b]	16.17	3.62	15.97	5.06	1.95**	−0.27
	($N = 50$)		($N = 50$)			
Age	39.90	4.90	39.88	6.53	1.78*	—
Schooling [a]	11.68	1.78	11.50	2.08	1.36	−0.46
Illness [b]	16.08	3.62	15.82	5.07	1.96*	−0.29
W-B FS IQ [c]	97.98	12.70	98.28	13.53	1.13	—

Note.—All subjects are white veterans.
[a] Highest grade completed.
[b] Duration of illness.
[c] Wechsler-Bellevue Scale, Form I, Full Scale IQ.
* Significant at the .05 level, two-tailed test.
** Significant at the .01 level, two-tailed test.

TABLE 3

CATEGORICAL GROUP AND SUBGROUP MATCHING [a]

Category	150 E [b]	150 C	69 E	72 C	50 E	50 C
Sex male	.53	.53	.87	.88	.82	.82
Paranoid	.47	.47	.52	.53	.54	.54
Hebephrenic	.17	.15	.12	.11	.12	.12
Catatonic	.22	.13	.19	.13	.26	.14
Simple	.01	.03	.00	.03	.00	.00
Undifferentiated	.07	.16	.13	.18	.06	.16
Mixed	.05	.05	.04	.03	.02	.04
Veteran	.54	.54	1.00	1.00	1.00	1.00
Tranquilizer	.74	.77	.88	.90	.86	.86
Hospital [c]	.94	—	.97	—	.84	

[a] Proportions of groups falling into criterion or descriptive categories.
[b] E, Experimentals; C, Controls. Groups 69 E, 72 C, 50 E, and 50 C are white veterans.
[c] Proportion of experimental group that is matched by having controls from the same hospital.

Included in each group of 150 subjects were 2 Negroes and 3 Mexicans. All others were Caucasians, except for 1 American Indian in the experimental group. All subjects in the subgroups were Caucasians.

All feasible steps were taken to increase the comparability of the lobotomized and control groups. Ideally, the controls would have been schizophrenic patients who had been chosen for lobotomy and then had been selected by the investigator to act as controls instead of having the operation. Such a design, of course, could never be carried out, due to extra-experimental considerations. An approximation to this ideal was to find subjects who had been recommended for a lobotomy at a medical staff conference and then had not received the operation because of some event extrinsic to the individual diagnosis and plan of treatment. Such events were the refusal of relatives to grant permission for the operation, transfer of the patient, or a change of policy with respect to lobotomy. At every hospital where controls were obtained, a search was made for such subjects; 14 were found.

This control group category, Lobotomy Recommended, was correlated with each of 16-test variables to see whether there was a tendency for the recommended subjects to do less well than the other controls. The correlation coefficients appear in Row k of Table B.[2] None is large enough to be statistically significant; 11 are negative and 5 are positive, while the chance expectation is 8 and 8. The interpretation of

[2] Table A—Correlation Matrix, Experimental Group; Table B—Correlation Matrix, Control Group; Table C—Centroid Matrix, Experimental Group; and Table D—Centroid Matrix, Control Group, have been deposited with the American Documentation Institute. Order Document No. 7260 from ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress; Washington 25, D.C., remitting in advance $1.25 for microfilm or $1.25 for photocopies. Make checks payable to: Chief, Photoduplication Service, Library of Congress.

these results is that the test performance of the control group would not have been of lower quality if all controls had been previously recommended for lobotomy. Supporting this interpretation is the fact that in the veteran sample, where 6 out of 72 subjects were recommended for lobtomy, the mean Wechsler-Bellevue Full Scale IQ of the 6 was 3.5 IQ points higher than that of the other 66 subjects, a nonsignificant difference.

Not shown in any table is the fact that an effort was made to obtain control subjects from the same wards as lobotomized subjects, and in the same proportions. To a large extent, this effort was successful. Bias in terms of interward variables was thus minimized, and it is felt that no advantage was given to the control group as a result of interward variables.

Another aspect of the attempt to avoid a bias in favor of the control group was the inclusion of 9 lobotomized veterans who were not, like all controls, residents in a hospital at the time of testing. Two were residents in the Domiciliary, 5 were in Family Care, 1 was on Trial Visit, and 1 was discharged. The discharged patient was still mildly delusional and hallucinated. Two of the subjects in Family Care were eliminated from the subgroup of 50 lobotomized subjects.

The comparability of experimentals and controls was increased in the intermediate subsample of 69 experimentals and 72 controls by restricting the subsample to white veterans. The elimination of nonwhites had the actual effect of reducing differences in mean test scores between experimentals and controls. Coincidentally, a small number of subjects were lost because of incomplete Wechsler-Bellevue data. No systematic trends in the reasons for the loss appeared; and this restriction likewise had the actual effect of reducing mean score differences between groups.

In the small subsample, intergroup comparability was increased further by equating subgroup means and variabilities on the Wechsler-Bellevue, Form I, Full Scale IQ for 50 experimentals and 50 controls drawn from the intermediate subsample. As can be seen in Tables 1, 2, and 3, each step of this progressive matching procedure was achieved without sacrificing any of the previous matching.

An attempt was made to obtain premorbid General Classification Test scores from the service records of all of the veteran subjects. Scores were actually available for 34 experimentals and 27 controls, Army, Navy, and Marine veterans of World War II and the Korean War. These two special subgroups were as fully and closely matched as the other pairs of groups, and were also matched as to their branch of service. The mean premorbid intelligence score, converted to a Wechsler-Bellevue IQ score, was 102.1 (*SD* 11.8) for the experimentals, 103.0 (*SD* 11.4) for the controls, a result consistent with the general assertion of intergroup comparability.

No effort was made to measure motivation in the two groups apart from their test performances where it is inseparable from ability. The investigator received no reports of differential motivation from t[he] examiners. Mean scores on the Digit Symbol test, [a] speeded test, were nearly identical (see Table 6) f[or] the two groups, which argues against a difference [in] motivation (Guilford, 1959).

No attempt was made to match the groups in term[s] of their histories of shock therapies. Notation [in] clinical records was neither consistent enough n[or] complete enough to justify assigning a numerica[l] value for each subject. It was assumed that tw[o] groups of patients of such comparability and chronic[-]ity would have had the benefit of the various form[s] of shock therapy in about equal measure. A possib[le] difference of policy on shock therapy from hospita[l] to hospital was controlled by a high degree of hos[-]pital matching (see Table 3). No subject had ha[d] shock treatment for 6 months preceding testing. I[n] a large majority of cases, the subjects had not ha[d] shock treatment for several years.

