1996

Design Fluency Test: Normative data, cognitive skills related to performance, and performance by individuals with temporal lobe pathology.

Sherri Leigh. Carter
University of Windsor

DESIGN FLUENCY TEST: NORMATIVE DATA, COGNITIVE SKILLS RELATED TO PERFORMANCE, AND PERFORMANCE BY INDIVIDUALS WITH TEMPORAL LOBE PATHOLOGY

by

Sherri L. Carter

B. A. University of Western Ontario, 1994

A Thesis
Submitted to the Faculty of Graduate Studies and Research
through the Department of Psychology
in Partial Fulfilment of the
Requirements for the Degree
of Master of Arts at the
University of Windsor
Windsor, Ontario, Canada
1996

0-612-30891-X

ABSTRACT

The Design Fluency Test (DFT) requires the invention of abstract designs within a limited amount of time. There are two conditions of the test: a free condition without restrictions and a fixed condition where designs must consist of exactly four lines. DFT performance was investigated for healthy controls (n = 66), left temporal lobectomy patients (n = 44), right temporal lobectomy patients (n = 40), right frontal resection patients (n = 8) and left frontal resection patients (n = 1). Groups of individuals with frontal resections were too small to include in data analyses. Each participant in the control group completed a battery of measures sensitive to verbal fluency/production, visuospatial skills, visual attention, psychomotor speed, and level of psychometric intelligence. Regression analyses demonstrated that performance on these measures was not able to predict DFT score in the free condition, although combinations of these skills were related to output in the fixed condition. Inter-rater reliability among 3 raters for 44 DFT protocols was generally good to excellent, although nameable errors and designs with the incorrect number of lines yielded poorer reliability coefficients. The right and left temporal lobectomy patients committed significantly more nameable errors in the free condition of the DFT than controls, but there were no differences among groups on indices of perseveration or novel output scores. DFT performance did not differ between pre-operative and post-operative assessment for groups with right or left temporal lobectomies. The results of this study do not discredit the clinical use of the DFT as a measure sensitive to possible right frontal-lobe dysfunction. Future research should address the possible contribution of "creativity" or divergent thought to DFT performance and the replication of group differences for nameable errors on the DFT.

ACKNOWLEDGEMENTS

I would like to express my appreciation for the time, effort, and valuable advice of my committee members, Drs. D. Shore, D. Reynolds and P. Weir. As my thesis advisor, I would like to thank Dr. D. Shore for his support and guidance throughout the year.

I would also like to express my gratitude to the Department of Psychological Services at London Health Sciences Centre - University Campus. I greatly appreciate the time and effort of Drs. M. Harnadek and C. Kubu in the scoring of Design Fluency protocols. I especially thank Dr. M. Harnadek whose extensive help and guidance throughout all stages of this project enabled its successful completion.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Background on the Design Fluency Test

The Design Fluency Test (DFT; Jones-Gotman & Milner, 1977) is a test that requires individuals to generate novel designs which are abstract and cannot be named. The instructions given during administration of the test also stipulate that scribbles are not accepted and that designs should be very different from one another. The test has both a free and a fixed (i.e., four-line) condition.

In the free condition, individuals are required to "invent" as many different abstract designs as possible within a five minute interval. In the fixed condition, individuals are still required to draw as many different abstract designs as possible but each design must consist of exactly four lines and the time limit is four rather than five minutes. Jones-Gotman (unpublished manuscript, no date) has reported that the free condition is more sensitive than the fixed condition. The reduced sensitivity reported for the fixed condition is likely related to the restrictions placed on output in this condition, which result in a limited range of possibilities for novel designs.

General guidelines for scoring the DFT are described in the original article by Jones-Gotman and Milner (1977).

Additional unpublished scoring instructions with specific criteria and examples are available from Jones-Gotman (unpublished manuscript, no date). Designs are scored as nameable errors when they represent actual objects, letters, numbers, conventional signs, or nameable abstract forms such as regular geometric shapes. Perseverative errors include rotations or minor variations of previous designs, variations on a theme, vague forms, and scribbles. The fixed condition has an additional type of error for designs with the incorrect number of lines.

The Design Fluency Test was originally developed as a nonverbal analogue of the Thurstone Word Fluency Test (Thurstone, 1938), a written measure of verbal fluency that involves generating words which begin with a particular letter. Consequently, the Design Fluency Test was considered to be sensitive to fluency in the nonverbal domain. In the original article, Jones-Gotman and Milner (1977) examined the DFT performance of individuals who had undergone the surgical excision of focal cortical areas as treatment for intractable epilepsy. These authors predicted that performance on the DFT would be more reliant on the right cerebral hemisphere and that individuals with anterior lesions would exhibit poorer nonverbal fluency relative to those with posterior lesions.

For the free condition, Jones-Gotman and Milner (1977) reported that participants with anterior (i.e., frontal, fronto-central, and temporal) lesions of the right hemisphere produced significantly fewer novel drawings than a control group. In contrast, the number of novel designs produced by a group of participants with left-hemisphere lesions in corresponding regions did not differ significantly from controls.

In addition, individuals with lesions of the frontal or fronto-central regions of the right hemisphere exhibited significantly reduced novel output relative to participants with lesions of other cortical areas (i.e., left frontal and fronto-central lesions; both right and left temporal and posterior lesions). These individuals also made significantly more perseverative errors relative to the control group.

The incidence of DFT protocols with more than one nameable error was significantly higher in the group with right temporal-lobe lesions relative to controls and the groups with left temporal-lobe lesions. The authors did not conduct a statistical analysis for the incidence of nameable errors in the groups with frontal lesions due to the small number of cases in these groups. Since individuals with right frontal or fronto-central lesions produced fewer novel designs than controls and than patient groups with lesions in other areas, and committed more perseverative errors than the control group, it is conceivable that this group of patients would also make more nameable errors.

In the fixed condition, both participants with right-hemispheric lesions and those with left-hemispheric lesions produced significantly fewer designs than the control group. Again, participants with right-hemispheric lesions in the frontal or fronto-central regions produced significantly fewer designs than all other lesion groups. A combined group of individuals with right-sided frontal and fronto-central lesions and those with left-sided frontal lesions made significantly more perseverative errors than individuals with lesions of other cortical areas.

Fluency Tasks and the Frontal Lobes

Damage to the frontal regions of the cerebral hemispheres has been associated with behaviour commonly described as decreased initiative and impersistence in the completion of tasks (Lezak, 1995). One facet of this behaviour that is tapped by neuropsychological assessment is decreased productivity, as evidenced by poor performance on verbal fluency tasks. Individuals with lesions of the left frontal lobe exhibit particularly poor performance on measures of verbal fluency, even in patients who demonstrate no evidence of aphasia (Milner, 1964). This association between left frontal lesions and impaired performance on both oral and written measures of verbal fluency has been well documented (Benton, 1968; Jones-Gotman & Milner, 1977; Milner, 1964; Perret, 1974; Ramier & Hecaen, 1970, as cited in Damasio & Anderson, 1993). Pendleton, Heaton, Lehman and Hulihan (1982) analyzed data on the Thurstone Word Fluency Test (Thurstone, 1938) from 203 individuals with cerebral damage and reported that individuals with frontal lesions or lesions that included frontal areas earned significantly lower scores on this test then patients with lesioned areas outside of the frontal lobes. For individuals with damage restricted to the frontal lobe, left-sided lesions produced more impairment on the fluency task than right-sided lesions. Again, this finding is consistent with reports in the literature which document the sensitivity of verbal fluency tasks to frontal lobe lesions, with more impairment in productivity noted for individuals with left-sided lesions.

Although the DFT is of a nonverbal rather than a verbal nature, it is similar to verbal fluency tasks in that it requires productivity which meets particular specifications. Thus, it is not surprising that groups of individuals with lesioned areas that include the frontal lobes exhibit the greatest degree of impairment on the DFT. In addition, the fact that the poorest performances on this task are by individuals with right-sided frontal lesions is consistent with the nonverbal nature of the DFT, since the right hemisphere appears to be more active in the processing of nonverbal material relative to the left hemisphere (for review see Springer & Deutsch, 1993). Therefore, the cerebral areas associated with impaired DFT performance in the original study (Jones-Gotman & Milner, 1977) are supported by neuropsychological theory and previous research documented in the literature.

Clinical Applications

Since the original article by Jones-Gotman and Milner (1977), impaired DFT performance relative to controls has been documented for individuals with probable Alzheimer's dementia (Mickanin, Grossman, Onishi, Auriacombe & Clark, 1994) and schizophrenia (Kolb & Whishaw, 1983). Other studies have documented poor performance on the DFT for individuals with Huntington's disease (Jason et al., 1988), frontal lobe syndrome (Canavan, Janota, & Schurr, 1985), and lesions of the right dorsomedial nucleus of the thalamus (Speedie & Heilman, 1983).

A recent study by Varney et al. (1996) investigated the performance of patients with closed head injury on the free condition of the DFT. These authors cite Jones-Gotman and Milner (1977) as the authors of the DFT but do not refer to the additional unpublished monograph of scoring criteria available from Jones-Gotman (no date). Hence, it is assumed that designs in the Varney et al. (1996) study were scored according to the test description contained in the original 1977 article (Jones-Gotman & Milner, 1977) but without the aid of the additional scoring instructions and examples.

In this study, individuals with closed head injury produced significantly fewer novel designs than a control group (Varney et al., 1996). Although failure on the DFT was more frequent for participants who had lost consciousness for more than ten minutes, the correlation between novel output score and duration of loss of consciousness was not significant. The authors attribute this low correlation to the limited number of patients in their sample who had experienced a loss of consciousness greater than ten minutes in duration. The results of this study suggest even more strongly that the DFT may be a useful tool in clinical assessment and an indicator of neuropsychological impairment in patients with brain injury.

Varney et al. (1996) reported that nameable responses comprised a significant proportion of total responses in the closed head injury group but not in the healthy control group. This finding is consistent with the performance of some groups of individuals who have undergone the surgical removal of cerebral areas as treatment for intractable epilepsy (Jones-Gotman & Milner, 1977).

In contrast, Varney et al. (1996) did not find the proportion of perseverative errors in the head-injured group to be significantly greater than in the control group. In part, this finding may be related to the specific population under investigation. Other research has documented a higher incidence of perseverative errors on DFT protocols for populations with dementia (Bigler, 1995, as cited in Varney et al., 1996) and individuals who have undergone removal of focal cortical areas as treatment for intractable epilepsy (Jones-Gotman & Milner, 1977). It is also possible, however, that this finding is related to DFT scores obtained without the use of the additional scoring criteria (Jones-Gotman, unpublished manuscript, no date). In the unpublished monograph, perseverative errors are defined more clearly than in the original article (Jones-Gotman & Milner, 1977), examples of different types of perseverative errors are given, and raters are instructed to score perseverations harshly. The low number of perseverative errors reported by Varney et al. (1996) for both healthy controls (95% made 3 or fewer repeated designs) and head-injured patients (3% made 4 or more repeated designs) may be partially explained by the omission of the additional scoring criteria.

Psychometric Characteristics

The DFT is used fairly widely in research settings and is a potentially useful clinical instrument (Woodard et al., 1992), but there is a paucity of normative data on the test. Until the recent studies by Varney et al. (1996) and Woodard et al. (1992), published normative data on the DFT was limited to the performance of the 34 control subjects in Jones-Gotman and Milner's (1977) original article (Ruff, Allen, Farrow, Niemann & Wylie, 1994). For each of these studies, WAIS-R Full Scale IQ scores were not reported for any of the control groups, and the means and standard deviations of DFT scores were not reported in the Jones-Gotman and Milner (1977) study other than in bar graphs.

Preliminary findings suggest that the free condition of the DFT is amenable to standardization for clinical use (Varney et al., 1996). In order to assist in the interpretation of DFT performance in clinical populations, it is crucial to establish a larger base of normative data on healthy participants with a clearly defined sample composition.

Varney et al. (1996) reported that for their control group, the mean number of designs produced in the free condition was 16.1 with substantial variation between individuals (SD = 9.1; Range = 4-51). The novel design score was normally distributed among controls. Novel output score did not correlate significantly with age and/or years of education and males and females did not differ significantly in the number of designs produced.

Woodard et al. (1992) have reported relatively good reliability for the DFT with independent scoring by two raters in a sample of 80 older healthy adults. DFT protocols were scored according to the original article by Jones-Gotman and Milner (1977), and a personal communication with Jones-Gotman (April 30, 1984, as cited in Woodard et al., 1992) suggests that the additional unpublished scoring criteria (Jones-Gotman, no date) may have been used, although this issue is not clear. Rater consistency was good to excellent according to criteria specified by Cicchetti and Sparrow (1981); the authors interpreted this to suggest that raters tended to maintain the same rank order of subjects in terms of error frequency. Varney et al. (1996) also reported that for protocols from their groups of participants, two trained, independent raters agreed 90% of the time on the number of novel designs produced.

Based on differences noted between the score distributions of the two raters, however, Woodard et al. (1992) suggested that "using a single rater may result in highly variable raw scores, depending on the rater's degree of scoring leniency." (p. 176). Consequently, these authors recommend the use of more than one rater in the establishment of a normative data base.

DFT performance appears to discriminate accurately between older and younger individuals and is sensitive to neuropsychological impairments associated with aging (Mittenberg, Seidenberg, O'Leary, & DiGiulio, 1989). Given that the sample used by Woodard et al. (1992) consisted of older subjects, with a mean age of 69.4 and standard deviation of 10.6 years, an investigation of DFT performance in younger individuals is required to develop adequate normative data. In order to replicate and extend the findings of Woodard et al. (1992), a younger sample and multiple raters are essential elements in subsequent investigations of the DFT.

Another inter-rater reliability study of DFT scores has been conducted for approximately 400 children aged 5 through 14 years (M. Jones-Gotman, personal communication, March 21, 1996). The three raters used in this study had all received similar training in the administration and scoring of the DFT and correlations among the raters were "excellent" (M. Jones-Gotman, personal communication, March 21, 1996). An investigation of the inter-rater reliability of DFT scores in individuals older than 14 and younger than 60 years of age is needed to extend the findings of these two previous inter-rater reliability studies.

Cognitive Skills Related to DFT Performance

An important issue that has not been addressed in the literature pertains to the types of cognitive skills that contribute to effective performance on the DFT. Skills and abilities related to this test of nonverbal fluency have yet to be empirically investigated. Given its reported sensitivity to right-hemispheric lesions and the significant role the right hemisphere plays in the processing of nonverbal material (for review see Springer & Deutsch, 1993), one hypothesis is that performance on the DFT may be related to skills and abilities that lie within the nonverbal domain. In fact, one group of researchers has reported that DFT performance appears to be related to performance on measures that they consider to be nonverbal (i.e., copy and delayed recall of the Rey-Osterrieth figure and visual memory span of the Wechsler Memory Scale-Revised) in a group of patients with probable Alzheimer's dementia (Mickanin et al., 1994).

It may be, then, that DFT performance is related to performance on other measures of nonverbal abilities such as those sensitive to visuospatial skills (e.g., closure, perceptual speed, mental rotation of figures in space, detection of embedded figures) and/or visual attention. Similarly, DFT performance may also be related to performance on the nonverbal components of standardized measures of psychometric intelligence (e.g., Block Design subtest of the Wechsler Adult Intelligence Scale - Revised (WAIS-R)). This possibility is suggested further by some evidence that performance on verbal elements of psychometric intelligence measures (e.g., Vocabulary subtest of the WAIS-R) may be related to performance on measures of verbal fluency (Parks et al., 1988).

It is unclear whether the ability to generate novel abstract designs is related to general intellectual level. Varney et al. (1996) reported that for a group of individuals with closed head injury, novel output scores did not correlate significantly with general intellectual level as measured by the Wechsler Adult Intelligence Scale (WAIS) Full Scale IQ. However, this study did not examine the relationship between DFT performance and psychometric intelligence in a healthy control group. It is important to clarify whether factors contributing to effective DFT performance differ between participants serving as controls and individuals with neuropsychological impairment.

