16 Executive Functions and Motor Performance

THE EXECUTIVE FUNCTIONS

As the most complex of behaviors, executive functions are intrinsic to the ability to respond in an adaptive manner to novel situations and are also the basis of many cognitive, emotional, and social skills. The executive functions can be conceptualized as having four components: (1) volition; (2) planning; (3) purposive action; and (4) effective performance. Each involves a distinctive set of activity-related behaviors. All are necessary for appropriate, socially responsible, and effectively self-serving adult conduct. Moreover, it is rare to find a patient with impaired capacity for self-direction or self-regulation who has defects in just one of these aspects of executive functioning. Rather, defective executive behavior typically involves a cluster of deficiencies of which one or two may be especially prominent. Paradoxically, such profound changes in behavior are sometimes missed in a highly structured examination (Lezak, 1982a).

A medically retired financial manager whose cardiac arrest was complicated by a hard fall onto his right temple was very responsive to his own needs and energetic in attempts to carry out plans. Unfortunately, he could no longer formulate plans well because of an inability to take all aspects of a situation into account and integrate them. His lack of awareness of his mistakes further aggravated this disability. Problems arising from his emotional lability and proneness to irritability were overshadowed by crises resulting from his efforts to carry out inappropriate and sometimes financially hazardous plans.

In these cases and in much of the literature concerning the executive functions, frontal lobe damage is implicated. This is not surprising since most patients who have had significant injury or disease of the prefrontal regions, particularly when orbital or medial structures are involved, experience behavioral and personality changes stemming from defective executive functions. The classic tale of Phineas Gage is the first careful documentation of a person whose personality was strikingly altered from conforming and productive to irresponsible and unruly when a misfired tamping iron lodged in his head "passing back of the left eye, and out at the top of his head," thus running through his left frontal lobe (Macmillan, 2000).

However, the executive functions are also sensitive to damage in other parts of the brain (E. Goldberg and Bilder, 1987; E. Goldberg, Bilder, and Hughes, 1989; Lezak, 1994). Subcortical as well as cortical damage can be involved (Dujardin et al., 2000; Eslinger and Grattan, 1993; Hashimoto et al., 1995). Disturbances in executive functions may result from anoxic conditions that involve limbic structures (Januzzi and McKhann, 2002) and can be among the sequelae of alcohol abuse (Munro et al., 2000) or inhalation of organic solvents (Arlien-Søborg et al., 1979; Hawkins, 1990; Tsushima and Towne, 1977). Korsakoff patients with lesions primarily in thalamic nuclei and other subcortical components of the limbic system typically exhibit profound disturbances in executive behavior. Many of them are virtually immobilized by apathy and inertia. Some Parkinson patients display diminished conceptual flexibility and impaired initiative and spontaneity. Moreover, patients with right hemisphere damage who can "talk a good game" and are neither inert nor apathetic may be ineffective because limitations in organizing conceptually all facets of an activity and integrating it with their behavior may keep them from carrying out their many intentions.

Executive functions can break down at any stage in the behavioral sequence that makes up planned or intentional activity. Systematic examination of the capacities that enter into the four aspects of executive activity will help to identify the stage or stages at which a breakdown in executive behavior takes place. Such a review of a patient's executive functions may also bring to light impairments in self-direction or self-regulation that would not become evident in the course of the usual examination or observation procedures.

A major obstacle to examining the executive functions is the paradoxical need to structure a situation in which patients can show whether and how well they can make structure for themselves. Typically, in formal examinations, the examiner determines what activity the subject is to do with what materials, when, where, and how. Most cognitive tests, for example, allow the subject little room for discretionary behavior, including many tests thought to be sensitive to executive—or frontal lobe—disorders (Frederiksen, 1986; Lezak, 1982a; Shallice and Burgess, 1991b). The problem for clinicians who want to examine the executive functions becomes how to transfer goal setting, structuring, and decision making from the clinician to the subject within the structured examination. A limited number of established examination techniques give the subject sufficient leeway to think of and choose alternatives as needed to demonstrate the main components of executive behavior.