Pilot Project

Before undertaking to test 300 subjects, it wa[s] thought advisable to find out just how difficult th[e] tests were for chronic schizophrenic patients an[d] whether any alterations in the time limits would b[e] desirable. Accordingly, the main battery, plus th[e] Logical Reasoning test (Hertzka & Guilford, 195[5]) was administered to 27 patients. Two were unlobot[-]omized; 9 had prefrontal, and 16 had transorbita[l] lobotomies. The 11 patients in the first two class[es] were included in the main study.

Working times of all tests, except the three Wech[s-]ler-Bellevue subtests, were modified. Scores were re[-]corded for the original time limits and the modifie[d] time limits. A frequency distribution was plotted fo[r] each test variable at each time limit.

It was found that extending the working time mad[e] very little difference in the shape of the distributio[n] or in the number of zero scores for Numerical Opera[-]tions, Ship Destination, Social Situations, Ideationa[l] Fluency, Brick Uses (shifts), or Alternative Use[s]. Reducing the working time for Verbal Comprehen[-]sion made very little difference, and so the shorte[r] time was adopted.

On the other hand, doubling the time limits o[n] Unusual Details, Seeing Problems, and Letter Seri[es] improved the distributions and reduced the numbe[r] of zero scores, and so the longer times were adopte[d]. Logical Reasoning was too difficult, even for e[x-]perimental use, and was dropped.

Administration and Scoring of Tests

Nonveterans were tested between November 195[9] and October 1960; veterans, between March 196[0] and January 1961. All subjects were tested individ[-]ually in two or more sessions. Each subject had abou[t] 3 hours of testing, or 4 hours if the entire Wechsler[-]Bellevue was given.

Twenty-one examiners, including the investigator [a]dministered the tests.[8] Eight were clinical psychol[o]gy trainees; five were hospital staff psychologists; and three were school psychologists. Of the remaining five, two had had graduate training in testing, two were regularly engaged in administering tests to hospitalized patients, and the fifth had had both graduate training and hospital testing experience.

The geographical scatter of the subjects and the extended testing schedules made it difficult to match the groups on examiners. However, 58% of the experimental subjects had controls who were tested by the same examiner. The experience of the examiners with the battery was quite comparable for the two groups. Serial administration effects were controlled by the application of six counterbalanced test orders distributed about equally among the subjects. Though there was a tendency for the lobotomized subjects to be tested earlier than the controls, there was a great deal of overlap.

The Wechsler-Bellevue subtests were administered according to the manual (Wechsler, 1944). When the balance of the Wechsler was given, it followed the main battery. The factor-analytic tests were administered with the help of a special manual prepared for the project, providing a procedure for adapting the standard group instructions to individual administration. The key passage was as follows:

> In administering these tests individually, the examiner should use the instructions judiciously, as a guide for himself and as information for the subject. Help the subject to understand the task as fully as possible. Do not read the instructions verbatim if the subject has any difficulty following them, but explain the tasks making use of the pertinent instructional information. It is not necessary to tell the subject how long the test is or how many parts it has. It is necessary to observe him closely to make sure that he is complying with the instructions. Give help, if needed, during the instruction period. During the test proper, do not give help, but encourage the subject to keep going if his interest flags.

Very few anomalies in test administration came to the attention of the investigator who scored all of the tests.

The test batteries were scored as received, which led to a great deal of overlap between the experimental and control groups. No blind scoring procedure was used, since most of the scoring was completely objective. Four tests, Seeing Problems, Ideational Fluency, Brick Uses (shifts), and Alternative Uses, required the scorer to evaluate the responses to a degree that might have permitted a scoring bias to affect the results. However, it is felt that the frequent alternation between experimental and control subjects was a safeguard against scoring bias.

[8] Subjects not tested by the investigator were tested by Alice Beach, Stephen Carraway, Harry Carritte, Beverly Collins, Thomas Coke, Raymond Conatser, Chester Cooley, Sheila Farley, Joseph Fischer, Edward Gould, Jacqueline Griffin, Barbara Griswold, Millard Madsen, Ira Nathanson, Margaret Noszlopi, Alexander Quenk, Rollin Rose, H. William Safford, Frederick Stoller, and Ruby Thompson.

Treatment of Data

At every stage of the treatment of data, and of the factor analyses, identical or logically equivalent mathematical procedures were applied in the two groups in order to minimize spurious intergroup differences.

Continuous variable data. Distributions of the 16 test scores[4] and of age, schooling, duration of illness, and remoteness of operation were plotted graphically and inspected for unimodality, symmetry, and amount of kurtosis. Arithmetic, Digit Symbol, and Ideational Fluency were found to be sufficiently isokurtic to justify the use of raw scores in the computation of Pearson product-moment correlations. Evaluations of the other variables were as follows: *Bimodal*—remoteness of operation; *Leptokurtic*—schooling (the mode was 12 years); *Mild positive skew*—age, duration of illness, Unusual Details, and Social Situations; *Moderate positive skew*—Numerical Operations, Verbal Comprehension, Seeing Problems I; *Severe positive skew*—Ship Destination, Letter Series, Seeing Problems II, Brick Uses (shifts), and Alternative Uses; and *Mild negative skew*—Vocabulary.

Variables that were bimodal, leptokurtic, or mildly or moderately skewed were normalized graphically (Guilford, 1956). Severely skewed variables were dichotomized as near as possible to the median (Guilford, 1956), subject to the provision that a zero score would not be included in the upper category. This provision resulted in some nonmedian splits, the most extreme of which was a 61–39% split.

Category data. Sex, veteran status, paranoid, hebephrenic, tranquilizer status, history of seizures, and recommendation for lobotomy were category variables for correlation. The other four diagnostic subcategories were not correlated because of insufficient numbers or unequal proportions in the two groups (cf. Table 3).

Intercorrelations. Pearson product-moment correlations were computed for all pairs of variables. Correlations involving dichotomized continuous variables were corrected for coarse grouping (Guilford, 1956). In three cases, the limitations for the application of the correction were exceeded, and tetrachoric correlations were computed for both groups and corrected for nonmedian dichotomization.

[4] Means and standard deviations of the 16 main-battery test variables, for both experimental and control groups, are shown in Table 6. Tables 4 and 5, the rotated factor matrices, give the communalities (h^2) for these variables, which may serve as lower-bound estimates of reliability. In Table 4, the obtained h^2 for Variable 4, Letter Series—odd, appears to be an overestimate due to the confounding of common and specific factor variances; the value from Table 5, 0.97, may be used for both groups as an estimate of the reliability of Variable 4.

In the original correlation matrices, an additional variable was included, Brick Uses (fluency), which is not independent of Brick Uses (shifts), and which was dropped because it seemed unlikely to contribute to the clarity of the eventual solution.