Time constraints placed on individuals during the adminstration of the DFT implicate speed as another factor that may contribute to effective performance on this task. Varney et al. (1996) reported that for their head-injured sample, novel output score on the DFT did not correlate significantly with any of the individual WAIS subtests except Digit Symbol, a subtest sensitive to graphomotor speed. Although the correlation was significant, it was not particularly high ($r = 0.22$), which suggests that psychomotor speed is not the main contributor to DFT performance in individuals with head injury. This preliminary evidence does, however, imply that graphomotor speed may have some impact on DFT performance. The relative importance of speed as a contributor to DFT performance may be clarified through an investigation of the relationship between performance levels on tasks sensitive to graphomotor speed and the DFT.

A further question is whether performance on measures of verbal fluency (e.g., generating words that begin with a particular letter) is related to fluency in the nonverbal domain. The DFT was originally developed by Jones-Gotman and Milner (1977) as an analogue of the Thurstone Word Fluency Test (Thurstone, 1938) and the validity of this claim has not yet been clearly established. Jones-Gotman and Milner (1977) reported that individuals with right fronto-central lesions exhibited impaired DFT performance but showed "little or no deficit" (p. 671) on the Thurstone Word Fluency Test. A double dissociation was present in that individuals with left central lesions generated low word fluency scores but relatively unimpaired scores on the DFT. This preliminary evidence suggests that right anterior lesions may be associated with impaired nonverbal fluency but relatively less impaired verbal fluency. In contrast, left frontal lesions may be related to less impairment in nonverbal fluency relative to verbal fluency.

Other preliminary evidence suggests that scores on the DFT may be related to scores on traditional measures of verbal fluency. Varney et al. (1996) reported that for individuals with closed head injury, novel output scores on the DFT correlated most highly with performance on Controlled Oral Word Association (COWA; $r = 0.34$, $p < .05$), a measure of verbal fluency. Although this correlation was not particularly high, the fact that COWA was the measure most related to novel output score on the DFT suggests that the two types of fluency may be associated. It is critical to determine whether a relationship exists between DFT performance and measures of verbal fluency in order to support or reject the notion that the DFT measures "fluency" as opposed to other constructs (e.g., visuospatial skills, psychometric intelligence).

The Varney et al. (1996) study is the first study to attempt to relate DFT performance to performance on other neuropsychological measures, and represents a significant step forward in delineating the relative contributions of cognitive skills to DFT performance. However, this study did not examine the relationship between DFT performance and performance on other neuropsychological measures in a healthy control group. It is possible that the factors which contribute to effective DFT performance may differ between healthy individuals and those with some form of neuropsychological deficit. It is even possible that contributing factors may differ somewhat according to the particular neuropsychological population in question (e.g., closed head injury versus surgical excision of cortical areas as treatment for epilepsy). Furthermore, findings from the Varney et al. (1996) study should be regarded as a preliminary exploration of the relationships between DFT performance and other measures because an investigation of this nature was not the primary focus of their research.

Purpose and Hypotheses

A major purpose of the present study is to generate normative data on the DFT for a sample of healthy control subjects. The goal of this data collection is to provide a larger normative data base that will contribute to standardization of the DFT for clinical use. The DFT protocols will be scored by multiple raters to assess inter-rater reliability in a younger sample than in the study by Woodard et al. (1992) and in an older sample than in the study by Jones-Gotman (personal communication, March 21, 1996).

A second aim of this study is to investigate the cognitive skills and abilities that contribute to performance on the DFT. An attempt to delineate the skills that the DFT is sensitive to will assist in the interpretation of test performance and the appropriateness of clinical inferences based on DFT performance. The nonverbal nature of the DFT suggests that performance on this task may be related to performance on measures sensitive to nonverbal skills and abilities. Also, time constraints on the test may be reflected in a relationship between DFT performance and measures sensitive to graphomotor speed. Further, because the DFT was originally designed as a nonverbal analogue of a word fluency measure and creates a similar demand for productivity under particular specifications, there may be a relationship between DFT performance and measures of verbal fluency. Psychometric intelligence is yet another element that may be related to DFT performance in a systematic fashion.

In current clinical practice, the DFT is often used as an indicator of the functional state of systems that subserve the frontal regions of the right cerebral hemisphere. In conjunction with other neuropsychological evidence, poor performance on the DFT may be interpreted as an indication of right frontal-lobe dysfunction in individuals who do not exhibit any evidence of right frontal lobe pathology on neurological diagnostic instruments (e.g., computerized tomography (CT), magnetic resonance imaging (MRI), electroencephalogram (EEG)). Provided that impaired performance on the DFT is truly characteristic of individuals with right frontal-lobe lesions relative to individuals with lesions of other cerebral areas, this assumption is justified. On the other hand, if poor performance on the DFT is associated with pathology in cerebral regions other than the right frontal lobe (e.g., right anterior temporal lesions), then interpretations of impaired DFT scores as an indication of right frontal-lobe dysfunction may not be warranted in every case.

Jones-Gotman and Milner (1977) reported that individuals with right frontal or fronto-central lesions produced significantly fewer novel designs than other lesion groups. However, they also found that individuals with anterior lesions of the right hemisphere, including temporal lesions, performed more poorly than controls while individuals with left-hemisphere lesions performed at a level equivalent to controls. These findings imply that while decreased productivity on the DFT appears to be most sensitive to right frontal-lobe damage, it may also be sensitive to right temporal-lobe lesions. If DFT performance is related to the presence of right temporal-lobe pathology, then it would be erroneous to infer right frontal-lobe dysfunction in each case of impaired DFT performance. If the DFT is sensitive to pathology of the right temporal lobe, one would expect these individuals to generate fewer novel designs relative to healthy controls. On the other hand, if individuals with right temporal lobe pathology perform at a level equivalent to controls, then the clinical use of the DFT as a measure specifically sensitive to right frontal-lobe dysfunction is not discredited.

To investigate these issues, normative data on DFT performance by healthy individuals with presumably intact cerebral functioning can be compared to the performance of individuals with circumscribed cerebral lesions. Patients who have undergone surgery for the removal of epileptogenic foci provide a good comparison group because their lesions are relatively circumscribed and the area of cerebral damage is known and documented. Another major purpose of this study, then, is to compare DFT performance in healthy controls, individuals with right or left lesions of the temporal lobe, and those with right or left lesions of the frontal lobe. The evidence to date suggests that decreased nonverbal fluency (i.e., poorer performance on the DFT) may be associated with right frontal or anterior temporal lobe damage, with poorer performance in individuals with right frontal lesions (Jones-Gotman & Milner, 1977). Conversely, individuals with left frontal or temporal lobe lesions may be less impaired on measures of nonverbal fluency. A comparison of DFT performance in these groups of individuals should help clarify the relationship between the cerebral areas associated with impaired performance on the DFT.

A final goal of this investigation is to determine whether there are differences in DFT performance before and after the surgical excision of cortical regions. In current clinical practice, the DFT is used as a predictive indicator of possible right frontal-lobe dysfunction. If DFT performance declines post-surgery for groups of individuals with temporal-lobe excisions, this deterioration would suggest that the DFT is sensitive to temporal-lobe dysfunction. On the other hand, several possibilities arise if DFT performances do not differ between pre-operative and post-operative assessments. One suggestion would be that functional systems served by the excised regions are not intimately involved in the cognitive skills required for effective DFT performance. Another possibility would be that excised regions did not affect DFT performance because the pathological processes present in those regions had already exerted an effect on DFT performance. A comparison of pre-operative and post-operative performances on the DFT will provide another source of information with respect to the sensitivity of the DFT to specific regions of cerebral dysfunction.

CHAPTER II

METHOD

Subjects

Healthy Control Group

77 unpaid volunteers were recruited as study participants for members of the healthy control group. 65 of these individuals were recruited from a pool of undergraduate students at the University of Windsor and participated for credit in an undergraduate psychology course. 12 participants were volunteers recruited from the community.

Participants were excluded from data analysis if they did not meet the following inclusion criteria: between the ages of 18 and 60, right-handed, with English as a first or main language, adequate intellectual ability, and no evidence of significant neurologic, systemic, or psychiatric illness. Participants with English as their first or main language were selected to ensure comprehension of instructions and familiarity with words used in the verbal tasks. Only right-handed subjects were recruited because they yield a reliable pattern of scores on functional cerebral lateralization tasks that tap typical hemispheric specializations (for review see Bryden, 1988). That is, speech is more likely to be represented in the left cerebral hemisphere in right-handed individuals (Bryden, 1988;

McCarthy & Warrington, 1990). Adequate intellectual ability was operationally defined as a WAIS-R Full Scale IQ estimate of greater than or equal to 80.

67 of the 77 participants met inclusion criteria; one of these subjects did not complete the entire test battery. Thus, test scores from 66 of the 77 healthy control participants were retained for data analysis. Of the 66 participants used in data analysis, 19 were male and 47 were female. The mean age of these individuals was 25.06 years (SD = 7.83) with a range from 19 to 56 years.

Patient Groups

Data were obtained from the files of 93 individuals who had undergone the surgical removal of epileptogenic foci as treatment for intractable epilepsy. 84 participants had undergone lobectomies or lesionectomies of the temporal regions. 44 of these individuals (22 male, 22 female) had received left-sided temporal lobectomies and 40 of these individuals (20 male, 20 female) had received right-sided temporal lobectomies. 9 individuals had undergone frontal resection surgery: 1 individual (male) had received a left-sided resection and 8 individuals (1 male, 7 female) had received a right-sided resection.

Files were selected from a data base at London Health Sciences Centre - University Campus (London, Ontario) which is a regional tertiary care hospital that includes a specialized epilepsy investigation and treatment centre. Individuals included in the study spoke English as a first or main language and had completed a pre- and/or post-surgical neuropsychological assessment that included the free condition of the DFT. Additional criteria to be met for inclusion in the study were: the absence of additional neurological history (e.g., head trauma), left-hemisphere dominance for speech, no evidence of psychiatric history or systemic illness, and adequate intellectual and language ability. Specific operational definitions of these criteria are presented in Table 1. Pre-operative neuropsychological assessment scores were available for all of these individuals, but post-operative neuropsychological assessment data were available for fewer cases. For the group of individuals with left temporal lobectomies, 30 patients had returned for post-operative assessment. Post-operative neuropsychological assessment scores were also available for 31 members of the group with right temporal lobectomies and for 2 members of the group with right frontal lobectomies.

Procedure

Each undergraduate participant was recruited through sign-up sheets distributed to sections of undergraduate psychology courses at the University of Windsor. Participants recruited from the community were approached

Table 1

Inclusion Criteria for Individuals with Frontal or Temporal Lobectomies

Criteria	Operational Definition
Absence of additional neurological history	No report of additional neurological disorders in the patient's file, and no evidence of additional pathology on pre-surgical diagnostic imaging (MRI, CT), EEG, or neurological testing.
Absence of psychiatric disorder	No reports of any psychiatric disorder in the patient's file.
Left cerebral hemisphere dominance for language	Pre-operative findings of: a right visual field advantage for letters presented tachistoscopically and a right ear advantage for words on dichotic listening task; or dysphasia following left but not right intracarotid injection of sodium amobarbital (Wada procedure).
Adequate intellectual and language ability	WAIS-R Verbal and Performance IQ scores greater than or equal to 80.

Note. Adapted from "The Validity of a Model-Based Procedure for Assessing Episodic Memory," by M. C. S. Harnadek, 1993, unpublished doctoral dissertation, University of Windsor, Ontario, p. 33.

individually and referred from other participants who had expressed interest in the study. Each testing session lasted approximately 90 minutes. Prior to administration of the test battery, participants provided background information such as age, programme of study, relevant medical history (questionnaire adapted from Spreen & Strauss, 1991; see Appendix A), and hand preference (questionnaire adapted from Reitan & Wolfson, 1993; see Appendix B). Each participant completed the following test battery: Perceptual Speed - Identical Forms (Thurstone & Jeffrey, 1987); Closure Speed - Gestalt Completion (Thurstone & Jeffrey, 1984); Design Fluency Test (free and fixed condition; Jones-Gotman & Milner, 1977); Wechsler Memory Scale-Revised (WMS-R) Visual Memory Span subtest (Wechsler, 1987); Controlled Oral Word Association (COWA; Spreen & Benton, 1969); Closure Flexibility - Concealed Figures (Thurstone & Jeffrey, 1984); Space Thinking - Flags (Thurstone & Jeffrey, 1956); Thurstone Word Fluency Test (Thurstone, 1938); the Vocabulary, Block Design, Arithmetic, Digit Symbol and Similarities subtests of the WAIS-R (Wechsler, 1981). Half of the control group completed the tests in the above order of administration; alternating participants received the reverse order of test administration to control for the effects of fatigue.

DFT protocols were scored according to the unpublished manuscript of scoring criteria specified by Jones-Gotman (no date). 44 DFT protocols were xeroxed and scored independently by three different raters: two neuropsychologists at London Health Sciences Centre and a graduate student in the clinical neuropsychology programme at the University of Windsor. All raters scored the DFT protocols based on their own independent understanding of the scoring criteria available in the unpublished manuscript (Jones-Gotman, no date).

Scores on the free condition of the DFT consisted of the number of nameable designs produced (errors), the number of perseverative designs produced (errors), a novel output score (total output minus nameable and perseverative errors), and percent perseveration (the percentage of perseverative errors relative to the total output minus all errors except perseverations).

Raw scores on the Thurstone Word Fluency Test were converted to standard scores based on normative data published by Heaton, Grant, and Matthews (1991) and adjusted for differences in age, gender, and education. Raw scores on the Controlled Oral Word Association test were converted to standard scores based on normative data collected by Ruff, Light, Parker, and Levin (in press). All other tests were scored and converted to standard scores in accordance with standardized scoring procedures outlined in the respective test manuals. The WAIS-R subtest scores were used to derive WAIS-R Full Scale IQ estimates from the tables published by Brooker and Cyr (1986) and based on the rationale proposed by Tellegen and Briggs (1967).

Statistical analyses were accomplished through the SPSS for Windows, version 6.0 statistical computer program. To assess inter-rater reliability, intraclass correlation coefficients were calculated for each of the scoring parameters of the DFT among the three raters. To assess the degree of linear relationship between DFT scores and each measure in the test battery, Pearson-*r* product-moment correlation coefficients were generated for each pair of variables. Standard multiple regression analyses were performed to determine which measures were best able to predict novel output score for both the free and the fixed condition. Multivariate analyses of variance were employed to compare DFT performance among participant groups and between pre-operative and post-operative assessments.

CHAPTER III

RESULTS

Data Screening

Prior to analysis, each variable was examined through various programs of the SPSS for Windows, version 6.0 statistical package for accuracy of data entry, missing values, and fit between their distributions and the assumptions of multivariate analysis. The variables were examined separately for the 66 healthy controls, 44 individuals with left temporal lobe pathology, 44 with right temporal lobe pathology, and 9 with frontal lobe pathology (8 right-sided, 1 left-sided).

One case in the control group had a single missing value on the WAIS-R Digit Symbol subtest. In accordance with one recommendation by Tabachnick and Fidell (1989) for dealing with missing data, the missing value for this case was replaced by the mean value for Digit Symbol in the control group.

Two cases in the healthy control group were considered univariate outliers because of their high Z scores on the percent perseveration variable for the free condition of the DFT. One case in the right temporal group (for both pre- and post-operative scores) and one in the left temporal group (post-operative scores only) were also considered univariate outliers because of their high Z scores on the percent perseveration variable. Furthermore, based on Cook's distance and Mahalanobis distance with $p < .001$, these two patient cases were identified as multivariate outliers in the right and left temporal groups respectively. For each of the four outliers, scores were modified on the percent perseveration variable to values just larger than the next most extreme score on that variable. This strategy reduced the influence of the outliers on data analyses, yet allowed the cases to be retained because they were considered to be part of the target population sampled (for rationale, see Tabachnick & Fidell, 1989).