The following review covers techniques that may be useful in exploring and elucidating this most subtle and central realm of human activity (see also Lezak, 1989; A.C. Roberts et al., 1998). Other instruments presented in this chapter test more peripheral but equally important executive capacities, such as those that enter into self-regulation and self-correction.

Volition

> The distinction between an action that is intentional and one that is not seems to have something to do with the consciousness of the goal of the action.
>
> *J.W. Brown, 1989*

Volition refers to the complex process of determining what one needs or wants and conceptualizing some kind of future realization of that need or want. In short, it is the capacity for intentional behavior. It requires the capacity to formulate a goal or, at a lower conceptual level, to form an intention. Motivation, including the ability to initiate activity, is one necessary precondition for volitional behavior. The other is awareness of oneself psychologically, physically, and in relation to one's surroundings. Each aspect of volition can be examined separately. Deficiencies in self-initiated behavior may occur because of disturbances in cognitive/affective processes due to damage to frontal/subcortical or frontolimbic circuitry (Stuss, Van Reekum, and Murphy, 2000), to the right hemisphere, or in diffuse conditions such as Alzheimer's disease (R.S. Marin et al., 1994).

Persons who lack volitional capacity simply do not think of anything to do. In extreme cases they may be apathetic, or unappreciative of themselves as distinctive persons (much as an infant or young child), or both. They may be unable to initiate activities except in response to internal stimuli such as bladder pressure or external stimuli, for example, an annoying mosquito. Such persons may be fully capable of performing complex activities and yet not carry them out unless instructed to do so. For instance, although able to use eating utensils properly, some will not eat what is set before them without continuing explicit instructions. Less impaired persons may eat or drink what is set before them, but will not seek nourishment spontaneously, even when hungry. Patients whose volitional capacity is only mildly impaired can do their usual chores and engage in familiar games and hobbies without prompting. However, they are typically unable to assume responsibilities requiring appreciation of long-term or abstract goals and do not enter into new activities independently. Without outside guidance, many wander aimlessly or sit in front of the television or at the same neighborhood bar or coffee shop when they have finished their routine activities.

In some cases, particularly when deficits are subtle, it becomes important to identify the presence of a volitional defect. In others, where passivity or apparent withdrawal are obvious behavioral problems, the examiner must try to distinguish the unmotivated, undirected, and disinterested anergia occurring on an organic basis from characterological (e.g., laziness, childish dependency) or psychiatric (e.g., depression, schizophrenia) disorders that superficially appear similar. However, there are no formal tests for examining volitional capacity. The examiner must rely instead on observations of these patients in the normal course of day-to-day living and reports by caregivers, family, and others who see them regularly. These reports are often the best sources of information about the patient's capacity for generating desires, formulating goals, and forming intentions. Thus the examination should include both the patient and those who know the patient best.

Examining motivational capacity

The direct examination of motivational capacity should inquire into patients' likes and dislikes, what they do for fun, and what makes them angry, as many volitionally impaired patients are apathetic with diminished or even absent capacity for emotional response. The patient's behavior in the examination can also provide valuable clues to volitional capacity. Volitionally competent persons make spontaneous—and appropriate—conversation or ask questions; or they participate actively in the examination proceedings by turning test cards or putting caps back on pens. Patients whose volitional capacity is seriously impaired typically volunteer little or nothing, even when responsive to what the examiner says or does. Some patients report what sound like normal activity programs when asked how they spend their leisure time or how they perform chores. Then the examiner needs to find out when they last dated or went on a camping trip, for example, or who plans the meals they cook. A patient may report that he likes to take his girlfriend to the movies but has not had a "girlfriend" since before his accident three years ago and has not gone to a theater since then either. Another who talks about her competence in the kitchen actually prepares the same few dishes over and over again exactly as taught since being impaired.