The final matrices of intercorrelations are presented in Tables A and B (see Footnote 2).

Statistics

At the beginning of this study, it was hoped that a suitable test for evaluating the statistical significance of differences between two factor structures would be found. Recent publications suggest that, though progress is being made, no test fulfilling the needs of the present study is yet available. When such a test arrives, it can be applied to the present factor solutions.

In the nonfactorial portion of the study, it was necessary to evaluate intergroup mean-score differences on tests and continuous matching variables. The variance ratio was evaluated in every case with a two-tailed F test, since no prediction was made that one group would be more variable than the other. Group mean scores on continuous matching variables are so close that the t ratios are obviously nonsignificant. On the test variables, a one-tailed t test was used, since the direction of the difference was specified in advance. In instances where the F ratio is significant, the t ratio was computed with an estimate of population variance made from both groups.

Extraction of Factors

Using Thurstone's complete centroid method (Thurstone, 1947), 9 factors were extracted from the 2 correlation matrices of 16 test variables. The distributions of residual coefficients after the extraction of the ninth factor were leptokurtic about zero and contained no coefficient as great as .06 in absolute value. The centroid matrices are presented in Tables C and D (see Footnote 2).

Rotation of Axes

Graphic orthogonal rotations were used (Zimmerman, 1946). Primary emphasis was laid on the criteria of simple structure and positive manifold (Thurstone, 1947). In the final rotations, the criterion of psychological meaning was imposed to effect minor improvements in the structures. The experimental centroid matrix was rotated first, and an attempt was made to make the control solution conform to the experimental solution. This attempt was only partially successful; substantial unpredicted factor loadings appeared in the control solution, and attempts to rotate them out of existence grossly violated the criterion of simple structure. Confirmation was sought by applying another rotation method.

Analytical rotations were computed, to the varimax criterion (Kaiser, 1958). The analytical and graphic solutions of the experimental rotation problem were very similar, calling for the same interpretation, though the graphic solution more nearly approached simple structure. The control problem, however, had markedly different solutions by the two methods of rotation. The analytical solution was further from simple structure, and it exaggerated and emphasized the unpredicted factor loadings found by the graphic method to the point of producing some quite different factors. It was concluded that the analytical solutions were confirmatory of, though less easily interpreted than, the graphic solutions. Particularly, the suggestion that the two factor structures were not essentially identical seemed confirmed by the analytical solutions.

The possibility of computing oblique analytical rotations was considered and rejected. It was felt that oblique rotations would tend to maximize the differences between the two groups. In the absence of statistical evaluation, it was thought that the more conservative orthogonal treatment would better support inferences about differences in factor structure.

In the experimental solution, 68 graphic rotations were made; in the control solution, 93. The rotated factor matrices are presented in Tables 4 and 5.

RESULTS

Interpretation of the Factors

For each factor, the test variables having significant loadings of .30 or more (Merrifield, Guilford, Christensen, & Frick, 1960) are listed in order of magnitude. The notations following the factor loading indicate loadings of .30 or more on other factors when there are any. The experimental group is listed first for each factor, as is indicated vertically at the left margin.

Factor A—NSI—Numerical Facility

E	14. Digit Symbol	.54	CSS .36, DMU .3
x	1. Numerical		
p	Operations	.45	CMS .35
C	14. Digit Symbol	.63	
o	1. Numerical		
n	Operations	.61	CMS .41, CSS .3
	6. Unusual		
	Details	.38	EMS .47, DMC .4

It has been suggested (Guilford & Merrifield, 1960) that Numerical Facility may be essentially a memory ability, involving long-term memory for symbolic implications rather than the convergent production of symbolic implications. The Digit Symbol test appears to require short-term memory, and Wechsler (1944) points out a marked correlation between Digit Symbol and Memory Span fo

TABLE 4

ROTATED FACTOR MATRIX, EXPERIMENTAL GROUP

Variable	A	B	C	D	E	F	G	H	R	h^2
1. Numerical Operations	45	04	35	27	00	18	16	10	23	52
2. Verbal Comprehension	11	74	09	24	15	06	16	09	27	76
3. Ship Destination	08	19	56	41	08	15	47	21	07	82
4. Letter Series—odd	—08	10	26	81	19	02	27	29	29	102
5. Letter Series—even	00	10	09	84	10	00	29	19	14	87
6. Unusual Details	25	30	07	28	21	15	41	25	29	62
7. Social Situations	22	38	12	38	23	17	43	02	04	62
8. Seeing Problems—I	00	11	15	07	83	17	14	13	28	87
9. Seeing Problems—II	16	04	09	06	86	03	00	09	17	82
10. Ideational Fluency—I–II	08	20	06	28	10	71	13	19	12	71
11. Ideational Fluency—III–IV	15	21	18	08	07	62	15	34	29	72
12. Brick Uses (shifts)	14	19	—01	01	28	15	—08	74	16	74
13. Alternative Uses	—02	22	15	01	26	08	21	64	27	67
14. Digit Symbol	54	01	06	36	16	34	10	10	—04	59
15. Vocabulary	10	79	24	09	07	06	22	17	20	82
16. Arithmetic	04	22	68	29	01	—01	04	15	14	64

Note.—Decimal points have been omitted. Key: A—NSI, Numerical Facility; B—CMU, Verbal Comprehension; C—CMS, General Reasoning; D—CSS, Symbolic Patterns; E—EMI, Sensitivity to Problems; F—DMU, Ideational Fluency; G—EMS, Experiential Evaluation; H—DMC, Semantic Spontaneous Flexibility; R—Residual.

TABLE 5

ROTATED FACTOR MATRIX, CONTROL GROUP

Variable	A	B	C	D	E	F	G	H	R	h^2
1. Numerical Operations	61	10	41	33	10	19	26	09	17	81
2. Verbal Comprehension	21	72	—03	33	16	14	10	25	01	79
3. Ship Destination	—05	20	51	43	10	13	15	50	11	80
4. Letter Series—odd	14	35	25	79	26	12	16	00	—17	97
5. Letter Series—even	06	13	26	85	25	16	23	09	13	98
6. Unusual Details	38	29	12	20	18	09	47	42	22	77
7. Social Situations	27	28	—01	13	23	22	37	20	14	47
8. Seeing Problems—I	25	14	15	02	73	20	09	37	14	84
9. Seeing Problems—II	18	02	13	19	66	02	15	63	—21	99
10. Ideational Fluency—I–II	20	22	14	07	05	69	05	20	26	70
11. Ideational Fluency—III–IV	04	25	10	10	23	72	27	08	19	77
12. Brick Uses (shifts)	08	09	22	01	09	54	10	58	—19	75
13. Alternative Uses	10	08	07	25	24	16	15	61	29	68
14. Digit Symbol	63	10	05	24	06	25	15	14	16	66
15. Vocabulary	10	76	14	18	12	07	23	29	04	80
16. Arithmetic	22	36	60	09	25	10	09	03	24	69

Note.—Decimal points have been omitted. Key: A—NSI, Numerical Facility; B—CMU, Verbal Comprehension; C—CMS, General Reasoning; D—CSS, Symbolic Patterns; E—EMI, Sensitivity to Problems; F—DMU, Ideational Fluency; G—EMS, Experiential Evaluation; H—DMC Semantic Spontaneous Flexibility; R—Residual.