Test Scores for Each Group

Means and standard deviations for the healthy control group for demographic variables and each of the measures in the test battery are presented in Table 2. Table 3 lists the means and standard deviations for each of the patient groups (right temporal, left temporal, right frontal, and left frontal) on demographic variables and WAIS-R Full Scale IQ (FSIQ) scores both pre- and post-operatively.

Table 4 summarizes mean DFT scores for the control group and each of the patient groups both pre- and post-operatively. The mean novel output score for the control group in the free condition was 13.12 (SD = 5.19), with a wide range from 2 to 28 designs. For the fixed condition, the mean novel output score for the control group was

Table 2

Means and Standard Deviations for Healthy Controls:

Demographic Variables and Test Scores (n = 66)

Variable	Mean (Standard Deviation)
Age	25.06 (7.83)
Education (Years Completed)	15.21 (1.60)
WAIS-R Full Scale IQ (Estimate)	100.85 (11.07)
WAIS-R Arithmetic (Age-Scaled Score)	9.64 (2.27)
WAIS-R Block Design (Age-Scaled Score)	11.27 (3.18)
WAIS-R Digit Symbol (Age-Scaled Score)	11.11 (2.05)
WAIS-R Similarities (Age-Scaled Score)	9.71 (1.89)
WAIS-R Vocabulary (Age-Scaled Score)	9.88 (2.20)
Controlled Oral Word Association (T-Score)	49.33 (9.19)
Thurstone Word Fluency (T-Score)	51.15 (9.55)
Closure Flexibility (T-Score)	52.39 (8.77)
Closure Speed (T-Score)	46.56 (8.54)
Perceptual Speed (T-Score)	54.62 (9.23)
Space Thinking - Flags (T-Score)	46.12 (7.63)
WMS-R Visual Memory Span (Raw Total)	16.38 (3.49)

Table 3

Means and Standard Deviations on Demographic Variables for Temporal and Frontal Lobectomy Groups

Group	Type of Assessment	Variable	Mean (SD)
Left Temporal	Pre-operative (n = 44)	Age (Years)	30 (9)
		Education	11 (2)
		WAIS-R FSIQ	87.28 (10.22)
	Post-operative (n = 30)	Age (Years)	32 (9)
		WAIS-R FSIQ	90.93 (13.10)
Right Temporal	Pre-operative (n = 40)	Age (Years)	31 (8)
		Education	12 (3)
		WAIS-R FSIQ	89.43 (12.02)
	Post-operative (n = 31)	Age (Years)	33 (8)
		WAIS-R FSIQ	89.86 (10.26)
Right Frontal	Pre-operative (n = 8)	Age (Years)	30 (8)
		Education	12 (1)
		WAIS-R FSIQ (Standard)	91.50 (14.12)
	Post-operative (n = 2)	Age (Years)	30 (10)
		WAIS-R FSIQ	92.01 (4.25)
Left Frontal	Pre-operative (n = 1)	Age (Years)	23 (0)
		Education	14 (0)
		WAIS-R FSIQ	97.00 (0.00)

Note. Education is measured in years completed.

Table 4

Means and Standard Deviations of Design Fluency Scores for Each Group

	Condition of the Design Fluency Test									
	Fixed Condition				Free Condition (Pre-Operative)			Free Condition (Post-Operative)		
Group	Novel Output	Name. Errors	Persev Errors	# Line Errors	Novel Output	Name. Errors	Persev Errors	Novel Output	Name. Errors	Persev Errors
NC N=66	15.15 (5.50)	0.24 (0.61)	7.14 (7.24)	2.85 (2.62)	13.12 (5.19)	0.14 (0.39)	7.92 (8.96)	N.A.	N.A.	N.A.
RT N1=40 N2=30	N.A.	N.A.	N.A.	N.A.	11.45 (5.60)	0.65 (1.17)	5.00 (6.00)	10.74 (6.57)	1.00 (2.00)	6.00 (9.00)
LT N1=44 N2=31	N.A.	N.A.	N.A.	N.A.	10.52 (7.26)	0.86 (1.55)	5.00 (8.00)	11.10 (5.95)	1.00 (1.00)	6.00 (7.00)
RF N1=8 N2=2	N.A.	N.A.	N.A.	N.A.	9.25 (3.62)	1.13 (1.73)	3.00 (3.00)	5.50 (0.71)	0.00 (0.00)	2.00 (3.00)
LF N1=1 N2=0	N.A.	N.A.	N.A.	N.A.	5.00 (0.00)	0.00 (0.00)	2.00 (0.00)	N.A.	N.A.	N.A.

Note. Name. Errors = Nameable errors; # Line Errors = Errors with the incorrect number of lines; Persev Errors = Perseverative errors; NC = Normal control group; RT = Right temporal group; LT = Left temporal group; RF = Right frontal group; LF = Left frontal group; N1 = Number of participants in the pre-operative group; N2 = Number of participants in the post-operative group.

15.15 (SD = 5.5), also with a large range from 6 to 30 designs. The pre-operative mean novel output score (free condition) was 11.45 (SD = 5.60) for the right temporal group and 10.52 (SD = 7.26) for the left temporal group. Pre-operative mean novel output scores for the frontal groups were lower (right = 9.25; left = 5.00), but these groups were comprised of only a limited number of cases. Post-operative novel output scores were in the same range as pre-operative scores for the left temporal group (Mean = 11.10; SD = 5.95), slightly lower for the right temporal group (Mean = 10.74; SD = 6.57), and lower for the right frontal group (Mean = 5.50; SD = 0.71). No post-operative scores were available for the one individual who had received a left frontal resection.

Relationship Between DFT Scores and Battery Measures in the Healthy Control Group

A correlation matrix of Pearson-*r* product-moment correlation coefficients was generated for each pair of variables to assess the degree of relatedness between DFT scores and each measure in the test battery. This matrix was produced through the CORRELATE program of the SPSS for Windows, version 6.0 statistical package. Significant correlations between DFT scores and measures in the test battery are presented in Table 5. Significant correlations between DFT scoring parameters for both the free and the

Table 5

Significant Pearson-r Correlation Coefficients for DFT Scores with Scores on Measures in the Test Battery

Test Measure	Pearson-r
FREE CONDITION:	
Correlations with Novel Output Score	
Controlled Oral Word Association	.31
Correlations with Percent Perseveration Score	
Age	.25
WMS-R Visual Memory Span Subtest	-.31
FIXED CONDITION:	
Correlations with Novel Output Score	
Perceptual Speed	.34
Closure Speed	.32
Closure Flexibility	.30
Controlled Oral Word Association	.30
Space Thinking - Flags	.27
Correlations with Percent Perseveration Score	
WMS-R Visual Memory Span Subtest	-.26

fixed condition are displayed in Table 6.

Novel output scores for the free condition of the DFT were significantly correlated with scores on Controlled Oral Word Association ($r = .31$, $p < .05$) but not with any of the other measures in the test battery. The percent perseveration score for the free condition was correlated with age ($r = .25$, $p <. 05$), with higher percentages of perseverative output associated with older ages. Novel output score on the fixed condition of the DFT was

Table 6

Significant Pearson-r Correlation Coefficients for DFT Scores in the Free and Fixed Conditions

DFT Score	Pearson-r
Correlations between Scores on the Free and Fixed Conditions	
Percent Perseveration Score	.81
Number of Perseverative Errors	.66
Novel Output Score	.59
Number of Nameable Errors	.38
Correlations with Percent Perseveration Score for the Free Condition	
Number of Perseverative Errors	
In the Free Condition	.69
In the Fixed Condition	.52
Fixed Condition Errors with the Incorrect Number of Lines	.43
Novel Output Score in the Free Condition	-.35
Correlations with Percent Perseveration Score for the Fixed Condition	
Number of Perseverative Errors	
In the Fixed Condition	.82
In the Free Condition	.60
Fixed Condition Errors with the Incorrect Number of Lines	.43
Novel Output Score in the Fixed Condition	-.24
Correlations with Fixed Condition Errors with the Incorrect Number of Lines	
Number of Perseverative Errors	
In the Free Condition	.43
In the Fixed Condition	.28
Novel Output Score in the Fixed Condition	-.26
Correlations with Fixed Condition Nameable Errors	
Novel Output Score in the Fixed Condition	.32

Note. Correlations reported in earlier sections of the table are not reported in subsequent sections.

significantly correlated with Perceptual Speed (r = .34, p < .01), Closure Speed (r = .32, p < .01), Closure Flexibility (r = .30, p < .05), Controlled Oral Word Association (r = .30, p < .05) and Space Thinking - Flags (r = .27, p < .05). Percent perseveration scores in both conditions were negatively correlated with scores on the WMS-R Visual Memory Span subtest (Free: r = -.31, p < .01; Fixed: r = -.26, p < .05) with higher percentages of perseverative output associated with poorer scores on this test.

Although the correlations between DFT scores and these variables were significant, the size of each of the correlations was relatively small. The highest correlation was between novel output score on the fixed condition of the DFT and Perceptual Speed (r = .34, p < .01), which accounted for only 11.56% of the variance. In other words, a large portion of the variance in novel output scores for both conditions of the DFT was not accounted for by any of the correlations with test measures.

As expected, novel output scores for the free and fixed conditions of the DFT were significantly correlated (r = .59, p < .001). The percent perseveration scores for both conditions were significantly correlated (r = .81, p < .001) as were the number of perseverative errors (r = .66, p < .001) and nameable errors (r = .38, p < .01). The number of perseverative errors committed in the free condition was significantly correlated with percent perseveration for both conditions (Free: $r = .69$, $p < .001$; Fixed: $r = .60$, $p < .001$). Similarly, the number of perseverative errors in the fixed condition was correlated with percent perseveration in both conditions (Free: $r = .52$, $p < .001$; Fixed: $r = .82$, $p < .001$).

Novel output score for the free condition was negatively correlated with percent perseveration ($r = -.35$, $p < .01$) with higher novel output scores associated with smaller proportions of perseverative errors. For the fixed condition, novel output score was negatively correlated with percent perseverative error ($r = -.24$, $p < .05$) and number of designs with the incorrect number of lines ($r = -.26$, $p < .05$). Paradoxically, novel output score in this condition was positively correlated with nameable errors ($r = .32$, $p < .01$), which may be related to a floor effect on nameable errors in the healthy control group. Fixed condition errors with the incorrect number of lines were correlated with the number of perseverative errors for both conditions (Free: $r = .43$, $p < .001$; Fixed: $r = .28$, $p < .05$) and percent perseveration in both conditions (Free: $r = .43$, $p < .001$; Fixed: $r = .43$, $p < .001$). Generally speaking, a higher incidence of one type of error was associated with a greater frequency of other error types and with lower novel output scores.

Multiple regression analyses were performed to assess which measures were best able to predict DFT performance.

In order to reduce the number of predictor variables in the regression equation, composite variables were created based upon the results of a factor analysis and reliability analyses of those factors. Principal components extraction with varimax rotation, through the FACTOR program of the SPSS for Windows version 6.0 statistical package, extracted five factors. A cutoff of .75 as a factor loading was used to determine which variables would contribute to the composite; a factor loading of .75 would account for 56.25% of the variance.

For the first factor, Space Thinking - Flags loaded .82, Closure Flexibility loaded .81, and the WAIS-R Block Design subtest loaded .77. A composite score was formed from the weighted sum of standard scores on these measures where the factor loadings served as weights for each variable. This first factor was interpreted as a measure sensitive to visuospatial skills. For a second factor, Controlled Oral Word Association loaded .88 and the Thurstone Word Fluency test loaded .87. The T-scores of these variables were then collapsed into a weighted sum. This composite variable was interpreted to reflect verbal fluency skills and/or the ability to produce output in a limited amount of time under particular specifications. The WAIS-R Digit Symbol subtest loaded .84 on a third factor; no other measures generated factor loadings large enough to meet the cutoff for inclusion in this composite. The highest loadings for the fourth factor were the WAIS-R Vocabulary subtest at .86 and the WAIS-R Full Scale IQ estimate at .81. This factor was interpreted as a measure sensitive to psychometric intelligence. A fifth and final factor was related to the Design Fluency tests, with free condition scores loading .85 and fixed condition scores loading .84. This fifth factor was not included in the regression analyses because scores on each condition of the DFT served as the criteria to be predicted.

A series of Cronbach's alpha tests was used to assess the reliability of each composite factor. These analyses were performed through the RELIABILITY program of the SPSS for Windows, version 6.0 statistical package. Factors with low reliability coefficients were not retained for inclusion in the regression analyses. In addition, if any one variable significantly decreased the reliability of a composite, then that variable was excluded from the composite score.

The factor related to visuospatial skills (i.e., the composite of scores on Space Thinking - Flags, Closure Flexibility, and WAIS-R Block Design) obtained a reliability coefficient of .87. The removal of each of these measures from the composite did not decrease the reliability coefficient significantly; thus, all three variables were retained in the spatial composite. Scores on the WMS-R Visual Memory Span subtest also loaded fairly highly on this first factor (.72), but because standard scores were not available on this measure, it was entered separately into the regression equations.

The factor related to verbal fluency skills and/or production (i.e., the composite of scores on Controlled Oral Word Association and the Thurstone Word Fluency Test) obtained a reliability coefficient of .81. Both of these variables were retained in the composite.

For the third factor with a high loading for the WAIS-R Digit Symbol subtest, two variables with factor loadings greater than .50 (Perceptual Speed loaded .65; Closure Speed loaded .52) were combined with Digit Symbol scores to assess reliability. Out of the possible combinations of these variables, the highest reliability coefficient obtained was .28. For this reason, Digit Symbol scores were entered separately into the regression equations as an element related to psychomotor speed. Closure Speed and Perceptual Speed were also included as separate predictors in the regression equations because of the poor reliability coefficients produced for possible composite variables in which they were included.

Two standard multiple regression analyses were performed through the REGRESSION program of the SPSS for Windows version 6.0 statistical package to determine which variables best predicted DFT performance. Novel output scores on the free condition and then the fixed condition of the DFT were the criteria and the following variables served as predictors: the composite related to verbal fluency/production, the composite related to visuospatial skills, WAIS-R Digit Symbol subtest score, Perceptual Speed score, Closure Speed score, WMS-R Visual Memory Span subtest score, and WAIS-R FSIQ estimate.

Novel output score on the free condition of the DFT served as the criterion for the first regression equation. For this analysis, Table 7 displays the multiple correlation (*R*), R^2, and adjusted R^2, the unstandardized regression coefficients (*B*) and intercept, and the standardized regression coefficients (Beta). The multiple correlation (*R*) was not significantly different from zero ($R = .38$; $F(7, 58) = 1.40$, $p > .05$). In other words, there was no evidence of a significant relationship between the predictors and novel output score on the free condition of the DFT.