Excerpts from an interview with a physically competent 26-year-old woman two years after she had become fully dependent as a result of massive frontal lobe damage incurred in a car vs. train accident shows how an interview can document a severe motivational impairment.
Q: What kind of work did you do? P: In a state park.
Q: Did you like that work? P: It was OK.
Q: How come you stopped doing that work? P: I don't know.
Q: Have you thought of going back to do it? P: I really don't care. . . .
Q: What would you do if your mother got sick and had to go to the hospital? P: If it was late I would put on my pyjamas and go to bed and go to sleep.
Q: And then what would you do the next day when no one was home? P: I would have to get up and eat breakfast and then go and get dressed.
Q: And then what would you do? P: Come in and turn on the TV and sit down.
Q: And then what would you do? P: After I watched TV, I would put on my shoes and socks and go back into the bedroom and sit down because I don't know anyone else to call.

Examination techniques can require the patient to initiate activity. Heilman and Watson (1991) scatter pennies on the table in front of patients, then blindfold them and tell them to pick up as many pennies as they can. The task thus requires exploratory behavior which may be lacking in patients whose capacity to initiate responses is impaired.

Examining the capacity for self-awareness

Assessment of self-awareness and awareness of one's surroundings also depends upon observations and interviews. Like other aspects of executive functioning, defective self-awareness occurs to varying degrees. Moreover, self-awareness is multifaceted as it includes physical awareness, awareness of self and of other persons, and social awareness. Mature self-awareness requires an integrated appreciation of one's physical status and ongoing physical relationship with the immediate external environment; an appreciation of being a distinctive person in a world which mainly exists outside of one's immediate awareness and is inhabited by many other distinctive individuals; and appreciation of oneself as an interactive part of the network of social relationships. Each of these facets of self-awareness can be disturbed by brain damage and each can be examined in its own right. Stuss and Alexander (2000) propose that the right frontal lobe with its hypothalamic limbic, posterior cortical, and cingulate connections plays an essential role in self-awareness.

Awareness of one's physical status. Inaccurate body images can occur as distortions, perceptions of more severe impairment than is the case, or as feelings of being intact when actually impaired. The most direct method for examining body image is to request a human figure drawing. Inquiry into vocational or career plans, or just plans for going home can elicit defective self-perceptions, as when a visually impaired youngster says he plans to be a pilot, or a wheelchair-bound patient assures the examiner he will be able to walk the flight of stairs to his apartment. An associated deficit can show up in impaired appreciation of one's physical strengths and limitations. Reduced or even absent appreciation of physical states and bodily functions usually involves loss of appetite or loss of satiation cues, sexual disinterest, with sleep disturbances not uncommon. Interviews with patient and family or other caregivers, and sensitive observation typically bring these problems to light.

Awareness of the environment and situational context. The extent to which patients are aware of and responsive to what goes on around them is likely to be reflected in their use of environmental cues. This can be examined with questions about the time of day, the season of the year, or other temporal events or situational circumstances (e.g., Christmas time, the dining hall, office, or waiting room, etc.) that can be easily deduced or verified by alert patients who are attentive to their surroundings.

A 58-year-old woman with mild frontotemporal dementia carried an armful of possessions to her appointment and placed them on the examination table. The examiner explained that the table must be clear for the examination and stood up to find somewhere to put the belongings. While the examiner's back was turned, the patient announced that she had found a place for her things—in the examiner's chair. Formal testing did not show many deficits, but numerous behaviors during the examination demonstrated impaired situational awareness.

Story and picture material from standard tests can also be used to examine the patient's ability to pay attention to situational cues. The Problems of Fact items of the 1973 revision of Terman and Merrill's Stanford-Binet scales require the patient to use cues to interpret a situation. The Cookie Theft picture from the Boston Diagnostic Aphasia Examination (Goodglass, Kaplan, and Barresi, 2000) or Picture Arrangement items from the WIS-A are excellent for testing the patient's ability to infer a story from a picture. The complexity and richness of responses may range from a single integrated story involving the important elements of the picture, to a bit-by-bit description of the picture that raises questions about whether the patient can integrate what is seen, to a disregard of all but one or two items because of impaired capacity to attend systematically or persevere in an activity.

Social awareness. Assessment of social awareness also depends upon observations and interviews. Lack of normal adult self-consciousness may show up in reports or observations of poor grooming and childish or crude behavior that contrast sharply with a premorbid history of social competence.