Digits. The appearance of Unusual Details on this factor, then, might reflect some memory content in Unusual Details for the control subjects.

Factor B—CMU—Verbal Comprehension

Exp	15. Vocabulary	.79	
	2. Verbal Comprehension	.74	
	7. Social Situations	.38	EMS .43, CSS .38
	6. Unusual Details	.30	EMS .41
Con	15. Vocabulary	.76	
	2. Verbal Comprehension	.72	CSS .33
	16. Arithmetic	.36	CMS .60
	4. Letter Series—odd	.35	CSS .79

It is not surprising that cognition of semantic units, or word meaning, should contribute to performance on Social Situations in which much reading is required, or Unusual Details in which 8 of the 32 items contain verbal cues. The replacement of these two by Arithmetic and Letter Series in the control group, and the separation of the odd and even items of Letter Series, are not readily explainable.

Factor C—CMS—General Reasoning

Exp	16. Arithmetic	.68	
	3. Ship Destination	.56	EMS .47, CSS .41
	1. Numerical Operations	.35	NSI .45
Con	16. Arithmetic	.60	CMU .36
	3. Ship Destination	.51	DMC .50, CSS .43
	1. Numerical Operations	.41	NSI .61, CSS .33

Though studies with normal subjects would not predict it, cognition of semantic systems, or general reasoning, appears to enter into Numerical Operations for both groups of subjects in the present sample. Guilford (1941) has pointed out that the factor composition of a test may change as its difficulty changes; and Zimmerman (1954) has suggested that, as test items are made more complex, general reasoning is more likely to enter into their factor composition. It seems possible that the Numerical Operations items are relatively more complex and unfamiliar for schizophrenic than for normal subjects and therefore involve general reasoning for schizophrenics though not for normals.

Factor D—CSS—Symbolic Patterns

Exp	5. Letter Series—even	.84	
	4. Letter Series—odd	.81	
	3. Ship Destination	.41	CMS .56, EMS .47
	7. Social Situations	.38	EMS .43, CMU .38
	14. Digit Symbol	.36	NSI .54, DMU .34
Con	5. Letter Series—even	.85	
	4. Letter Series—odd	.79	CMU .35
	3. Ship Destination	.43	CMS .51, DMC .50
	1. Numerical Operations	.33	NSI .61, CMS .41
	2. Verbal Comprehension	.33	CMU .72

Looking first at the experimental group, Variables 5, 4, 3, and 14 all require being able to grasp the nature of an unfamiliar system of symbols. Social Situations seems out of place. Its presence may signify a confounding of CSS variance with format variance; for in this study, Variables 5, 4, 3, and 7, and they alone, require circling the right symbol to indicate the answer.

Looking next at the control group, Variables 5, 4, and 3 appear again. Numerical Operations also is a symbolic test, though its content is meant to be familiar; with chronic schizophrenic subjects, the familiarity may be substantially reduced. The presence of Verbal Comprehension is not readily explainable.

Factor E—EMI—Sensitivity to Problems

Exp	9. Seeing Problems—II	.86	
	8. Seeing Problems—I	.83	
Con	8. Seeing Problems—I	.73	DMC .37
	9. Seeing Problems—II	.66	DMC .63

The increased factorial complexity of Seeing Problems in the control group is of interest and will be discussed in connection with Factor H. Evaluation of semantic implications, or Sensitivity to Problems, is the most clearly defined factor for the experimental group.

Factor F—DMU—Ideational Fluency

Exp	10. Ideational Fluency—I-II	.71	
	11. Ideational Fluency—III-IV	.62	DMC .34
	14. Digit Symbol	.34	NSI .54, CSS .36
Con	11. Ideational Fluency—III-IV	.72	
	10. Ideational Fluency—I-II	.69	
	12. Brick Uses (shifts)	.54	DMC .58

Divergent production of semantic units, or Ideational Fluency, has not entered into Brick Uses (shifts) to any substantial degree with normal subjects; with unlobotomized schizophrenics, it does here. The marked difference between the groups with respect to Brick Uses (shifts) will be dealt with in the discussion of Factor H. There is some previous support for relating Digit Symbol to fluency factors (Davis, 1956). This factor is the most clearly defined factor for the control group.

Factor G—EMS—Experiential Evaluation

Exp	3. Ship Destination	.47	CMS .56, CSS .41
	7. Social Situations	.43	CMU .38, CSS .38
	6. Unusual Details	.41	CMU .30
Con	6. Unusual Details	.47	DMC .42, NSI .38
	7. Social Situations	.37	

Evaluation of semantic systems, or Experiential Evaluation, appears in both groups much as expected, except for the presence of Ship Destination in the experimental group. The lobotomized subjects are apparently more differentiated in the degree to which evaluation plays a part in their performance on that test, which is also a systems test, but primarily a cognitive one.

Divergent production of semantic classes, also known as Semantic Spontaneous Flexibility, appears in the experimental group much as expected. The presence of one of the Ideational Fluency variables is understandable; a slight change in the response class on that test can increase the number of responses.

Factor H—DMC—Semantic Spontaneous Flexibility

Exp	12. Brick Uses (shifts)	.74	
	13. Alternative Uses	.64	
	11. Ideational Fluency—III-IV	.34	DMU .62
Con	9. Seeing Problems—II	.63	EMI .66
	13. Alternative Uses	.61	
	12. Brick Uses (shifts)	.58	DMU .54
	3. Ship Destination	.50	CMS .51, CSS .43
	6. Unusual Details	.42	EMS .47, NSI .38
	8. Seeing Problems—I	.37	EMI .73

In the control group there is a striking departure from the normally predicted pattern. The two marker tests, Alternative Uses and Brick Uses (shifts), appear as predicted, except that a substantial portion of the variance of Brick Uses (shifts) has been lost to the Ideational Fluency factor, also a semantic divergent production factor. What is unexpected is that four other variables also appear. The phenomenon is orderly, since all four of the additional variables are semantic in content, and since flexibility of thinking can easily be seen as a contributor to their differentiation among subjects.

Factor R—Residual

No loadings greater than .29 in absolute value are found on this factor for either group.