For the second regression equation, novel output score on the fixed condition served as the criterion and the predictors were the same ones used in the previous analysis. Values of *R*, R^2, and adjusted R^2, the unstandardized regression coefficients (*B*) and intercept, and the standardized regression coefficients (Beta) are presented in Table 8. The multiple correlation (*R*) was significantly different from zero ($R = .47$; $F(7, 58) = 2.40$, $p < .05$). These results indicate that variables in the regression

Table 7

Summary Statistics for Standard Multiple Regression with Novel Output Score in the Free Condition of the DFT as Criterion

Variable	B	Beta
Verbal Fluency Composite	0.045	0.130
Visuospatial Composite	-0.064	-0.177
WAIS-R Digit Symbol	0.435	0.172
Perceptual Speed	0.055	0.098
Closure Speed	0.085	0.139
WMS-R Visual Memory Span	0.323	0.218
WAIS-R FSIQ (estimate)	-0.035	-0.075
		Multiple R = .38 Multiple R^2 = .14 Adjusted R^2 = .04

equation were able to predict novel output score on the fixed condition of the DFT to a significant degree. For tests of individual predictors, none of the variables yielded a significant value. Although the combined group of variables was able to predict novel output score for the fixed condition of the DFT, no individual variable appeared to make a unique contribution to the regression equation. The combined group of predictors was able to account for approximately 22% of the variation in novel output score for

Table 8

Summary Statistics for Standard Multiple Regression with Novel Output Score in the Fixed Condition of the DFT as Criterion

Variable	B	Beta
Verbal Fluency Composite	0.069	0.189
Visuospatial Composite	0.089	0.230
WAIS-R Digit Symbol	-0.149	-0.056
Perceptual Speed	0.104	0.174
Closure Speed	0.137	0.213
WMS-R Visual Memory Span	0.063	0.040
WAIS-R FSIQ (estimate)	-0.109	-0.220
		Multiple $R = .47^*$ Multiple $R^2 = .22$ Adjusted $R^2 = .13$

* $p < .05$

the fixed condition.

Inter-rater Reliability of DFT Scores

Scores by the three raters for 44 DFT protocols of healthy control participants were analyzed for inter-rater reliability and consistency. Rater means and standard deviations for each of the scoring parameters in the free and fixed condition of the DFT are summarized in Table 9. Intraclass correlation coefficients (ICC) were calculated

Table 9

Means for Each Rater on DFT Scores for 44 Protocols

DFT Condition and Type of Score	Rater 1	Rater 2	Rater 3
Free Condition			
Total Number of Designs	22.27 (11.94)	22.34 (12.02)	22.27 (11.89)
Novel Output Score	13.11 (6.17)	16.80 (8.43)	14.59 (7.97)
Perseverative Errors	9.05 (10.48)	5.43 (8.18)	7.43 (9.31)
Nameable Errors	0.11 (0.39)	0.11 (0.39)	0.25 (0.61)
Fixed Condition			
Total Number of Designs	26.27 (12.04)	26.30 (12.03)	26.25 (12.07)
Novel Output Score	15.82 (6.52)	17.45 (6.25)	19.05 (8.17)
Perseverative Errors	7.84 (8.35)	5.82 (5.80)	5.57 (5.36)
Nameable Errors	0.18 (0.66)	0.32 (1.12)	0.41 (1.26)
Incorrect Number of Lines	2.45 (2.60)	2.45 (3.58)	1.23 (2.03)

for indices of both inter-rater agreement and consistency. The ICC was used as a measure of inter-rater reliability because it takes into account more than one source of variance (Sechrest, 1984). Berk (1979) reviews the advantages of the ICC as a measure of reliability. The formulae used for calculation of the ICC were based on an analysis of variance model and reported by Shrout and Fleiss

(1979). For inter-rater agreement, the raters were considered to be random effects, and the following formula was used:

$$ICC(2,1) = \frac{BMS - EMS}{BMS + (k - 1)EMS + k(JMS - EMS)/n}$$

where BMS represented the mean square value between cases, EMS represented the residual (i.e., error) mean square value, k was equal to the number of raters, and JMS represented the mean square value between raters. For rater consistency, the raters were considered to be fixed effects, and ICCs were calculated in accordance with the following formula given by Shrout and Fleiss (1979):

$$ICC(3,1) = \frac{BMS - EMS}{BMS + (k - 1)EMS}$$

with BMS, EMS and k defined as in the previous equation. The analysis of variance mean square values for each of the scoring parameters for each condition of the DFT were generated through the RELIABILITY program of the SPSS for Windows, version 6.0 statistical package.

ICC values of inter-rater agreement and consistency for each scoring parameter of both the free and the fixed condition are listed in Table 10. ICC values were evaluated for clinical significance according to criteria set out by Cicchetti and Sparrow (1981). These authors suggested that correlation coefficients of .75 or greater represented excellent inter-rater reliability, and coefficients from .60 to .74 represented good inter-rater agreement. ICC values from .40 to .59 were interpreted as fair inter-rater reliability, and values less than .40 were indicative of poor inter-rater agreement. These levels of predictive significance suggested by Cicchetti and Sparrow (1981) are presented for each ICC in Table 10.

In all cases, indices of rater consistency were equal to or greater than those of inter-rater agreement. The highest ICC values were produced for total number of designs in both the free and the fixed conditions, where rater agreement and consistency rounded to perfect values. Novel output scores and perseverative errors for both conditions generated good to excellent reliability coefficients. Designs with the incorrect number of lines in the fixed condition yielded fair reliability indices. Nameable errors yielded fair reliability coefficients for the free condition but poor reliability indices for the fixed condition.

Poorer reliability for nameable errors and designs with the incorrect number of lines is likely related to the infrequent occurrence of these errors. When errors are less frequent, differences in scores among raters exert a greater effect because the means and standard deviations among raters become correlated (Woodard et al., 1992). Because the ICC is based on an analysis of variance model, less variance among DFT protocols relative to greater variance among raters yields a lower reliability coefficient. This situation applies to the case of errors that are less

Table 10

Intraclass Correlation Coefficients for DFT Scores Among 3 Raters

DFT Condition and Type of Score	Inter-rater Agreement	Inter-rater Consistency	Clinical Significance
Free Condition			
Total Designs	1.00	1.00	Excellent
Novel Output	0.78	0.83	Excellent
Perseverative Errors	0.85	0.88	Excellent
Nameable Errors	0.44	0.44	Fair
Fixed Condition			
Total Designs	1.00	1.00	Excellent
Novel Output	0.72	0.75	Good; Excellent
Perseverative Errors	0.72	0.74	Good
Nameable Errors	0.33	0.33	Poor
Incorrect Number of Lines	0.46	0.49	Fair

frequent. Poorer reliability indices for these errors are not completely attributable to their infrequent occurrence, however, because nameable errors in the fixed condition were less reliable than those in the free condition and yet occurred more often.

Differences between Healthy Controls and Patient Groups

Group differences on demographic variables.

To assess possible differences between participant groups on demographic variables prior to analysis for differences in DFT performance, these variables were analyzed through the ONEWAY program of the SPSS for Windows, version 6.0 statistical package. The groups of individuals with frontal lobe pathology were not included in the analyses due to the small sample size (n = 9). Separate univariate analyses of variance were performed with group membership as the independent variable and age, education, and WAIS-R FSIQ (or estimate) as dependent variables.

The analyses revealed significant differences between the groups for age (F (2, 147) = 7.53, p < .001), education (F (2, 147) = 58.74, p < .001), and WAIS-R FSIQ (F (2, 146) = 23.86, p < .001). A Bonferroni pair-wise comparison of means, which adjusts for multiple comparisons, was used to determine the groups that differed from one another on these variables. These comparisons revealed that the healthy control group was significantly younger (p < .05) and more educated (p < .05) than the two temporal lobectomy groups. The healthy control group also had significantly greater WAIS-R FSIQ scores than both the left and right temporal lobectomy groups (p < .05). The left-sided and right-sided temporal lobectomy groups did not differ significantly on age, education, or WAIS-R FSIQ scores.

Differences among DFT performances.

Scores on the free condition of the DFT were compared between healthy controls and patient groups, using pre-operative assessment scores. A between-subjects multivariate analysis of covariance was performed on four dependent variables associated with the DFT: novel output score, nameable errors, perseverative errors, and percent perseveration. Adjustment was made for three covariates: years of completed education, age, and WAIS-R FSIQ. The independent variable was group membership with three levels: healthy controls (n = 66), individuals with right temporal-lobe pathology (n = 40), and those with left temporal-lobe pathology (n = 42). The group of patients with frontal lobe pathology was not included in the analysis because of the small sample size (n = 9).

The MANOVA program of the SPSS for Windows, version 6.0 statistical package was used for the analyses with the sequential adjustment for nonorthogonality due to unequal cell sizes. With the use of Pillais' criterion, the combined dependent variables were significantly related to group membership (approximate F (8, 280) = 2.16, $p < .05$), but not to the combined covariates (approximate F (12, 423) = 1.21, $p > .05$). Effects of group membership on the dependent variables after adjustment for covariates were investigated in univariate and stepdown analyses. For the stepdown analysis, novel output score was given the highest priority, percent perseveration second priority (i.e., adjustment was made for novel output score in addition to the three covariates), perseverative errors third priority (i.e., adjustment was made for novel output score, percent perseveration, and the three covariates) and nameable errors fourth priority (i.e., adjustment was made for novel output score, percent perseveration, perseverative errors, and the three covariates). Results of this analysis are summarized in Table 11. An experimentwise error rate of 5% for each effect was attained by apportioning alpha according to the values shown in Table 11.

After adjusting for differences on the covariates, the number of nameable errors made a significant contribution to the linear equation of dependent variables that discriminated best between the groups (stepdown F (2, 139) = 6.07, $p < .01$). Post-hoc comparisons of group means with Tukey's HSD test showed that both individuals in the group with left temporal-lobe pathology (q (3, 139) = 0.82, $p < .05$) and those with right temporal-lobe pathology (q (3, 139) = 1.03, $p < .05$) made significantly more nameable errors than members of the healthy control group. These results suggest that when adjustments are made for differences in education level, age, and psychometric intelligence, individuals with pathology of either the left or right temporal lobes commit more nameable errors than healthy controls. It is important to note, however, that

Table 11

Univariate and Stepdown F Tests for Significant Differences in DFT Scores Among Controls and Temporal Lobectomy Groups

Effect	DV	F^1	df	F^2	df	Alpha
Covariates	Novel Output	1.27	3,142	1.27	3,142	.01
	Percent Persev.	0.86	3,142	0.89	3,141	.01
	Persev. Errors	0.81	3,142	1.69	3,140	.01
	Nameable Errors	1.06	3,142	1.01	3,139	.01
Group	Novel Output	0.15	2,142	0.15	2,142	.01
	Percent Persev.	1.37	2,142	1.37	2,141	.01
	Persev. Errors	0.75	2,142	1.19	2,140	.01
	Nameable Errors	6.91	2,142	6.07*	2,139	.01

Note. Persev. Errors = Number of perseverative errors; Percent Persev. = Percent perseveration score; F^1 = Univariate F; F^2 = Stepdown F.

* $p < .01$

the number of nameable errors made by the patient groups was still low, with less than one nameable error (left temporals: adjusted mean = 0.96; right temporals: adjusted mean = 0.75).

Pre- versus Post-Operative DFT Performance

A mixed design multivariate analysis of variance of between and within-subjects factors was performed with the four scores on the DFT as dependent measures (i.e., novel output score, percent perseveration, perseverative errors, and nameable errors). The between-subjects independent variable was group membership in either the right temporal group or the left temporal group. The within-subjects factor was type of assessment related to one of two occasions: pre-operative versus post-operative. Individuals in the frontal groups were not included in this analysis because only two participants returned for post-operative assessments.

The SPSS for Windows version 6.0 MANOVA program was used for the analysis with the sequential adjustment for nonorthogonality due to unequal cell sizes. The results of this analysis are summarized in Table 12. With the use of Pillais' criterion, the combined dependent variables were not significantly affected by group membership (F (4, 56) = 1.14, $p > .05$), type of assessment (F (4, 56) = .55, $p > .05$), or the interaction between group membership and type of assessment (F (4, 56) = .72, $p > .05$). That is, DFT scores did not differ significantly from pre-operative to post-operative neuropsychological assessment in these groups of patients. In addition, DFT scores did not differ significantly between groups of individuals with left and

Table 12

Mixed Design MANOVA: Main Effects and Interactions for Group Membership and Type of Assessment

Effect	Multivariate F df (4,56)	DV	Univariate F df (1,59)
Group	1.14 n.s.	Novel Output Score	0.01 n.s.
		% Perseveration	0.0004 n.s.
		Perseverative Errors	0.007 n.s.
		Nameable Errors	1.57 n.s.
Type of Assessment	0.55 n.s.	Novel Output Score	0.00 n.s.
		% Perseveration	1.31 n.s.
		Perseverative Errors	0.04 n.s.
		Nameable Errors	0.34 n.s.
Group by Type of Assessment	0.72 n.s.	Novel Output Score	0.04 n.s.
		% Perseveration	1.39 n.s.
		Perseverative Errors	1.10 n.s.
		Nameable Errors	0.79 n.s.

right temporal-lobe pathology on pre- versus post-operative occasions.

CHAPTER IV

DISCUSSION

Normative Data Collection and Psychometric Characteristics of the DFT

The normative data collected on the sample of healthy control participants in this study yielded a mean output of novel designs in the free condition that was slightly lower than those reported in previous studies. In her unpublished manuscript, Jones-Gotman reported a mean novel output score of 15.9 for the healthy controls. Varney et al. (1996) also reported a similar mean score of 16.1 novel designs for their control group. In contrast, the mean number of novel designs produced by the participants in this study was 13.1, although it did vary somewhat among raters. Woodard et al. (1992) also reported mean novel output scores from two different raters that varied to some extent (i.e., 20.3; 13.7).

The slight differences between the mean novel output scores generated in this study and those reported previously may be related to the population from which the control groups were drawn. Jones-Gotman and Milner (1977) and Varney et al. (1996) did not define clearly the sample composition of their control groups and the approach used to recruit these individuals. In contrast, the majority of participants in this study are known to be from an undergraduate student population with an average level of psychometric intelligence. This sample may differ from the composition of normal control groups used in previous studies in a manner that accounts for the slight discrepancy in novel output score for the free condition of the DFT. Differences in mean novel output score across normative samples may also be related to the inconsistent use of the unpublished scoring criteria (Jones-Gotman, no date) and some variations in scoring leniency among raters.

The mean novel output score for the fixed condition in the present sample was slightly higher than that of the free condition (15.2). For both the free and the fixed condition, there was considerable variation among individual performances on novel output score, a finding that is consistent with previous research (e.g., Jones-Gotman & Milner, 1977; Varney et al., 1996; Woodard et al., 1992). Novel output scores for both the free and the fixed condition were normally distributed. This psychometric characteristic adds to the credibility of the DFT as a measure that can be standardized for use in clinical assessment.

The inter-rater reliability of scoring parameters for both conditions of the DFT generally fell within the good to excellent range, according to criteria given by Cicchetti and Sparrow (1981). Inter-rater agreement and consistency was good to excellent for the total number of designs, novel output score, and perseverative errors for both the free and the fixed condition. Fixed condition errors with the incorrect number of lines yielded fair reliability indices. Reliability indices for nameable errors were fair for the free condition, but poor for the fixed condition. To some extent, the poorer reliability indices for nameable errors and designs with the incorrect number of lines may be due to the greater impact of rater differences on less frequent types of errors (Woodard et al., 1992). However, there may also be larger differences among raters on some of these scoring parameters, particularly for nameable errors in the fixed condition.

In general, inter-rater reliability indices for DFT scoring parameters were greater than those reported by Woodard et al. (1992). There were two exceptions to this tendency: nameable errors in the free condition and designs with the incorrect number of lines in the fixed condition yielded good reliability coefficients in the prior study as opposed to the fair coefficients produced by this investigation. In agreement with the present findings, Woodard et al. (1992) also reported the poorest reliability coefficients for nameable errors in the fixed condition. One key difference between the two studies is that in the present study inter-rater agreement for perseverative errors was considerably greater than in the Woodard et al. (1992) study (Free Condition: .85 vs. .57; Fixed Condition: .72 vs.

.41).