The same woman described above with frontotemporal dementia (autopsy confirmed) startled strangers in a grocery store by punching them on the shoulder and saying in a loud voice, "My daughter went to Cal Tech." Obviously proud of her daughter, the woman was unable to appreciate the inappropriateness of her behavior.

At the other extreme, excessive politeness may also expose impaired social awareness.

A very bright Vietnam veteran who had sustained a blow that crushed the anterior portion of his right frontal lobe was still able to complete a university level accounting program and qualify as a CPA. Even after working for more than ten years in his profession and in a major metropolis he continued to address women as "Ma'am," including those with whom he worked, as he had been taught to do as a child. Loneliness and feeling out of touch socially were persistent problems for him.

How patients dress and groom themselves, how they relate to the examiner or to other clinical staff, and how they interact with their family members can provide important information regarding their appreciation of social roles and accepted codes of social behavior. Interviews with the patient and family members can be invaluable in making social disturbances evident. Test responses, too, may offer insight into the patient's social understandings. For example, patients who say they would "shout fire" in a theater (WAIS and WAIS-R item) are out of touch with what is both socially acceptable and socially responsible behavior.

Planning

The identification and organization of the steps and elements (e.g., skills, material, other persons) needed to carry out an intention or achieve a goal constitute planning and involve a number of capacities. In order to plan, one must be able to conceptualize changes from present circumstances (i.e., look ahead), deal objectively with oneself in relation to the environment, and view the environment objectively (i.e., take the abstract attitude; see p. 84). The planner must also be able to conceive of alternatives, weigh and make choices, and entertain both sequential and hierarchical ideas necessary for the development of a conceptual framework or structure that will give direction to the carrying out of a plan. Good impulse control and reasonably intact memory functions are also necessary. Moreover, all of this conceptual activity requires a capacity for sustained attention. Patients who are unable to form a realistic intention also cannot plan. However, some patients who generate motives and initiate goal-directed activity spontaneously fail to achieve their goals because one or more of the abilities required for effective planning is impaired.

Use of standard examination procedures

There are few formal tests of planning ability per se. However, the patient's handling of many of the standard psychological tests will provide insight into the status of these important conceptual activities.

Responses to storytelling tasks, such as the Thematic Apperception Test, reflect the patient's handling of sequential verbal ideas. Stories told to these pictures may be complex and highly organized, have simple and straight story lines, be organized by accretion, or consist of loose or disjointed associations or descriptions (W.E. Henry, 1947). How patients address such highly structured tests as Picture Arrangement and Block Design will provide information about whether they order and plan ahead naturally and effectively, laboriously, inconsistently, or not at all. Sentence Arrangement of the WAIS-RNI affords a good opportunity to see whether patients can organize their thoughts into a sensible and linguistically acceptable construct.

The Complex Figure Test also elicits planning behavior. Osterrieth's (1944) analysis of how people go about copying the complex figure provides standards for evaluating how systematic is the patient's response to this task. A haphazard, fragmented mode of response suggests poor planning; while a systematic approach beginning with the basic structure of the figure is generally the hallmark of someone who plans well. Some examiner techniques capture the sequence of the drawing and a representation of the plan (see pp. 544–547). Several scoring systems assess the organizational approach used to copy the figure (e.g., Deckersbach et al., 2000). A time-consuming system, the Boston Qualita-

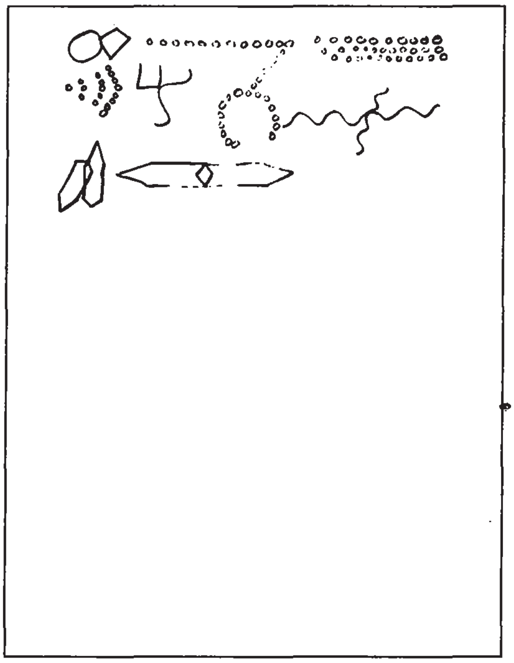

FIGURE 16.1 Bender-Gestalt copy trial rendered by a 42-year-old interior designer a year after she had sustained a mild anterior subarachnoid hemorrhage. Note that although the design configurations are essentially preserved, she used only one-third of the page, drawing several of the designs as close to each other as to elements within these designs.