Hypothesis Testing

The 16 variables of the main test battery, and 14 variables of the Wechsler-Bellevue Intelligence Scale, Form I, were used to test the hypothesis: the control subjects have group mean scores that are not different from, or are lower than, those of the lobotomized subjects.

Mean scores and variabilities for matched groups of 150 lobotomized and 150 control subjects on the main battery variables are presented and compared in Table 6. Ten of the 16 variables show a significantly higher mean score for the control group. Though the intergroup differences are not great in abso-

TABLE 6

COMPARISON OF MEANS AND VARIABILITIES OF MAIN-BATTERY VARIABLES

Variable	Experimental (N = 150)		Control (N = 150)		Ratios	
	M	SD	M	SD	F	t
1. Numerical Operations	41.89	12.51	47.49	15.57	1.55**	3.36**
2. Verbal Comprehension	29.47	11.57	32.91	12.27	1.12	2.49**
3. Ship Destination	14.28	5.53	16.85	7.35	1.77**	3.35**
4. Letter Series—odd	4.03	3.11	5.42	3.48	1.25	3.63**
5. Letter Series—even	4.31	3.23	5.74	3.16	1.04	3.87**
6. Unusual Details	11.02	6.29	12.20	7.32	1.35*	1.49
7. Social Situations	11.15	3.22	12.27	3.85	1.43**	2.67**
8. Seeing Problems—I	4.75	4.22	6.80	5.19	1.51**	3.65**
9. Seeing Problems—II	3.25	3.94	5.46	4.84	1.51**	4.20**
10. Ideational Fluency—I–II	13.38	6.36	13.87	6.42	1.02	0.66
11. Ideational Fluency—III–IV	17.05	8.37	18.67	8.67	1.07	1.64
12. Brick Uses (shifts)	2.63	3.09	2.99	3.80	1.51**	0.87
13. Alternative Uses	4.34	4.20	6.15	5.45	1.68**	3.16**
14. Digit Symbol	29.09	9.45	29.15	10.52	1.23	0.05
15. Vocabulary	20.57	6.88	21.62	6.47	1.13	1.36
16. Arithmetic	5.48	2.73	6.09	2.79	1.05	1.92*

Note.—F ratio has two-tailed test of significance; t ratio has one-tailed test.
* Significant at the .05 level.
** Significant at the .01 level.

TABLE 7

COMPARISON OF MEANS AND VARIABILITIES OF WECHSLER-BELLEVUE SCALE, FORM I, SUBTESTS [a] AND IQs

Variable	Experimental [b] (N = 69)		Control [b] (N = 72)		Ratios	
	M	SD	M	SD	F	t
Information	13.43	4.43	14.39	4.69	1.12	1.24
Comprehension	9.06	3.94	9.60	3.86	1.04	0.82
Digit Span	9.35	1.88	10.08	2.10	1.25	2.15*
Arithmetic	6.01	2.67	6.72	2.83	1.13	1.51
Similarities	9.13	4.61	10.92	4.56	1.02	2.29*
Vocabulary	21.62	6.64	22.28	6.64	1.00	0.58
Picture Arrangement	7.41	3.54	8.43	4.05	1.31	1.58
Picture Completion	8.90	3.19	9.72	3.14	1.03	1.53
Block Design	16.32	8.03	18.47	7.68	1.09	1.61
Object Assembly	16.93	4.88	17.51	4.08	1.44	0.76
Digit Symbol	28.91	9.52	30.07	10.02	1.11	0.70
Verbal IQ	95.04	12.67	99.63	13.88	1.20	2.03*
Performance IQ	96.90	15.92	101.49	13.77	1.34	1.82*
Full Scale IQ	95.70	13.49	100.46	14.22	1.11	2.03*

Note.—F ratio has two-tailed test of significance; t ratio has one-tailed test.
[a] Raw scores.
[b] Matched groups of white veterans (see Tables 2 and 3).
* Significant at the .05 level.

lute value, Numerical Operations, Verbal Comprehension, Ship Destination, the Letter Series variables, Social Situations, the Seeing Problems variables, and Alternative Uses differentiate the groups at a high level of significance. One Wechsler-Bellevue subtest, Arithmetic, differentiates the groups at the .05 level. The other six variables show nonsignificantly higher scores for the control group. The first-stage comparison does not support the hypothesis.

Table 7 presents a similar comparison for matched groups of 69 lobotomized and 72 control white veterans. The variables are the 11 subtest scores and 3 IQ scores of the Wechsler-Bellevue scale, Form I. The Digit Span and Similarities subtests, and the 3 IQs, show significantly higher mean scores for the control subjects. The 9 other subtests show nonsignificantly higher mean scores for the controls. The second-stage comparison does not support the hypothesis.

Table 8 compares mean scores and variabilities of the 10 factor-analytic tests (with parts recombined) for matched groups of 50-lobotomized and 50-control white veterans. Subjects have been eliminated in order to equate the mean and variability (cf. Table 1) of the Wechsler-Bellevue Full Scale IQ for the two groups. In spite of this matching on a correlated variable (Davis, 1956), Numerical Operations, Ship Destination, Letter Series, and Social Situations show significantly higher mean scores for the control group. The third-stage comparison does not support the hypothesis, which must be rejected.

Incidental Findings

Inspection of the correlation matrices, Tables A and B (see Footnote 2), reveals that the general level of intertest correlations is high in both groups. There are, in fact, only two nonsignificant correlations, both in the experimental group. In an earlier study (Marks, Guilford, & Merrifield, 1959), which made use of Ship Destination, Social Situations, Verbal Comprehension, and Unusual Details, applying these tests to 204 normal adults, the average intercorrelation of the four tests was significantly lower than for either group in the present study (de Mille, 1961). It would appear that these schizophrenic subjects were more variable than normals along some unmeasured continuum that was responsible for an unexpectedly high level of correlation.