In part, the greater coefficients of inter-rater agreement and consistency for perseverative errors reported in the present study may be due to the exposure of two of the raters to the DFT in a clinical setting. The two raters in the Woodard et al. (1992) study had no experience with the DFT prior to the research in question. Training in the use of the DFT and clinical exposure to impaired DFT performances may both augment inter-rater reliability. Differences may also be attributable to the use of the unpublished scoring guidelines available from Jones-Gotman (no date). Although a personal communication with Jones-Gotman by the authors of the Woodard et al. (1992) study suggests they may have had access to this manuscript, it is unclear whether it was used. The specific criteria and examples provided in the unpublished manuscript likely increase the chances that raters will score DFT protocols in a similar manner.

The ICC values for rater consistency generated by the present study suggest that despite some differences in scoring leniency, in general raters tended to rank DFT protocols in the same order in terms of error frequency. The fair rater consistency for errors with the incorrect number of lines indicates that there was more variability in rank order for this scoring parameter. Lower rater consistency for nameable errors in the free condition and poor rater consistency for these errors in the fixed condition implies that there were considerable differences on this scoring parameter. Particularly with respect to nameable errors, training and practice in appropriate approaches to scoring the DFT may be of benefit prior to clinical use of the test.

Despite the variability among raters for nameable errors in both conditions and errors with the incorrect number of lines in the fixed condition, the DFT appears to have fairly good inter-rater reliability. Scores with poorer reliability indices may be partially accounted for by the infrequent occurrence of these error types. This likelihood, combined with good to excellent reliability indices for perseverative errors, novel output scores, and the total number of designs in both conditions suggest that the test is appropriate for clinical use. Although the DFT does not have excellent inter-rater reliability for all scoring parameters, training sessions on how to administer and score the DFT in addition to access to normative data based on more than one rater could both contribute to a more standardized use of the DFT.

The normative data yielded by this study represent an important contribution to the psychometric information available on the DFT. The healthy control group in this investigation is one of the largest samples for which normative data on both the free and the fixed condition of the DFT have been collected. In addition, participants in this study were considerably younger than those in the only other study with a large normative sample for both conditions of the DFT (i.e., Woodard et al., 1992) and thus, it provides replication as well as an extension of earlier research. Consequently, the normative data and psychometric information on the DFT collected in the course of this study constitutes an original contribution to the field.

Cognitive Skills Related to DFT Performance

The small size of the correlations between novel output scores on either condition of the DFT and scores on other measures in the test battery suggested that no single predictor would account for the bulk of variance in novel output scores. This finding was confirmed by the results of the regression analyses, where no individual variable was a significant predictor of novel output score for either the free or the fixed condition.

For the free condition of the DFT, the combined linear aggregate of the variables in the regression equation was not able to predict novel output score to a significant degree. This finding suggests that novel output score on the free condition is related to a large extent to factors other than verbal fluency/production, visuospatial skills, psychomotor speed, visual attention, and level of psychometric intelligence. The finding that level of psychometric intelligence was not significantly associated with DFT performance in healthy controls is consistent with evidence of no association between the two measures in individuals with closed head injury (Varney et al., 1996).

On the other hand, the combined linear aggregate of the variables in the regression equation was able to predict novel output score for the fixed condition of the DFT. As mentioned above, none of the individual variables contributed significantly to the prediction of novel output score. Because no single variable was able to predict novel output score for the fixed condition to a significant degree, this measure does not appear to be related solely to verbal fluency/production, visuospatial skills, psychomotor speed, visual attention, or level of psychometric intelligence.

It is noteworthy that the verbal fluency composite was not able to predict novel output score on either the free or the fixed condition of the DFT. Although intuitively the DFT and verbal fluency tasks appear to share similar requirements, the inability of verbal fluency scores to predict DFT scores suggests that the DFT is not a direct nonverbal analogue of verbal fluency tasks, although that was the original intention (Jones-Gotman & Milner, 1977). Performance on one of the verbal fluency measures, Controlled Oral Word Association, was significantly associated with DFT performance, and was the task most highly correlated with DFT performance in the free condition. Despite this evidence of a relationship between the two measures, the small magnitude of this correlation and the inability of the verbal fluency composite to predict DFT scores implies constructs other than verbal fluency are more strongly related to performance on the DFT.

The results of this study suggest that novel output score in both the free and the fixed condition of the DFT are related largely to some element(s) untapped by the measures used in this study. However, skills tapped by these measures do appear to have some influence on DFT performance in the fixed condition since the combined aggregate of the variables was able to significantly predict novel output score. It may be that individual differences result in several different combinations of varying degrees of these cognitive skills, or subsets of these skills, which can lead to successful performance on the DFT. In other words, although no one individual factor investigated in this study was able to predict DFT performance, various combinations of aptitudes for verbal fluency/production, visuospatial skills, visual attention, psychomotor speed, and level of psychometric intelligence appear to be related to the number of novel designs produced in the fixed condition.

The differences between the free condition and the fixed condition are emphasized by the finding that the combined predictors were able to predict novel output score for the fixed condition but not for the free condition. This outcome shows clearly that the restrictions placed on output in the fixed condition change the nature of the task somewhat and the cognitive skills that are able to predict performance. Within the restrictions imposed by the fixed condition, the prediction of novel output scores may become more related to skills that are less important in the prediction of free condition scores (i.e., verbal fluency/production, visuospatial skills, psychomotor speed, visual attention, and psychometric intelligence). It is unclear whether cognitive skills that are more directly related to DFT performance in the free condition would be related to fixed-condition performance to the same degree, given that the skills investigated in this study can predict novel output in one condition but not the other. A further indication of the differences between cognitive skills tapped by the two conditions is provided by evidence that performance in each condition is affected differently by lesions of the same cerebral regions (Jones-Gotman and Milner, 1977).

The observation that no one individual variable was able to predict DFT performance suggests that the task is sensitive to cognitive skills, or sets of skills, which are unique, perhaps even orthogonal, to those investigated in the present study. One possibility is that a large degree of variability in the novel output score of the two conditions may be accounted for by a factor such as divergent thought, or "creativity", which would enable an individual to produce as many unique designs as possible. The concept of divergent thought involves the generation and exploration of many different solutions to a problem, whereas convergent thought attempts to discover one unique, correct solution to the problem. In the fixed condition, certain combinations of cognitive skills such as visuospatial skills and verbal fluency/production appear to make a smaller but significant contribution to performance. These skills may become more predictive in the fixed condition because output is limited to the possible combinations of four lines.

Performance on measures sensitive to divergent thinking may more accurately reflect the cognitive skills involved in performance on the DFT. A measure such as a possible jobs test may involve components of both verbal fluency skills and divergent thought. In this type of task, participants are presented with a drawing of an object and are required to generate as many possible uses for it as they can within a limited amount of time. Individuals with frontal lobe tumours perform more poorly on this task than individuals with intracranial tumours which spare the frontal lobes (Correa & Butler, 1996). It would be interesting to investigate whether performance on similar measures is related to DFT performance and if so, whether the relationship differs between the free and the fixed conditions of the test.

Comparison of DFT Performance in Healthy Controls and Individuals with Temporal-Lobe Pathology

The weighted aggregate of four scores related to DFT performance in the free condition (i.e., novel output score, percent perseveration, number of perseverative errors, and number of nameable errors) differed significantly between the healthy control group and groups of individuals with right-sided and left-sided pathology of the temporal lobe. Individual tests of each type of error indicated that this difference between groups was due largely to a greater number of nameable errors committed by both the left and right temporal-lobe groups relative to the control subjects.

Jones-Gotman and Milner (1977) reported that individuals with lesions of the right temporal lobe or the right frontal lobe accounted for 75% of their participants who made more than one nameable error. In their study, the incidence of individuals who made more than one nameable error was significantly higher in the right temporal-lobe group than in the left temporal-lobe group and than in the control group. These researchers chose one nameable error as a cutoff because participants who exceeded this cutoff continued to make nameable errors despite a warning given after the first error occurred.

The results of the present study are consistent with those of the original article (Jones-Gotman & Milner, 1977) in that relatively few nameable errors were made by all groups and individuals with pathology of the right temporal lobe made more nameable errors than controls. However, individuals with pathology of the left temporal lobe also made more nameable errors than the control group, a finding that has not been reported previously.

It is not surprising that individuals with cerebral dysfunction commit more errors on a task than healthy controls, but it is unclear why nameable errors in particular are more likely to occur in these patient groups. A higher occurrence of nameable errors in individuals with cerebral dysfunction may be related to more general cognitive impairment relative to controls which does not represent a specific deficit. On the other hand, committing more nameable errors may reflect a specific type of difficulty these patients have with the abstract nature of the task. It is difficult to make conclusive judgements about the significance of this finding and its specificity to groups of individuals with temporal-lobe pathology because data from the frontal groups were not statistically analyzed.

In any case, it is important to keep in mind that the mean number of nameable errors committed by each group, including patient groups, was quite low (i.e., less than one nameable error). Moreover, the lower coefficients of inter-rater agreement and consistency for nameable errors relative to other error types suggests that group differences on this scoring parameter may not be revealed consistently. One would predict that groups of individuals with frontal-lobe lesions would commit more nameable errors on the DFT. Future research endeavours in this area would help to suggest possible explanations for this outcome and perhaps clarify the discrepancy between the original article (Jones-Gotman & Milner, 1977) and the present findings.

The observation that the two groups with temporal lobe pathology did not differ significantly from healthy control participants on percent perseveration scores is consistent with the original article (Jones-Gotman & Milner, 1977). According to the original source, only groups of individuals with right frontal lesions and those with right fronto-central lesions produced impaired percent perseveration scores relative to the control group.

Novel output score did not differ significantly among controls and individuals with left or right temporal-lobe pathology. Jones-Gotman and Milner (1977) also reported that individuals with left temporal-lobe lesions attained novel output scores comparable to those of healthy controls. In contrast, a combined group of individuals with anterior lesions of the right hemisphere, including those with right temporal-lobe lesions, did exhibit impaired novel output scores relative to controls. However, the other two patient groups included (i.e., groups with right frontal and right fronto-central lesions) produced significantly fewer novel designs than all other lesion groups and may have been responsible for the impaired output exhibited by the combined group relative to controls. In other words, it is possible that if the right temporal group in the Jones-Gotman and Milner (1977) study had been compared to the control group without the additional frontal cases, they may have performed at a level equivalent to controls.

In a more recent article (Jones-Gotman, 1991), impaired novel output scores on the DFT were reported for groups of individuals with surgical excisions of the right frontal, right frontotemporal, and right central regions. Jones-Gotman (1991) reported that for individuals with right frontotemporal excisions, the temporal-lobe lesions did not exacerbate impairment on the task. Thus, the impaired output for individuals with right frontotemporal excisions appeared to be attributable to the frontal-lobe lesion rather than the temporal-lobe lesion. This finding further strengthens the argument that individuals with right temporal-lobe excisions do not exhibit impaired novel output scores relative to healthy controls.

In the present study, the failure to find a difference in novel output score between controls and individuals with temporal-lobe pathology concurs with the DFT's reported sensitivity to right frontal-lobe damage. Although the number of cases in the frontal-lobe groups was too small to justify their inclusion in the data analysis (n = 9), the qualitative observation of lower mean novel output scores in these groups relative to other groups is also consistent with this assumption. The results of the present study do not contradict the assumption that the DFT is sensitive to right frontal-lobe dysfunction. Given that most scoring parameters on the DFT do not appear to be sensitive to temporal-lobe pathology, it may even be a test that is not only sensitive to cerebral dysfunction, but specific to right frontal-lobe dysfunction.

Comparison of DFT Performance Before and After Surgical Excision of Cortical Regions

For both groups of patients with temporal-lobe pathology (i.e., left-sided and right-sided), the four scores related to DFT performance (i.e., novel output score, percent perseveration, number of perseverative errors, and number of nameable errors) did not differ significantly from pre-operative to post-operative assessment. This finding suggests that DFT performance in individuals with temporal-lobe pathology is not adversely affected by the surgical removal of focal cortical areas of the affected temporal lobe. The implication is that any differences in DFT performance that exist between temporal-lobe groups and controls (i.e., a higher mean number of nameable errors) are present prior to surgical removal of the epileptogenic foci. In other words, pathology of those cerebral areas appear to be just as likely to produce any group differences as the surgical excision of cortical areas in the temporal groups.

These results do not exclude the possibility that DFT performance may be adversely affected by the surgical removal of focal cortical areas in other regions despite existing pathology in those regions prior to surgery. If DFT performance is dependent upon the integrity of functional systems which subserve the right frontal lobe, then it may be that differences in the production of novel designs would be evident in individuals before and after the surgical excision of epileptogenic foci of the right frontal-lobe regions. Although the frontal group was not included in statistical analyses, a qualitative examination of the data suggests that there was a greater difference between pre-operative and post-operative mean novel output scores in this group than in the temporal groups. This observation is also consistent with the assumption that the DFT is sensitive to right frontal-lobe damage. The failure to find differences in DFT performance before and after surgery in individuals with temporal-lobe pathology further supports the notion that the DFT is sensitive to cognitive skills subserved by systems outside of the temporal lobes.

Conclusions and Future Directions

The DFT appears to be a test that is amenable to standardization for clinical use, given that novel output scores on both conditions of the test were normally distributed. Although mean scores for normal control groups tend to vary slightly, sufficient knowledge of the sample composition should assist in the clinical interpretation of an individual's performance on the task. In any case, previous studies (Jones-Gotman & Milner, 1977; Jones-Gotman, 1991) suggest that impaired scores on the DFT would fall well below the mean scores for any of the normal samples. Inter-rater agreement and consistency was generally good to excellent, although coefficient values were lower for nameable errors and designs with the incorrect number of lines. Training in the administration and scoring of DFT protocols would likely eliminate the greater part of these discrepancies.

Novel output and perseveration on the DFT do not appear to differentiate between healthy controls and individuals with temporal-lobe pathology of either the left or the right side, which suggests that it may be a test specific to right frontal-lobe dysfunction. The fact that DFT performance does not differ between pre-operative and post-operative assessments in individuals with temporal-lobe pathology further supports this notion. The clinical use of impaired DFT scores as an indication of possible frontal lobe dysfunction in individuals with temporal-lobe pathology is not discredited by the results of this study.

Despite the large range of domains tapped by the measures used in the present study, the cognitive skills related to performance on the DFT are still unclear. No single measure or composite sensitive to verbal fluency/production, visuospatial skills, visual attention, psychomotor speed or level of psychometric intelligence appears to be related to the ability to produce novel, abstract designs within a limited time period. Combinations of these skills appear to be related to performance on the fixed condition of the DFT, and their relative contribution likely varies between individuals. Future research is needed to assess the possible contribution of factors such as divergent thinking, creativity, or other similar elements to DFT performance and whether the contributions of these factors differ between the free and fixed condition.

The results of this study suggest that individuals with anterior temporal-lobe pathology of either the right or left side generally perform at a level similar to healthy controls on the DFT. The finding that both temporal-lobe groups commit more nameable errors on the DFT than healthy controls is surprising, given that perseverative error rates and the net production of novel designs were at comparable levels among groups. Given the lower reliability indices for nameable errors relative to other scoring parameters on the DFT, it is unclear whether this finding is replicable. This development necessitates future research to determine the stability of the finding and clarify its significance.

In conclusion, although the DFT does not demonstrate excellent inter-rater reliability on all scoring parameters and the cognitive skills it taps are still unclear, the possibility that it is specifically sensitive to cerebral dysfunction of the right frontal lobe merits its inclusion in clinical neuropsychological assessment batteries.