tive Scoring system, includes Planning as one of its main scores (Somerville et al., 2000; R.A. Stern et al., 1999).

The patient's use of space in drawings can provide a concrete demonstration of planning defects. The Bender-Gestalt designs are particularly well-suited to this purpose (see Fig. 16.1); but free drawings (e.g., human figures, house, etc.) may also elicit planning problems (see Fig. 16.2).

Questioning can bring out defective planning. How patients who are living alone or keeping house describe food purchasing and preparation may reveal how well they can organize and plan. Other issues that may bring out organizing and planning abilities concern personal care, appreciation of how disability affects the patient's activities and family, what accommodations the patient has made to disability, to altered financial and vocational status, etc. Hebb (1939) offered a pertinent question used by his colleague, Dr. W.T.B. Mitchell: "What should you do before beginning something important?" (to which a patient who had undergone a left frontal lobectomy replied, after some delay, "I can't get it into my head"). Some patients, particularly those whose lesions are in the right hemisphere, may give lucid and appropriate answers to questions involving organization and planning of impersonal situations or events but show poor judgment in unrealistic, confused, often illogical, or nonexistent plans for themselves, or lack the judgment to recognize that they need to make plans if they are to remain independent (Lezak, 1994).

Working memory tasks

Efficient planning entails decision making and developing strategies for setting priorities. Atkinson and Shiffrin (1968) described working memory as involving, among other processes, decision making. Using a similar model, Baddeley and Hitch (1974) proposed that working memory depends on an attentional controller, which they called the *central executive* and later the *supervisory attentional system* (Baddeley, 1986; Norman and Shallice, 1986). In this model, the supervisory attentional system selects and operates strategies for maintaining and switching attention as needs arise.

> Executive processes include various kinds of judgments made on stimuli held in short-term memory (e.g., comparison of the relative recency of stimuli, judgments of the relative saliency or novelty of stimuli, etc.) and active (voluntary) retrieval of specific information held in long-term form. (Petrides, 1994, p. 73)

Failure of this sytem produces the *dysexecutive syndrome* (Baddeley, 1986; Baddeley and Wilson, 1988).

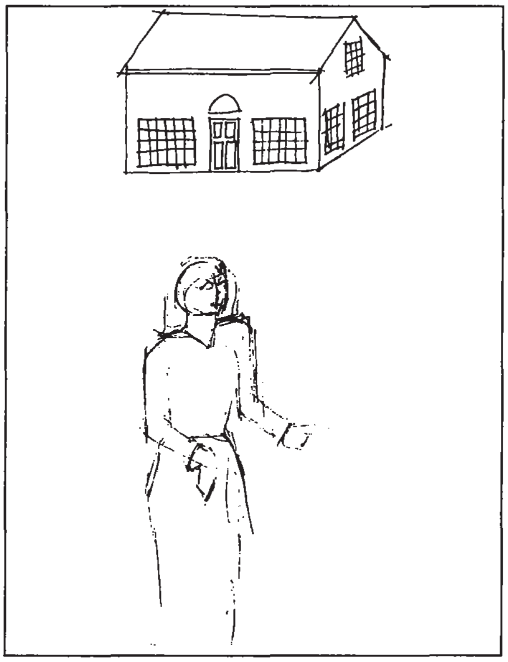

FIGURE 16.2 House and Person drawings by the interior designer whose Bender-Gestalt copy trial is given in Figure 16.1. Note absence of chimney on a highly detailed house drawing and placement and size of woman too low and too large to fit all of her on the page.