The degree of illness at the time of testing would seem a likely hidden source of variance. Ratings of the degree of illness by clinical personnel could not be used in this study, since raters were not available for all subjects. If a check list of relevant items such as are

TABLE 8

COMPARISON OF MEANS AND VARIABILITIES OF FACTOR-ANALYTIC TESTS

Test	Experimental [a] (N = 50)		Control [a] (N = 50)		Ratios	
	M	SD	M	SD	F	t
1. Numerical Operations	42.14	14.82	47.42	14.39	1.06	1.79*
2. Verbal Comprehension	30.20	12.01	31.88	12.09	1.01	0.69
3. Ship Destination	15.02	6.05	17.24	8.09	1.79*	1.75*
4–5. Letter Series	8.48	5.35	10.58	6.32	1.40	1.78*
6. Unusual Details	13.06	6.33	12.76	7.22	1.30	—0.22
7. Social Situations	11.24	3.17	12.34	3.13	1.02	1.73*
8–9. Seeing Problems	9.62	8.79	11.20	9.19	1.09	0.87
10–11. Ideational Fluency	31.92	13.11	31.24	13.75	1.10	—0.25
12. Brick Uses (shifts)	3.18	3.17	2.88	4.33	1.87*	—0.39
13. Alternative Uses	5.32	4.67	6.34	6.80	2.12**	0.86

Note.—F ratio has two-tailed test of significance; t ratio has one-tailed test.
[a] Matched groups of white veterans with equal group mean IQs on the Wechsler-Bellevue scale. See Tables 2 and 3.
* Significant at the .05 level.
** Significant at the .01 level.

found in the Psychotic Reaction Profile (Lorr, O'Connor, & Stafford, 1960) could be applied to the 300 clinical records as of the date of testing, some usable indicators of the degree of illness might be selected. Since the records are quite heterogeneous, however, objective criteria such as whether the patient was on an open ward or a closed ward might prove more practical. New factor analyses including several of these indicators should further simplify the factor structures that were found.

Another unmeasured variable related to test performance is, of course, level of motivation. Schizophrenic patients might be expected to be more variably motivated than the normal subjects to whom the tests have been applied. Variance left unaccounted for by the level of illness might be accounted for by level of motivation.

A look at the nontest rows of the intercorrelation matrices, Tables A and B (see Footnote 2), reveals a number of provocative apparent relationships. Weinstein and Teuber (1957) found no relationship between level of schooling and decrement in general intelligence after brain injury in normals. In the present study, correlations of the level of schooling with test variables appear to be different in the two groups: the degree of schooling seems to have mattered less to performance after prefrontal lobotomy.

In both groups, the males apparently did better on the general reasoning tests. The paranoid and hebephrenic categories appear to have patterns of correlation with tests that are distinct from each other and distinct in the two groups. In the lobotomized group, there are six negative correlations between test performance and being on a tranquilizer. In the control group there are none; possibly there was an interaction between lobotomy and tranquilizers that impaired performance.

Smith and Kinder (1959) attributed performance decrements of topectomized patients to neurological degeneration during an 8-year period. The postoperative interval in the present study ranged from 5 to 15 years, and yet no significant relationship with any test variable appeared.

In general, correlations between test and nontest variables were very low, except for the apparent trends that have been mentioned.[5]

Discussion

Consequences of Lobotomy

Within the limitations of the samples of tests and subjects, it has been shown that the structure of intellect was different for lobotomized and unlobotomized schizophrenics, as were the factor compositions of certain tests. The tests were: Seeing Problems, Ship Destination, Brick Uses (shifts), and Unusual Details. Evaluative and productive factors, especially the flexibility factor, made markedly different contributions to test performance in the two groups. Semantic Spontaneous Flexibility was a poorly differentiated ability for the control group, entering into many tests that had substantial loadings on other factors; for the lobotomized group it was well differentiated, entering into tests much as it has with normal subjects.

At the same time, it has been shown that the performance of the lobotomized subjects was significantly inferior to that of the control subjects on most of the factor-analytic tests.

[5] These observations call attention to the fact that a good deal of unanalyzed information remains in this study. The effects of selected nontest variables are still to be evaluated. The consequences of lobotomy could be more directly assessed by combining the groups and factoring anew, with a group indicator as one of the variables. The relationships among the 11 Wechsler-Bellevue subtest scores and the other test variables remain to be determined. In these pursuits, both orthogonal and oblique rotations should be made to extract as much information as possible and to discover what second-order factors may be found. Finally, it may be possible to develop a factor-score equation maximizing the predictive ability of the Wechsler-Bellevue scale with respect to frontal lesions. Given the opportunity, the investigator proposes to carry out these analyses and report on them at a later date.

Two additional findings have been reported elsewhere. The first (de Mille, 1962a) involved a comparison of group mean scores on 13 tests between a group of 22 subjects who had received bilateral transorbital lobotomies and a matched group of 22 subjects who had received bilateral prefrontal lobotomies. The prefrontal group had higher mean scores on all tests. The second finding (de Mille, 1962b) was entirely unexpected. A tendency was noticed for patients who had been diagnosed as catatonics before operation to develop postlobotomy epilepsy in greater proportion than the other lobotomized schizophrenics ($p < .001$).

Even after group matching on a test-correlated variable, Wechsler-Bellevue IQ, the lobotomized subjects had lower mean scores on factor-analytic systems tests and a test that may reflect long-term memory.

Combining these observations, it appears that schizophrenic patients may have a pervasive problem with rigidity of thinking that is modified by a prefrontal lobotomy, but at a cost of impairment of a number of factor-defined abilities. The patients become less able and at the same time less characteristically schizophrenic.

Just how a lobotomy could bring about a more normal contribution of flexibility to thinking is not easily explained. The impairment of the lobotomy group, in general, was not very great in comparison with the control group. Possibly a relatively small additional impairment, not greatly aggravating the existing general schizophrenic deficit, of some capacity such as the ability to hold a set, to maintain a systematic determining tendency, to remain "programed," or as Pribram (1960) suggests, to hold on to or generate an intention, might reduce rigidity of thinking.

It must be remembered that most of the impairment took place before lobotomy. One is left with the task of extrapolating from small intergroup differences, both in test performance and in the factorial composition of tests, to the presumably greater differences that would accompany lobotomy in nonpsychotic (Tow, 1955) or nonpsychiatric (Klebanoff, Singer, & Wilensky, 1954; Rosvold & Mishkin, 1950) groups. "It is still possible that long-term planning and initiative, creative work and capacity for radical readjustment may be notably impaired by lesions of the frontal lobes" (Hebb, 1945, p. 23).

The changes attributed to lobotomy in the present study are limited by the factor content of the battery that was selected. It cannot be inferred that these changes exhaust the possible effects of lobotomy. Applications of the complete factor-analytic test armamentarium would be necessary before it could be said that the consequences of lobotomy were known even within the limits of the existing instruments. Certainly, memory factors and long and short range planning abilities would have to be studied. In the nonintellective area, questions of effects on anxiety and the self concept would likewise demand study.

Factor-Defined versus Task-Defined Tests

Tables 6, 7, and 8 reveal what is apparently a marked superiority of the factor-defined over the task-defined tests in differentiating the lobotomized from the unlobotomized subjects. This performance is all the more impressive when one remembers that the factor-defined tests were designed as group tests and yield skewed score distributions, while the task-defined tests were designed for individual administration and yield nearly normal distributions.