Appendix A

History Questionnaire for Control Group

Date:__________________ Gender:______Race:____________

Date of Birth:__________________ Age:____________

First Language:____________________________________

Programme of Study:____________________________Yr.______

Education:____

Previous psychological testing:

Academic problems:

Vision:

Hearing:

Previous hospitalizations:

Medical conditions:

Head injuries/LOC:

Motor vehicle accidents:

Neurological testing (MRI,CAT,EEG):

Medications (type & dose):

Adapted from Spreen & Strauss (1991)

Appendix B

Hand Preference Questionnaire for Control Group

Show me how you would:

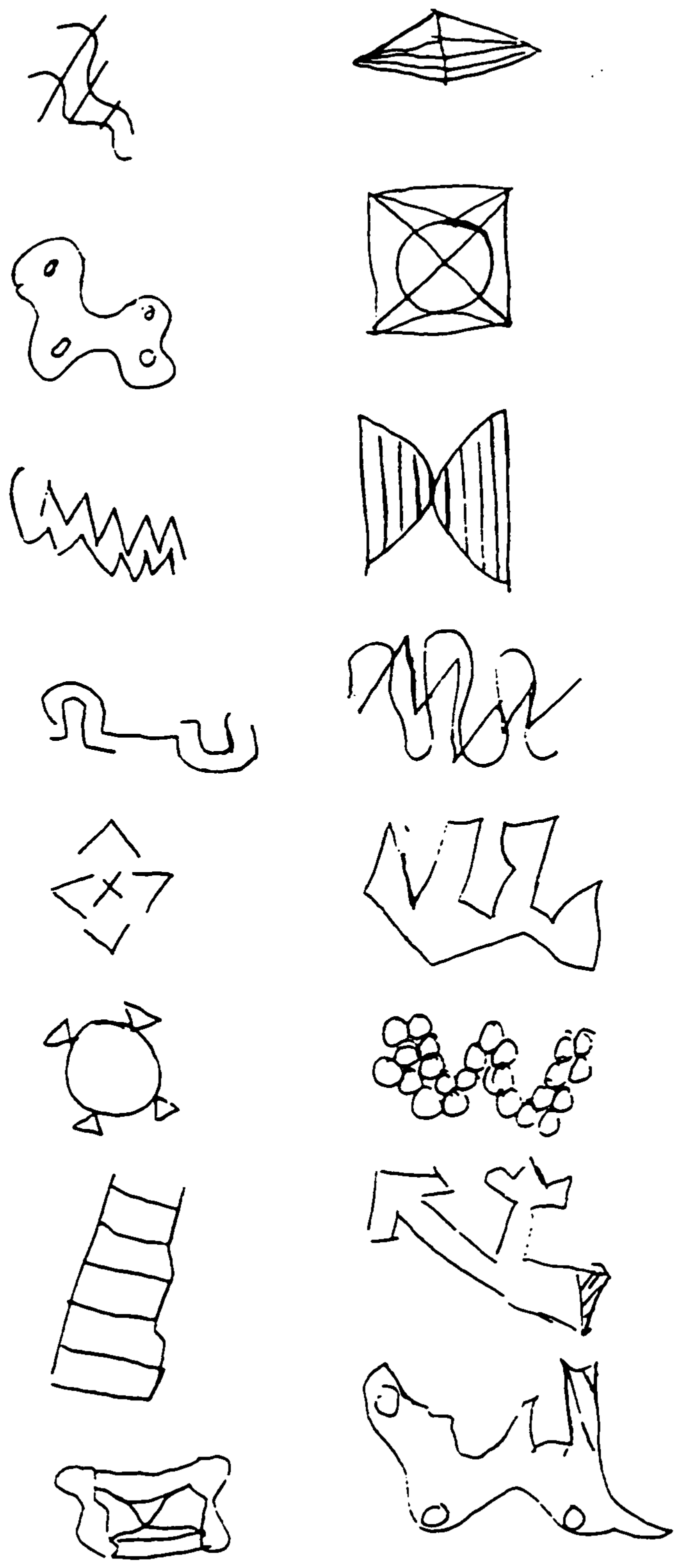
	Right or Left
1. Throw a ball.	________
2. Hammer a nail.	________
3. Cut with a knife.	________
4. Turn a door knob.	________
5. Use scissors.	________
6. Use an eraser.	________
7. Write your name.	________

Adapted from Reitan & Wolfson (1993).

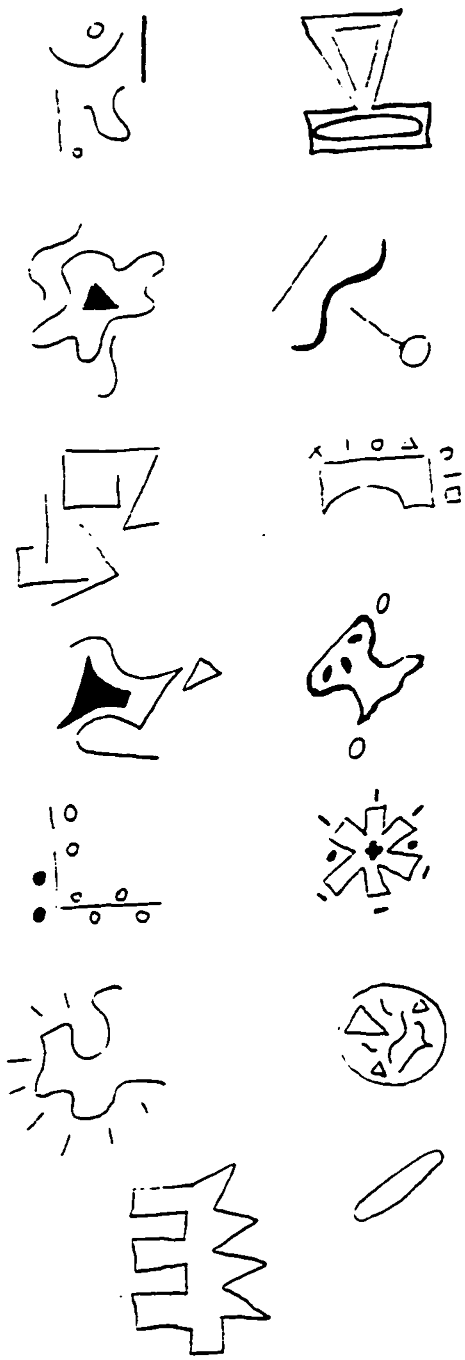
Appendix C

Examples of Design Fluency Protocols for Control Group

Example 1. Error-Free Protocol in the Free Condition

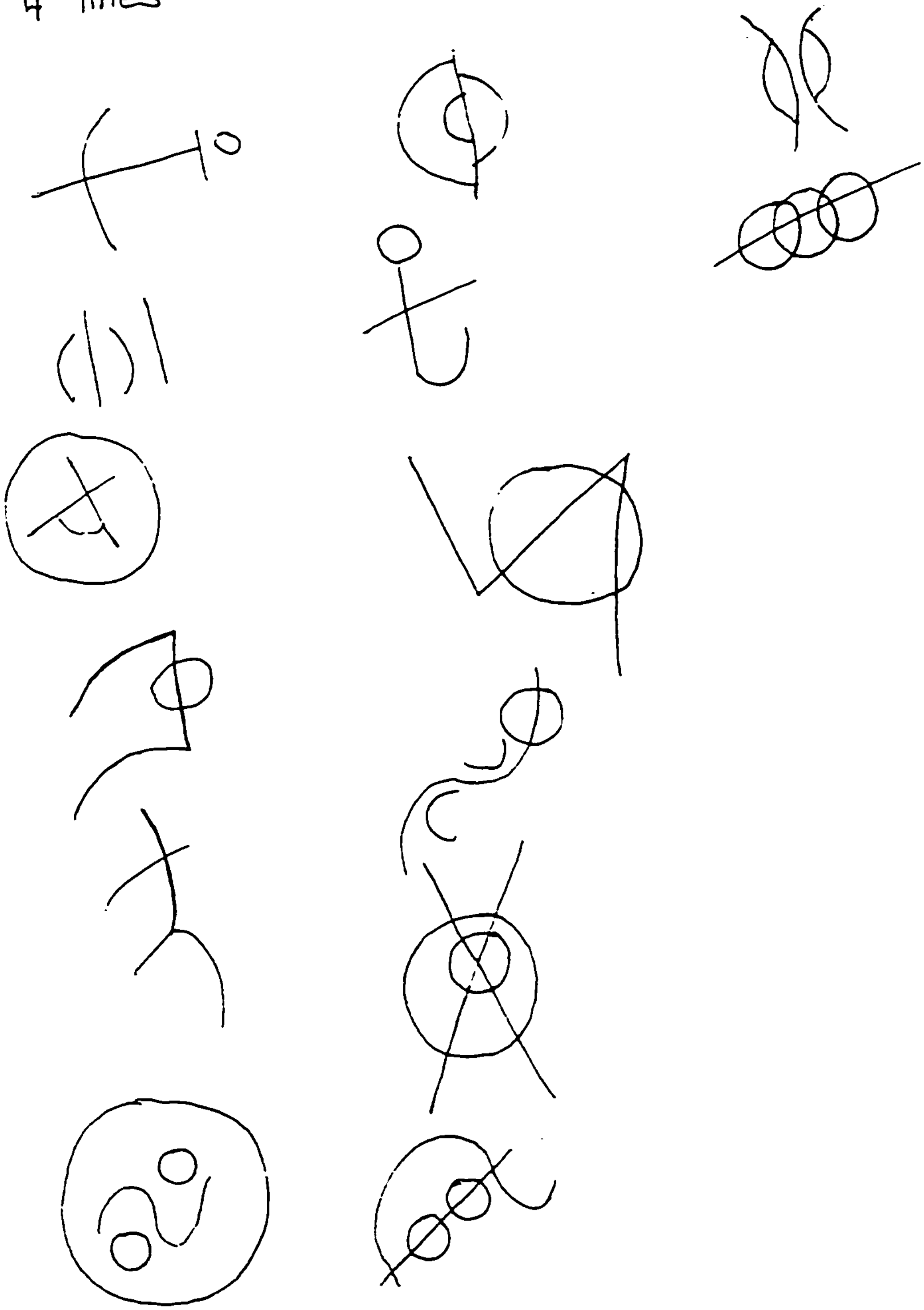


Example 2. Error-Free Protocol in the Free Condition



Example 3. Fixed Condition Protocol with Few Errors

4 lines

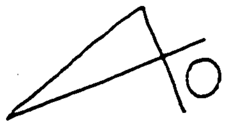


Example 4. Fixed Condition Protocol with Few Errors

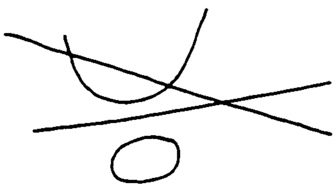
4 lines

1.

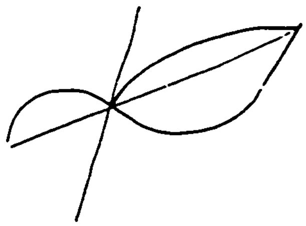
2.



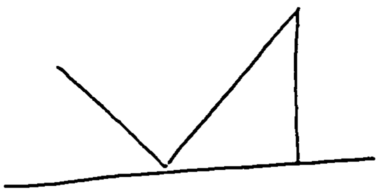
3.



4.

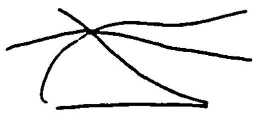


5.

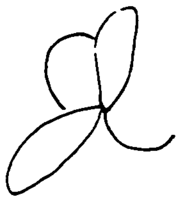


6.

7.

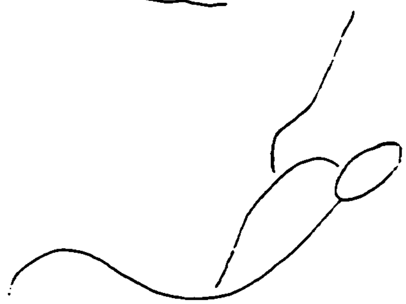


8.



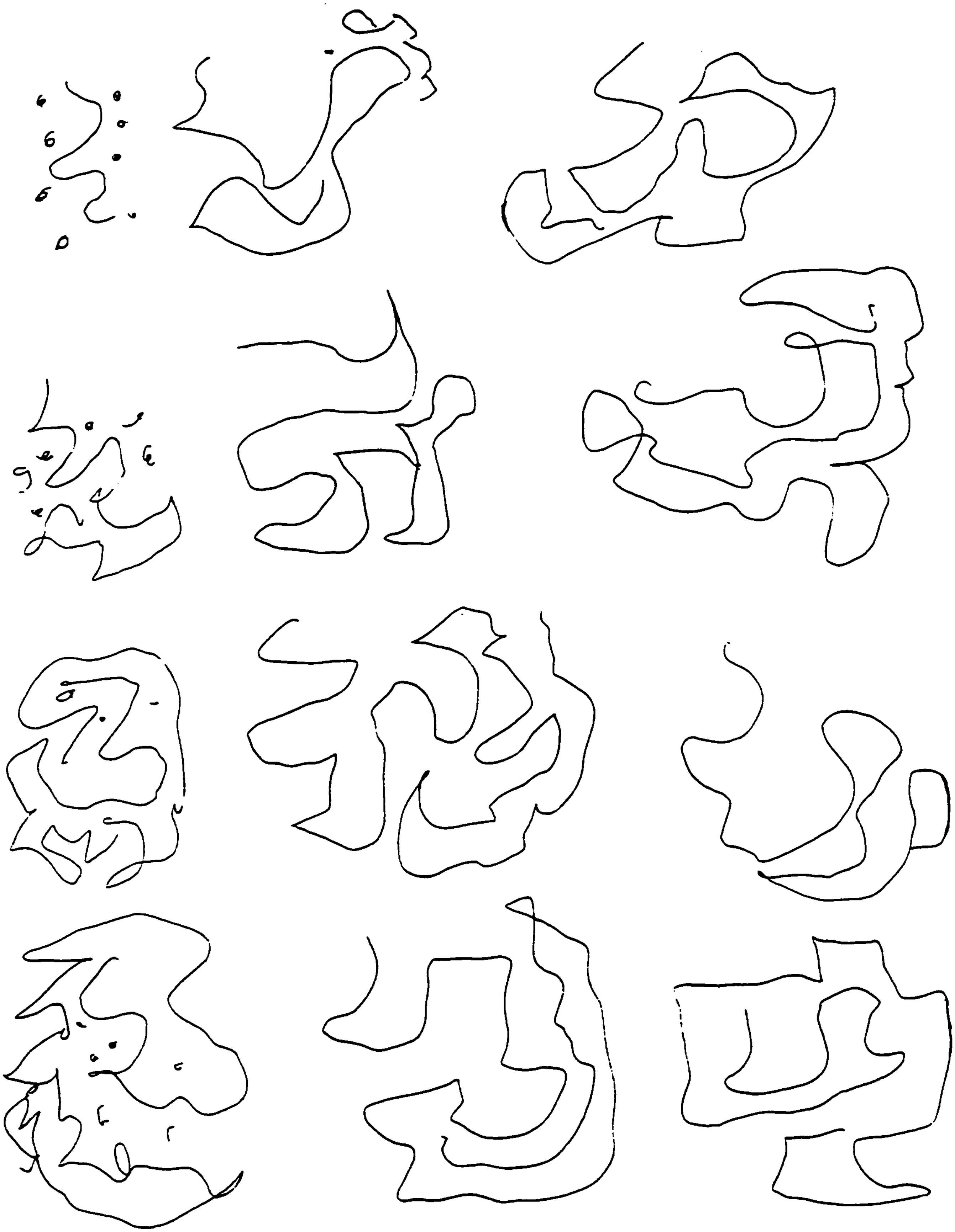
9.