The dorsal prefrontal cortex appears critical for allocating attentional resources during working memory tasks (Koechlin, Basso, et al., 1999; Koechlin, Corrado, et al., 2000; Petrides, 1994).

Self-Ordered Pointing Test (Petrides and Milner, 1982)

Tests calling for self-ordered responses assess strategy use and self-monitoring. In the *Self-Ordered Pointing Test,* the examiner asks subjects to point to a stimulus (e.g., abstract designs, line drawings; see Spreen and Strauss, 1998, pp. 209 and 210) not seen on previous trials in an array of stimuli on each trial. The position of the stimuli shifts from trial to trial so that the subject must try to monitor previous choices from memory. Patients with frontal lesions were impaired on this task compared to those with temporal lesions; the authors attributed this relative impairment to poor organizational strategies, poor monitoring of responses, or both. Age effects have been reported for this task (Daigneault and Braun, 1993). Deficits also have been observed in patients with Huntington's disease (Rich, Blysma, and Brandt, 1996) and Parkinson's disease (Gabrieli et al., 1996; West et al., 1998). West and his colleagues observed that most errors of Parkinson patients occurred toward the end of a trial regardless of set size, which they suggested resulted from failure to monitor how far they had proceeded in the trial.

Maze tracing

The maze tracing task was designed to yield data about the highest levels of mental functioning involving planning and foresight, i.e., "the process of choosing, trying, and rejecting or adopting alternative courses of conduct or thought. At a simple level, this is similar to solving a very complex maze" (Porteus, 1959, p. 7). The ideal approach to finding the path through the maze is by making a preliminary investigation of the maze in order to envisage a path that does not go down blind alleys. Despite the sensitivity of maze tests in eliciting planning deficits, these tests are not commonly used, perhaps because the original set requires considerable time and some administration challenges.

Porteus Maze Test (Porteus, 1959, 1965)

Three sets of this test are currently in use: the Vineland Revision, which contains 12 mazes for years III through XII, year XIV, and Adult; the eight-maze Porteus Maze Extension, covering years VII through XII, year XIV, and Adult; and the Porteus Maze Supplement, which also has eight mazes for years VII through XII, XIV, and Adult (Porteus, 1965; see Fig. 16.3). The latter two series were developed to compensate for practice effects in retesting so that the maze at each year of the Porteus Maze Extension is a little more difficult than its counterpart in the Vineland Revision and each year of the Porteus Maze Supplement is more difficult than its corresponding test in the Extension series.

To achieve a successful trial, the subject must trace the maze without entering any blind alleys. The mazes range in difficulty from the simplest at year III to the most complex developed for adults. The rule for the number of failures required to discontinue the test varies with the difficulty level, with up to four trials

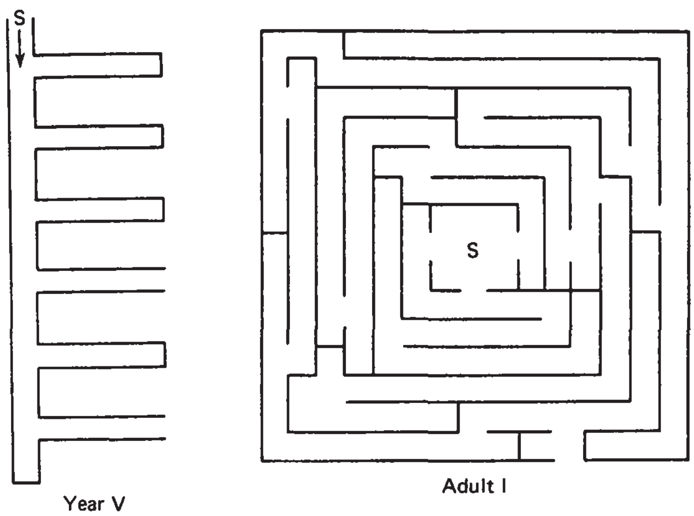

FIGURE 16.3 Two of the Porteus mazes. (Reproduced by permission. © 1933, 1946, 1950 by S.D. Porteus, published by The Psychological Corporation. All rights reserved.)

given on the most difficult mazes. The test is not timed and may take some patients an hour or more to complete all the mazes given to them.