Verbal Comprehension, Seeing Problems, and Alternative Uses discriminate between the groups before IQ matching, but not after. One may infer that they are more highly correlated with IQ than are Numerical Operations, Ship Destination, Letter Series, and Social Situations, whose ability to discriminate seems markedly independent of manipulation of the IQ.

The greater articulation of our picture of intellect through the use of factor-analytic tests makes it possible to say not only that lobotomy results in a measurable impairment of intellect but, also, that the impairment appears in some areas of intellect though not in others. The results imply, for example, that lobotomized patients may be relatively impaired on one productive ability (flexibility) but not on another (fluency). The development of factor scores and their application in new studies would help to clarify the pattern of factor-defined abilities that are impaired by lobotomy in schizophrenia, and that may be impaired by other types of frontal lesions.

Three factors, Experiential Evaluation (EMS), Ideational Fluency (DMU), and, particularly, Flexibility (DMC), make different contributions to the factor compositions of tests for the two groups in this study. In support of Harvey's (1950) hypothesis, it may be shown with these results how a difference in a factor-defined ability may be obscured by a change in the factor composition of a test. Comparison of group mean scores (see Table 6) reveals the following indicative relationships. Ideational Fluency, representing the fluency factor primarily, does not differentiate the groups. It may be inferred from this that the groups do not differ in the factor-defined ability, Ideational Fluency. Alternative Uses, representing flexibility, does differentiate the groups, suggesting that the groups are different in the factor-defined ability, Semantic Spontaneous Flexibility. Brick Uses (shifts), however, represents flexibility for the experimental group, but represents flexibility and fluency in about equal measure for the control group; it does not differentiate the groups. If Brick Uses (shifts) were the only test employed, and if it were supposed to be a test of flexibility for both groups, then its failure to differentiate would obscure the difference between the groups in the factor, Semantic Spontaneous Flexibility.

The foregoing illustrates a strength of the factor-analytic approach. It also serves as a warning that when tests are known to have a different factor composition for two groups, or when their factor composition is not known, caution must be exercised in interpreting raw score differences. Statements about intergroup differences, when factor loadings are different for the groups, should be based upon weighted scores that take the factor loadings into account.

Failures to measure intellectual consequences of psychosurgery may have been due in part to this kind of experimental shortcoming. It was noted, for example, in the first Columbia-Greystone report (Mettler, 1949) that "tests which had shown rather high correlation before operation may decline in their relation after operation. . . . Tests which show no preoperative correlation may show one after operation" (p. 176). A change in the factor composition of the tests may be inferred from this observation. The fact that no difference was observed in test performance may have obscured a realignment of factor-defined abilities consequent to the operations.

Implications of the Study

In the introductory remarks it was stated that a goal of the study was to encourage and facilitate increased application of multifactorial methods to problems in physiological, abnormal, and clinical psychology. It would seem that progress has been made toward that goal, for an old and familiar problem that stubbornly refused to yield to a task-defined approach has proved less impervious to a factor-defined approach, even with tests developed for quite a different population.

Many of the difficulties attending the present investigation would be absent in studies of schizophrenia or mental deficiency, or in studies making use of psychopharmacological variables. It is too late for a thorough elucidation of the effects of psychosurgery, but the present study shows what might have been done had suitable arrays of factor-defined tests been available. More important, a direction has been given for future research. The surface has been scratched, and no more, in the development of factor-analytic instruments for use with abnormal populations, but there can be little doubt whether such instruments are needed for theoretical work in clinical and physiological psychology.

SUMMARY

Factor-analytic measurement has contributed little to physiologico-behavioral problems in spite of Lashley's prediction, in 1941, that it would speed the coalescence of neurology and psychology. At the same time, the literature on intellectual consequences of brain lesions has amply demonstrated the failure of "general intelligence" tests to analyze or often, even to detect the effects of brain modifications such as those of psychosurgery. Work with animals is bringing the specification of unitary functional brain systems to a point where it outstrips psychological specification of unitary behavioral parameters. If research is to be fruitful, psychological variables must be as well conceived and specified as physiological variables.

This investigation sought to improve measurement technique in physiological, abnormal and clinical psychology by attacking an old but still unsolved problem, the measurement of intellectual effects of psychosurgery. Both traditional and contemporary methods of measurement were used, and their relative effectiveness was assessed.

A 16-variable factor-analytic test battery

relating to Guilford's structure-of-intellect model, was individually administered to 150 lobotomized schizophrenics and 150 schizophrenic controls, group-matched for age, schooling, duration of illness, sex, race, diagnostic subgroup, veteran status, hospital, and presence of a tranquilizer. Intermediate matched subgroups of 69 lobotomized, and 72 control, white veterans were tested with the Wechsler-Bellevue Intelligence Scale, Form I. From these intermediate subgroups, a further selection was made, to equate group mean Wechsler-Bellevue IQs for matched subgroups of 50 lobotomized and 50 control subjects.

Nine factors were extracted from the intercorrelations of the 16 factor-analytic test variables for each of the two main groups of 150 subjects. Graphic orthogonal rotations were made, to the criteria of simple structure, positive manifold, and, finally, psychological meaning. Analytical rotations were made and compared. Eight interpretable factors and one residual appeared for each group. It was apparent that the two factor structures were not the same, the most striking dissimilarity involving the factor of Semantic Spontaneous Flexibility, with regard to which the lobotomized group, more than the control group, resembled normals. There were intergroup differences in the factor composition of four of the tests. No suitable test of the statistical significance of the factor-structure differences was found to be available.

Group mean test score comparisons were made for the large, intermediate, and small pairs of groups. The control subjects had significantly higher mean scores on most of the factor-analytic tests, on all three Wechsler-Bellevue IQs, and on three Wechsler-Bellevue subtests. Four of the factor-analytic tests differentiated the groups that had been matched on Wechsler-Bellevue IQ.

It was concluded that prefrontal lobotomy may modify an apparently pervasive inflexibility of thinking so as to make the patient less characteristically schizophrenic, but at a cost of impairment of a number of factor-defined intellectual abilities. The factor-analytic battery was found to be both more sensitive and more differentiated than the general intelligence test in its discrimination between the groups, lending support to the contention that factor-analytic instruments and methods are indispensable for some phases of research in physiological, abnormal, and clinical psychology. If differential impairment of intellectual abilities after lobotomy can be shown in subjects already much impaired by chronic schizophrenia, through the use of factor-analytic instruments and factor-analytic treatment of data, then the use of such tests and methods in more current problems dealing with intellectual, physiological, psychopharmacological, and psychopathological variables should prove most fruitful.

A warning was given that caution must be exercised in interpreting scores when the factor composition of a test for each class of subjects under study is not taken into account. Actual differences may be obscured, and observed differences may be misinterpreted, when factor composition is ignored. The almost universal failure to measure intellectual consequences of psychosurgery may have been due in part to this kind of experimental oversight.