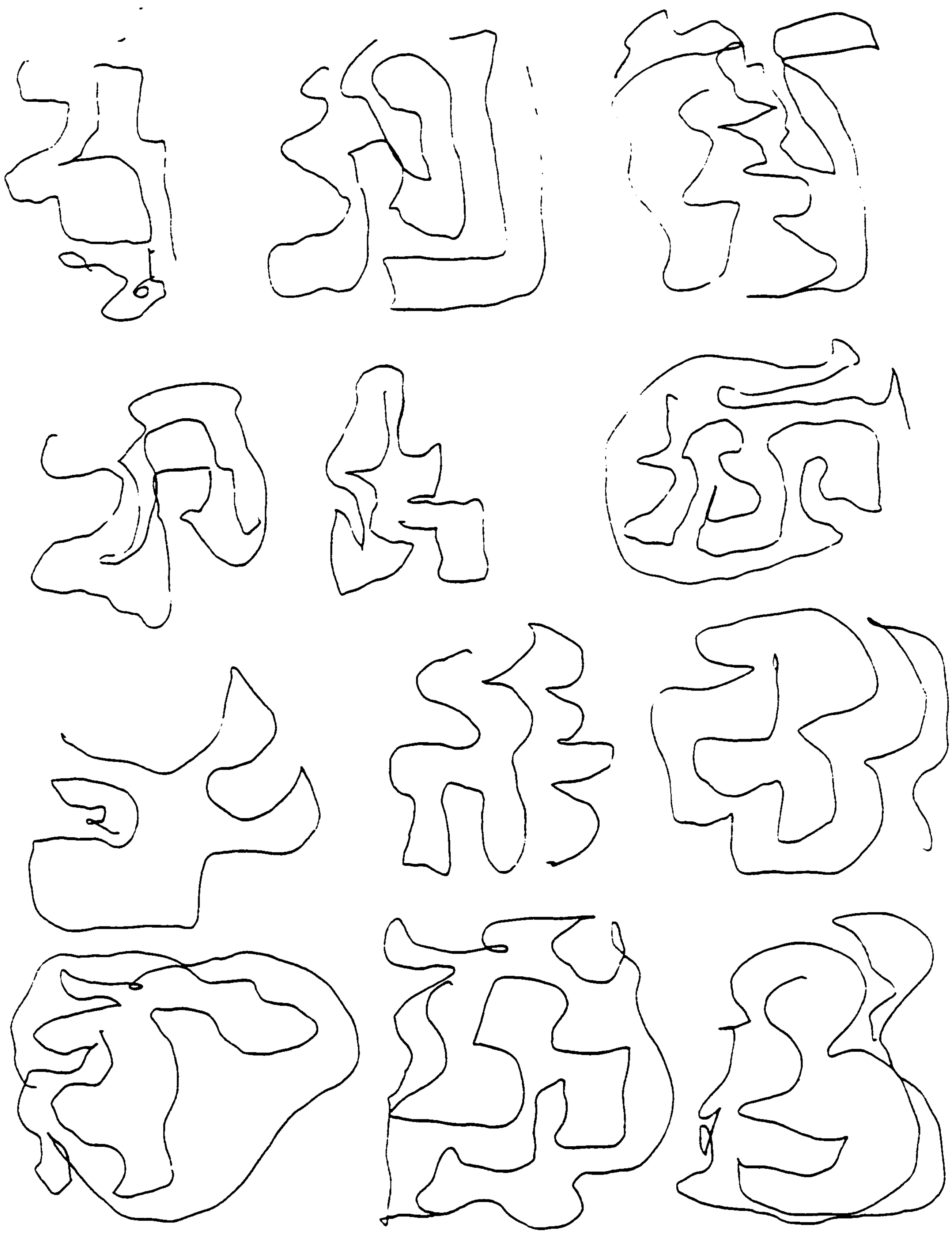
10.

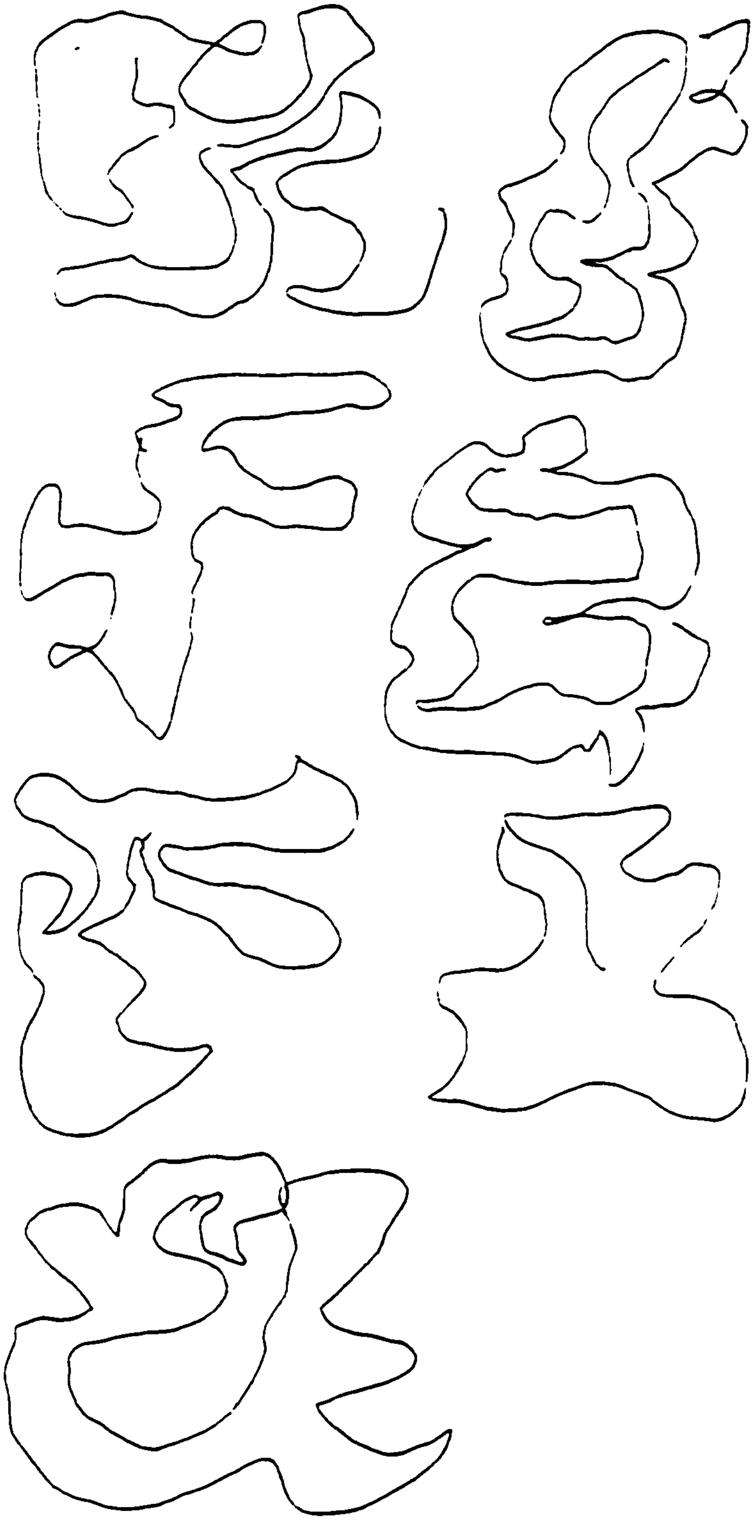


11.