Scores are reported in terms of test age (TA), which is the age level of the most difficult maze the patient completes successfully minus a half-year for every failed trial. The upper score is 17 for success on the adult level maze. Porteus also used eight qualitative error scores: *First Third Errors, Last Third Errors, Wrong Direction, Cut Corner, Cross Line, Lift Pencil, Wavy Line,* and *Total Qualitative Errors.* Other kinds of score have been used. Time to completion scores of frontal leucotomy patients pre- and postoperatively showed that psychosurgery resulted in slowing, and more errors occurred postoperatively as well (Tow, 1955). Subtracting the time to trace over an already drawn path on a similar maze from the time to solution produced a time score free of the motor component of this task (H.S. Levin, Goldstein, Williams, and Eisenberg, 1991). The number of repeated entries into the same blind alley can measure perseverative tendencies (Daigneault et al., 1992).

Test characteristics. Ardila and Rosselli (1989) reported education effects but as many as one-third of their subject group had four or fewer years of school, which raises some question as to the generalizability of these findings. Age effects have shown up in 45- to 65-year-olds as these subjects made more perseverative errors than younger ones (Daigneault et al., 1992). Age effects have also appeared in the 55 to over 76 age range (Ardila and Rosselli, 1989).

Studying older persons, Daigneault and her colleagues (1992) used a battery composed of tests selected for their supposed sensitivity to frontal lobe damage and found that the Porteus Mazes loaded on a "planning" factor. In a much larger battery that included several construction tasks, the Maze test was associated with "visuospatial and visuomotor tasks" (Ardila and Rosselli, 1989). While these findings are suggestive regarding the nature of the Maze tracing task, they also illustrate how much the outcome of factor analyses depends on their input. A moderate correlation (r = .41) exists between performances by children and young adults on the Porteus Maze and the Tower of London, another task with a large planning component (Krikorian et al., 1994). With a young TBI group, Maze test error and time scores correlated significantly with both an untimed test contributing to Daigneault's "planning factor" (Wisconsin Card Sorting Test) and tests of visuomotor tracking (Trail Making Test A and B), implicating sensitivity to executive disorders in all three tasks (Segalowitz, Unsal, and Dywan, 1992). The Mazes error score, along with the other tests, correlated significantly ($p \leq .05$) with a physiological measure of frontal dysfunction.

Neuropsychological findings. The Porteus Maze Test can be quite sensitive to brain disorders. Perhaps the most notable research was undertaken by A. Smith (1960), who did an eight-year follow-up study of psychosurgical patients comparing younger and older groups who had undergone superior or orbital topectomy with younger and older patient controls. Following a score rise in a second preoperative testing, scores on tests taken within three months after surgery were lower than the second preoperative scores in all cases. The superior topectomy group's scores dropped still lower during the eight-year interval to a mean score significantly ($p \leq .05$) lower than the original mean. The control group mean scores climbed slightly following the first and second retest but the eight-year and the original Maze test scores were essentially the same.

Maze test scores have successfully predicted the severity of brain disease (Meier, Ettinger, and Arthur, 1982). Those patients who achieved test age (TA) scores of VIII or above during the first week after a stroke made significant spontaneous gains in lost motor functions, whereas those whose scores fell below this standard showed relatively little spontaneous improvement. In another study, a set of Maze test performances including both brain damaged and intact subjects correlated significantly (r = .77) with scores on actual driving tasks (Sivak et al., 1981). A small group of TBI patients with severe frontal lobe injuries solved the Porteus Mazes more slowly than either TBI patients with severe posterior damage or matched control subjects, this difference holding up even when motor speed was taken into account (H.S. Levin, Goldstein, Williams, and Eisenberg, 1991). Yet 15 of 20 anosmic TBI patients achieved scores above the failure level defined by Porteus (1965); although all of them displayed psychosocial deficits, 16 were reported to have planning problems, and only four were employed two or more years postinjury (Martzke et al., 1991; see p. 626 for a fuller description). Most patients with mild to moderate Alzheimer's disease had low Test Age scores compared to control subjects, although some overlap of scores existed between groups (Mack and Patterson, 1995). Alzheimer patients' Test Age scores correlated with ratings on activities of daily living. These patients also had higher *First Third Errors* and *Last Third Errors.*