REFERENCES

CHRISTENSEN, P. R., & GUILFORD, J. P. *Ship Destination Test: Form A*. Beverly Hills, Calif.: Sheridan Supply, 1955.

CHRISTENSEN, P. R., & GUILFORD, J. P. *Ideational Fluency I: Form A*. Beverly Hills, Calif.: Sheridan Supply, 1957.

DAVIS, P. C. A factor analysis of the Wechsler-Bellevue scale. *Educ. psychol. Measmt.*, 1956, **16**, 127–146.

DE MILLE, R. Intellect after lobotomy in schizophrenia: A factor-analytic study. Unpublished doctoral dissertation, University of Southern California, 1961.

DE MILLE, R. Intellectual effects of transorbital versus prefrontal lobotomy in schizophrenia: A follow-up study. *J. clin. Psychol.*, 1962, **18**, 61–62. (a)

DE MILLE, R. An incidental observation of epilepsy in catatonic schizophrenics after prefrontal lobotomy. *J. nerv. ment. Dis.*, 1962, **134**, 182–183. (b)

FREEMAN, W., & WATTS, J. W. *Psychosurgery*. Springfield, Ill.: Charles C Thomas, 1942.

GUILFORD, J. P. The difficulty of a test and its factor composition. *Psychometrika*, 1941, **6**, 67–77.

GUILFORD, J. P. *Fundamental statistics in psychology and education*. (3rd ed.) New York: McGraw-Hill, 1956.

GUILFORD, J. P. *Personality*. New York: McGraw-Hill, 1959.

GUILFORD, J. P. Factorial angles to psychology. *Psychol. Rev.*, 1961, **68**, 1–20.

GUILFORD, J. P., CHRISTENSEN, P. R., WILSON, R. C., & MERRIFIELD, P. R. *Alternate Uses: Form A.* Beverly Hills, Calif.: Sheridan Supply, 1959.

GUILFORD, J. P., & MERRIFIELD, P. R. The structure of intellect model: Its uses and implications. *Univer. Sth. Calif. Psychol. Lab. Rep.*, 1960, No. 24.

GUILFORD, J. P., & ZIMMERMAN, W. S. *The Guilford-Zimmerman Aptitude Survey.* Beverly Hills, Calif.: Sheridan Supply, 1947.

HALSTEAD, W. C. *Brain and intelligence.* Chicago: Univer. Chicago Press, 1947.

HARVEY, H. Cognitive changes in brain-injured children: A preliminary survey. Unpublished doctoral dissertation, University of Southern California, 1950.

HEBB, D. O. Man's frontal lobes: A critical review. *AMA Arch. Neurol. Psychiat.*, 1945, **54**, 10–24.

HERTZKA, A. F., & GUILFORD, J. P. *Logical Reasoning: Form A.* Beverly Hills, Calif.: Sheridan Supply, 1955.

KAISER, H. F. The varimax criterion for analytical rotation in factor analysis. *Psychometrika*, 1958, **23**, 187–200.

KLEBANOFF, S. G. Psychological changes in organic brain lesions and ablations. *Psychol. Bull.*, 1945, **42**, 585–623.

KLEBANOFF, S. G., SINGER, J. L., & WILENSKY, H. Psychological consequences of brain lesions and ablations. *Psychol. Bull.*, 1954, **51**, 1–41.

LASHLEY, K. S. Coalescence of neurology and psychology. *Proc. Amer. Phil. Soc.*, 1941, **84**, 461–470.

LEWIS, N. D. C., LANDIS, C., & KING, H. E. (Eds.) *Studies in topectomy.* New York: Grune & Stratton, 1956.

LORR, M., O'CONNOR, J. P., & STAFFORD, J. W. The Psychotic Reaction Profile. *J. clin. Psychol.*, 1960, **16**, 241–245.

MARKS, A., GUILFORD, J. P., & MERRIFIELD, P. R. A study of military leadership in relation to selected intellectual factors. *U. Sth. Calif. Psychol. Lab. Rep.*, 1959, No. 21.

MERRIFIELD, P. R., GUILFORD, J. P., CHRISTENSE[N,] P. R., & FRICK, J. W. A factor-analytic study [of] problem-solving abilities. *U. Sth. Calif. Psych[ol.] Lab. Rep.*, 1960, No. 22.

METTLER, F. A. (Ed.) *Selective partial ablation of [the] frontal cortex.* New York: Hoeber, 1949.

PRIBRAM, K. H. A review of theory in physiologi[cal] psychology. *Annu. Rev. Psychol.*, 1960, **11**, 1–

ROSVOLD, H. E., & MISHKIN, M. Evaluation of [the] effects of prefrontal lobotomy on intelligen[ce.] *Canad. J. Psychol.*, 1950, **4**, 122–126.

SCHERER, I. W., KLETT, C. J., & WINNE, J. F. Ps[y]chological changes over a five-year period followi[ng] bilateral prefrontal lobotomy *J. consult. Psych[ol.]* 1957, **21**, 291–295.

SMITH, A., & KINDER, ELAINE F. Changes in ps[y]chological test performance of brain-operat[ed] schizophrenics after 8 years. *Science*, 1959, **1[…]**, 149–150.

THURSTONE, L. L. *Multiple factor analysis.* Chicag[o:] Univer. Chicago Press, 1947.

THURSTONE, L. L., & THURSTONE, T. G. *SRA Prima[ry] Mental Abilities: Intermediate, Form AH.* Chicag[o:] Science Research Associates, 1947.

TOW, P. M. *Personality changes following fron[tal] leucotomy.* London: Oxford Univer. Press, 1955.

WECHSLER, D. *The measurement of adult intelligen[ce.]* Baltimore: Williams & Wilkins, 1944.

WECHSLER, D. *Wechsler-Bellevue Scale for Adol[es]cents and Adults: Form I.* New York: Psycholo[gi]cal Corporation, 1947.

WECHSLER, D. *The measurement and appraisal [of] adult intelligence.* Baltimore: Williams & Wilki[ns,] 1958.

WEINSTEIN, S., & TEUBER, H. L. The role of preinju[ry] education and intelligence level in intellectual l[oss] after brain injury. *J. comp. physiol. Psych[ol.]*, 1957, **50**, 535–539.

ZIMMERMAN, W. S. A simple graphical method [for] orthogonal rotation of axes. *Psychometrika*, 19[46,] **11**, 51–55.

ZIMMERMAN, W. S. Hypotheses concerning the natu[re] of the spatial factors. *Educ. psychol. Meas[mt.],* 1954, **14**, 396–400.

(Received June 15, 1961)