Example 5. Free Condition Protocol with Many Perseverative Errors

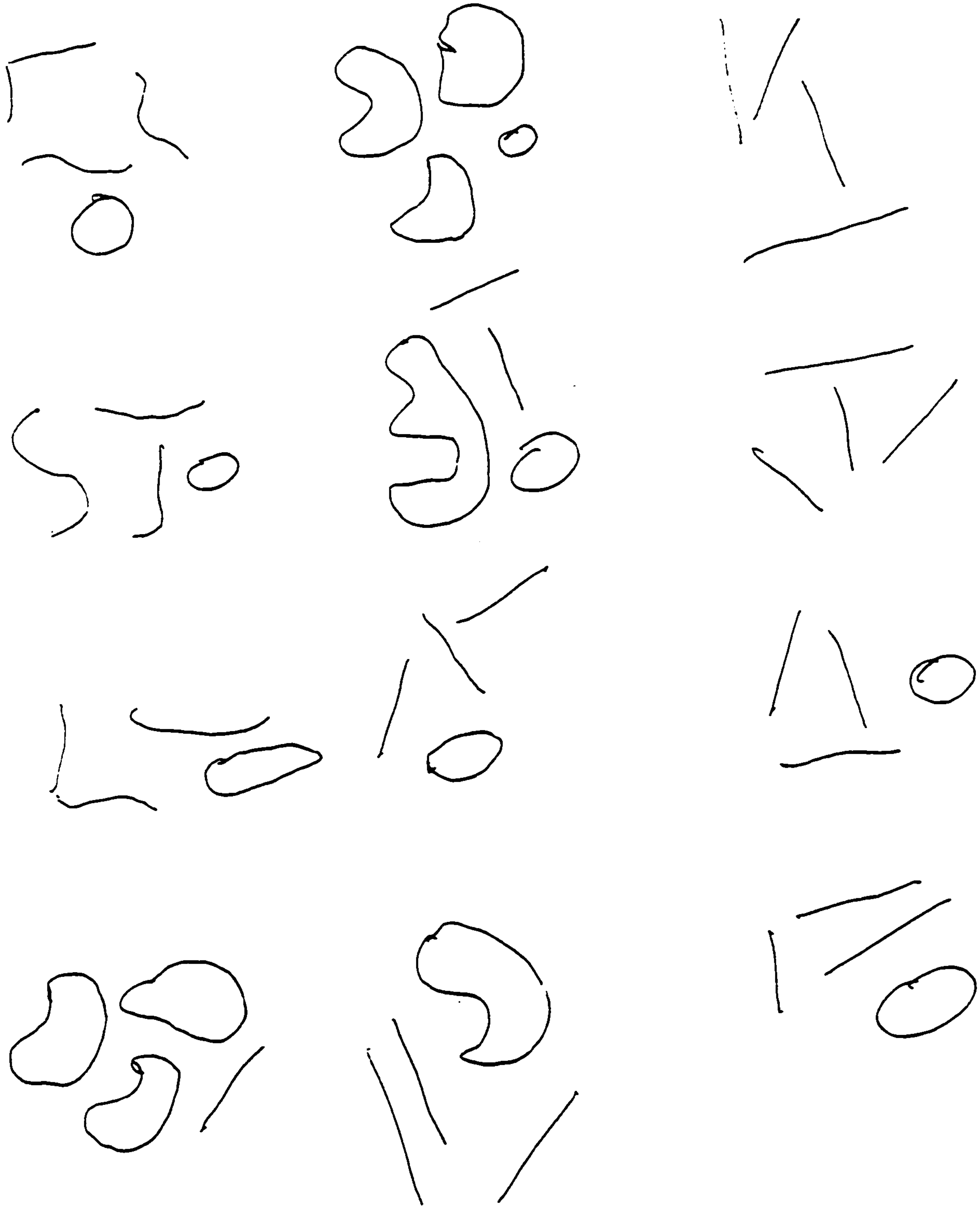




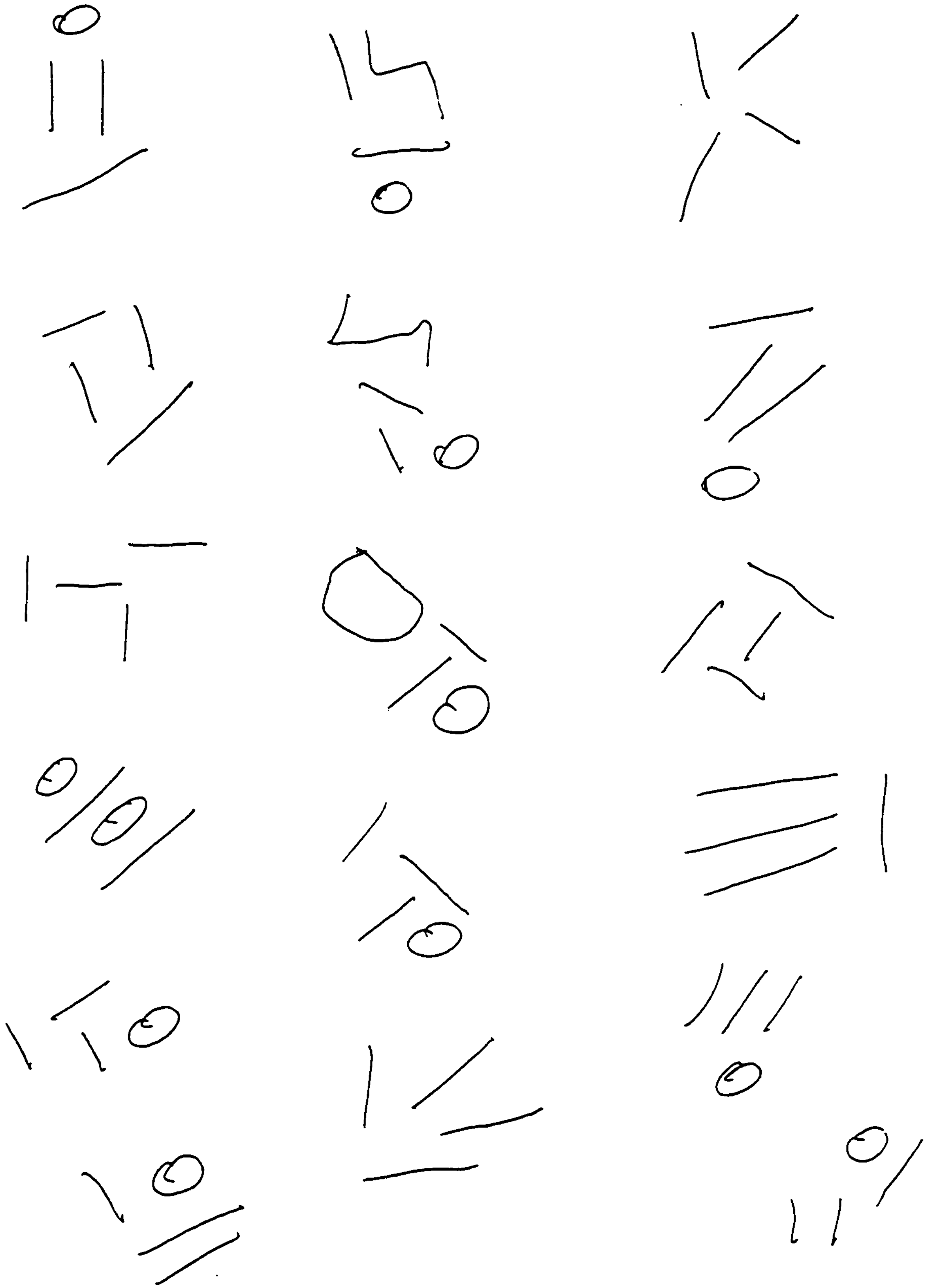


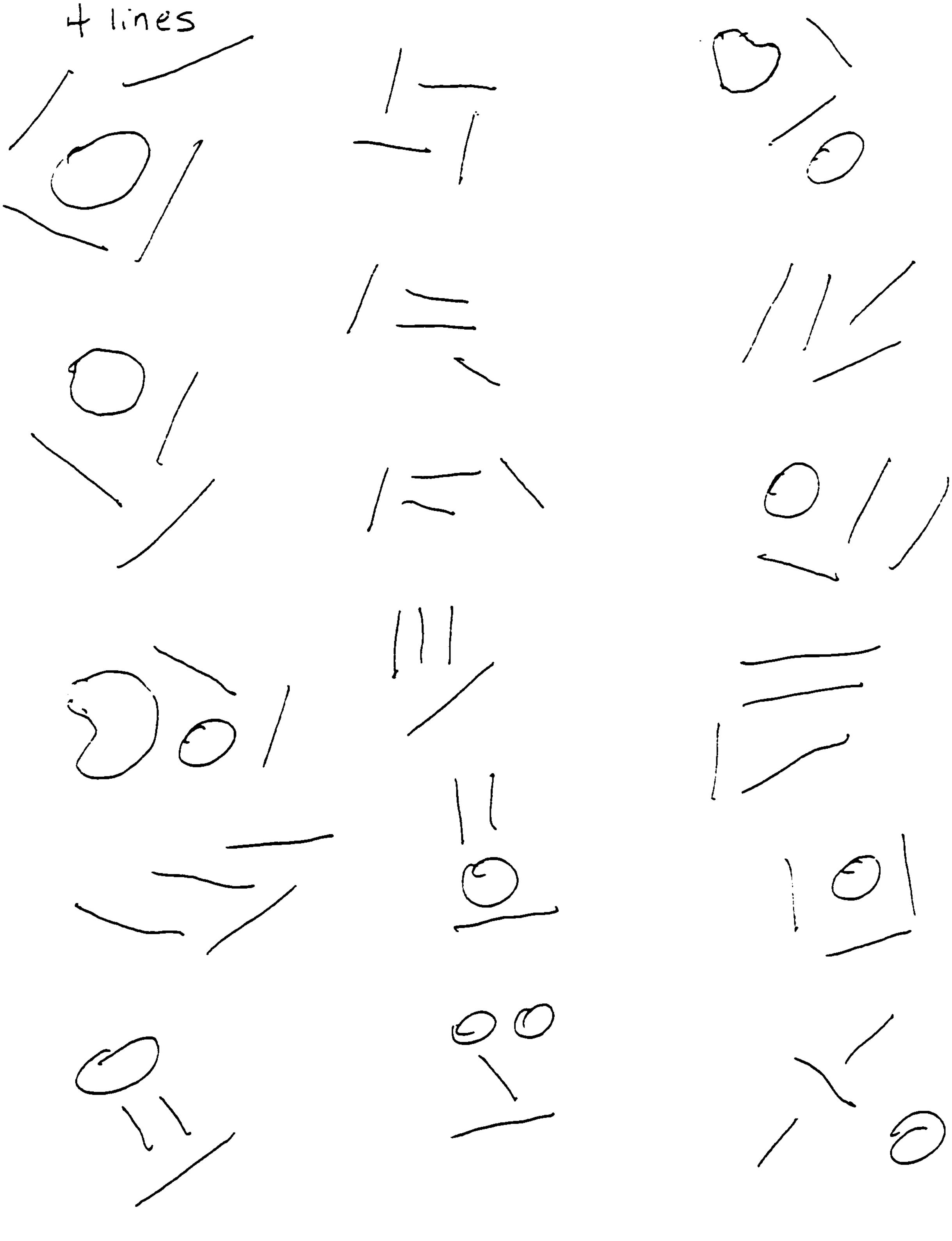
Example 6. Fixed Condition Protocol with Perseverative Errors and Incorrect Number of Lines

4 lines



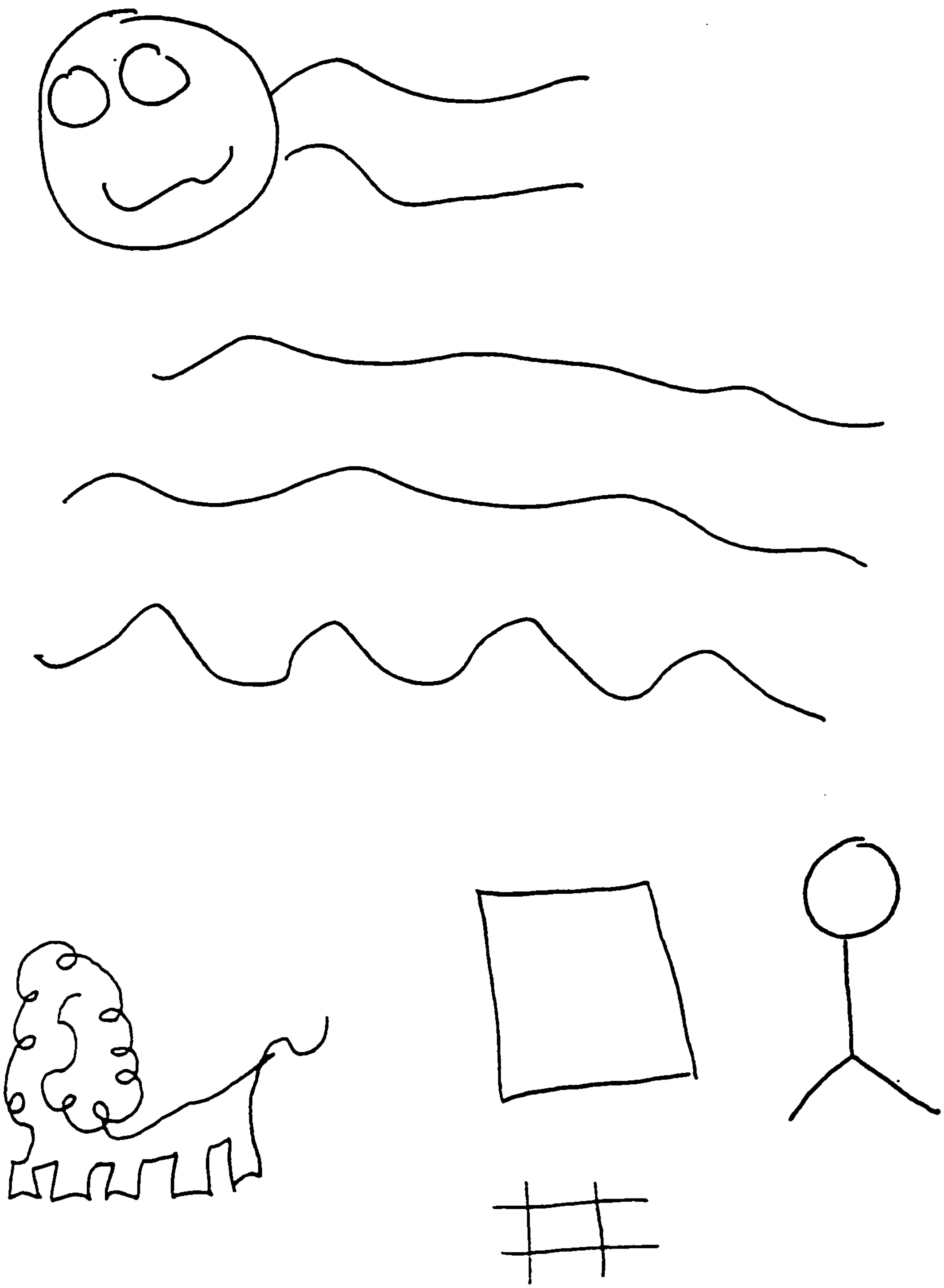
lines




4 lines

Example 7. Nameable Errors from Both

the Free and Fixed Conditions



References

Benton, A. L. (1968). Differential behavioral effects in frontal lobe disease. Neuropsychologia, 6, 53-60.

Berk, R. A. (1979). Generalizability of behavioural observations: A classification of interobserver agreement and interobserver reliability. American Journal of Mental Deficiency, 83, 460-472.

Brooker, B. H., & Cyr, J. J. (1986). Tables for clinicians to use to convert WAIS-R short forms. Journal of Clinical Psychology, 42(6), 982-986.

Bryden, M. P. (1988). Cerebral specialization: Clinical and experimental assessment. In F. Boller & J. Grafman (Eds.), Handbook of neuropsychology (Vol 1, pp. 143-159). Amsterdam: Elsevier Science.

Canavan, A. G. M., Janota, I., & Schurr, P. H. (1985). Luria's frontal lobe syndrome: psychological and anatomical considerations. Journal of Neurology, Neurosurgery, and Psychiatry, 48, 1049-1053.

Cicchetti, D. V., & Sparrow, S. A. (1981). Developing criteria for establishing interrater reliability of specific items: Applications to assessment of adaptive behavior. Journal of Mental Deficiency, 86(2), 127-137.

Correa, D.D., & Butler, R. W. (1996). Complex verbal fluency: Differential impairment in patients with frontal and nonfrontal intracranial tumors (Abstract). Journal of the International Neuropsychological Society, 2(1), 69.

Damasio, A. R., & Anderson, S. W. (1993). The frontal lobes. In K. M. Heilman & E. Valenstein (Eds.) Clinical neuropsychology (3rd ed., pp. 409-460). New York: Oxford University Press.

Heaton, R. K., Grant, I., & Matthews, C. G. (1991). Comprehensive norms for an expanded Halstead-Reitan battery: Demographic corrections, research findings, and clinical applications. Odessa, FL: Psychological Assessment Resources.

Jason, G. W., Pajurkova, E. M., Suchowersky, O., Hewitt, J., Hilbert, C., Reed, J., & Hayden, M. R. (1988). Presymptomatic neuropsychological impairment in Huntington's disease. Archives of Neurology, 45, 769-773.

Jones-Gotman, M. (no date). Design Fluency scoring instructions. Unpublished manuscript.

Jones-Gotman, M. (1991). Localization of lesions by neuropsychological testing. Epilepsia, 32 (Suppl. 5), S41-S52.

Jones-Gotman, M., & Milner, B. (1977). Design Fluency: The invention of nonsense drawings after focal cortical lesions. Neuropsychologia, 15, 653-674.

Kolb, B., & Whishaw, I. Q. (1983). Performance of schizophrenic patients on tests sensitive to left or right frontal, temporal, or parietal function in neurological patients. The Journal of Mental and Nervous Disease, 171(7), 435-443.

Lezak, M. D. (1995). Neuropsychological Assessment (3rd ed.). New York: Oxford University Press.

McCarthy, R. A., & Warrington, E. K. (1990). Cognitive neuropsychology: A clinical introduction. San Diego: Academic Press.

Mickanin, J., Grossman, M., Onishi, K., Auriacombe, S., & Clark, C. (1994). Verbal and nonverbal fluency in patients with probable Alzheimer's disease. Neuropsychology, 8(3), 385-394.

Milner, B. (1964). Some effects of frontal lobectomy in man. In J. M. Warren & K. Akert (Eds.) The Frontal Granular Cortex and Behavior. New York: McGraw-Hill.

Mittenberg, W., Seidenberg, M., O'Leary, D. S., & DiGiulio, D. V. (1989). Changes in cerebral functioning associated with normal aging. Journal of Clinical and Experimental Neuropsychology, 11, 918-932.

Parks, R. W., Loewenstein, D. A., Dodrill, K. L., Barker, W. W., Yoshii, F., Chang, J. Y., Emran, A., Apicella, A., Sheramata, W. A., & Duara, R. (1988). Cerebral metabolic effects of a verbal fluency test: A PET scan study. Journal of Clinical and Experimental Neuropsychology, 10(5), 565-575.

Pendleton, M. G., Heaton, R. K., Lehman, R. A. W., & Hulihan, D. (1982). Diagnostic utility of the Thurstone word fluency test in neuropsychological evaluations. Journal of Clinical Neuropsychology, 4(4), 307-317.

Perret, E. (1974). The left frontal lobe of man and the suppression of habitual responses in verbal categorical behaviour. Neuropsychologia, 12, 323-330.

Reitan, R. M., & Wolfson, D. (1993). The Halstead-Reitan neuropsychological test battery. Tucson, AZ: Neuropsychology Press.

Ruff, R. M., Allen, C. C., Farrow, C. E., Niemann, H., & Wylie, T. (1994). Figural fluency: Differential impairment in patients with left versus right frontal lobe lesions. Archives of Clinical Neuropsychology, 9, 41-55.

Ruff, R. M., Light, R. H., Parker, S. B., & Levin, H. S. (in press). Controlled oral word association test: Reliability and updated norms.

Sechrest, L. (1984). Reliability and validity. In A. S. Bellack & M. Hersen (Eds.), Research methods in clinical psychology (pp. 24-54). Toronto: Pergamon.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations uses in assessing rater reliability. Psychological Bulletin, 86, 420-428.

Speedie, L. J., & Heilman, K. M. (1983). Anterograde memory deficits for visuospatial material after infarction of the right thalamus. Archives of Neurology, 40, 183-186.

Spreen, O., & Benton, A. L. (1969). Neurosensory Center Comprehensive Examination for Aphasia (NCCEA). Victoria: University of Victoria, Neuropsychological Laboratory.

Spreen, O., & Strauss, E. (1991). A compendium of neuropsychological tests: Administration, norms and commentary. New York: Oxford University Press.

Springer, S. P., & Deutsch, G. (1993). Left brain, right brain. (2nd ed). New York: Freeman.

Tabachnick, B. G., & Fidell, L. S. (1989). Using multivariate statistics (2nd ed.). New York: Harper Collins.

Tellegen, A., & Briggs, P. F. (1967). Old wine in new skins: Grouping Wechsler subtests into new scales. Journal of Consulting Psychology, 31, 499-506.

Thurstone, L. L. (1938). Primary Mental Abilities. Chicago: University of Chicago Press.

Thurstone, L. L., & Jeffrey, T. E. (1956). Space Thinking (Flags). Park Ridge, IL: London House.

Thurstone, L. L., & Jeffrey, T. E. (1984). Closure Speed (Gestalt Completions). Park Ridge, IL: London House.

Thurstone, L. L., & Jeffrey, T. E. (1984). Closure Flexibility (Concealed Figures). Park Ridge, IL: London House.

Thurstone, L. L., & Jeffrey, T. E. (1987). Perceptual Speed (Identical Forms). Park Ridge, IL: London House.

Wechsler, D. (1981). WAIS-R manual. New York: Psychological Corporation.

Wechsler, D. (1987). WMS-R manual. New York: Psychological Corporation.

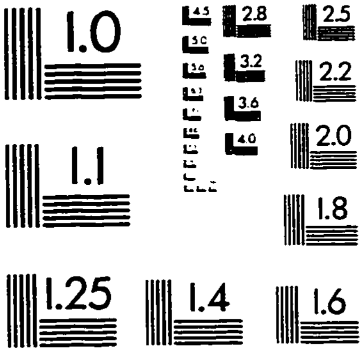
Woodard, J. L., Axelrod, B. N., & Henry, R. R. (1992). Interrater reliability of scoring parameters for the Design Fluency Test. Neuropsychology, 6(2), 173-178.

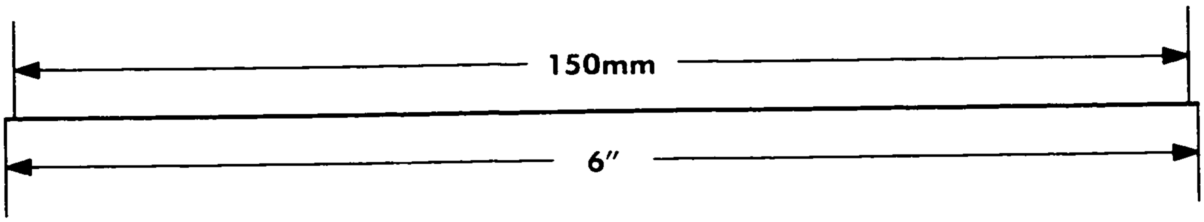
Varney, N. R., Roberts, R. J., Struchen, M. A., Hanson, T. V., Franzen, K. M., & Connell, S. K. (1996). Design fluency among normals and patients with closed head injury. Archives of Clinical Neuropsychology, 11(4), 345-353.

VITA AUCTORIS

Sherri L. Carter was born on August 21, 1971 in Mississauga, Ontario. In June 1990, she graduated from Gordon Graydon Memorial High School, Mississauga, Ontario. In September, 1990 she enrolled at the University of Western Ontario. She graduated with a Bachelor of Arts (Honours) degree in Psychology in June, 1994. Since September 1994 she has been in enrolled in the Master's programme in Clinical Neuropsychology at the University of Windsor.

IMAGE EVALUATION TEST TARGET (QA-3)





APPLIED IMAGE . Inc
1653 East Main Street
Rochester, NY 14609 USA
Phone: 716/482-0300
Fax: 716/288-5989