Mazes in the Wechsler Intelligence Scales for Children (WISC-R, WISC-III) (Wechsler, 1974, 1991)

These test batteries contain a shorter maze test with time limits and an error scoring system. The most difficult item is almost as complex as the most difficult items in the Porteus series. The highest (15 years 10 months) norms allow the examiner to make a rough estimate of the adequacy of the adult patient's performance. Moreover, the format and time limits make these mazes easy to give. For most clinical purposes, they are a practical and satisfactory substitute for the lengthier Porteus test.

Tower Tests: London, Hanoi, and Toronto

These "brain teasers," familiar to puzzle lovers, get to the heart of planning disorders. To arrive at the best (most direct, fewest moves) solution of the *Tower of London* test, the subject must look ahead to determine the order of moves necessary to rearrange three colored rings or balls from their initial position on two of three upright sticks to a new set of predetermined positions on one or more of the sticks (Shallice, 1982) (see Fig. 16.4). The constraints are that only one piece may be moved at a time, each piece may be moved only from peg to peg, and only a specified number of pieces may be left on each peg at a time. The original task consists of 12 test items of graded levels of difficulty. Levels of difficulty of the test items depend on the number and complexity of subgoals required to achieve the desired arrangement. A problem is scored correct if the solution is achieved with the minimum number of moves necessary. Three trials are allowed for each problem.

Young adults (mean age 21.6 years) correctly solved 92.2% of the problems (Krikorian et al., 1994). Functional imaging has shown a major role for the prefrontal cortex during task performance (S.C. Baker et al., 1996; Lazeron, Rombouts, et al., 2000). Although this test is typically used to measure ability to plan ahead, other factors are important for successful performance such as working memory, response inhibition, and visuospatial memory (D. Carlin et al., 2000; L.H. Phillips et al., 1999; M.C. Welsh et al., 2000).

In an early study of brain injured persons in which the score was the number of correct solutions, patients with predominantly left anterior lesions performed least well while those with either left or right posterior lesions did as well as normal control subjects (Shallice, 1982; Shallice and Burgess, 1991a). Patients in the right anterior lesion group performed less well than control subjects only on the 5-move (most difficult) problems. A more recent study found that patients with frontal lobe lesions and those with frontal lobe dementia had normal planning times (D. Carlin et al., 2000). However, compared to control subjects, patients with focal lesions made more moves, used a trial and error strategy, and were slower to arrive at a solution, while patients with frontal lobe dementia made more moves, committed more rule violations, made more incorrect solutions, and were slower in executing moves. Patients with Huntington's disease also tend to be impaired on this task (L.H. Watkins et al., 2000). TBI patients with anterior lesions performed at essentially the same level as control subjects and, on the most complex item (5 moves), better than those with nonfrontal lesions (H.S. Levin, Goldstein, Williams, and Eisenberg, 1991). The relative insensitivity of this test to the cognitive impairment associated with TBI has been replicated in a sample of patients with severe TBI (Cockburn, 1995).

The *Tower of Hanoi* puzzle is more complex in that, instead of same size pieces, the objects to be rearranged are five rings of varying sizes. The goal and general procedures are the same as for the Tower of London: rings are moved from peg to peg to achieve a final goal with as few moves as possible. As with the Tower of London, only one ring may be moved at a time and any ring not being currently moved must remain on a peg. Instead of a restriction on the number of rings allowed for each peg as for the Tower of London, the restriction for the Tower of Hanoi is that a larger ring may not be placed on a smaller ring. Multiple forms of this puzzle are used and it can be computer-administered. A number of strategies work for achieving the goal; the common strategy requires establishing subgoals and a counterintuitive backward move (Goel and Grafman, 1995; a *subgoal* involves a move that is essential for the solution of the puzzle but does not place a ring into its goal position. A computerized program for this task comes from the U.S. Military Personnel Assessment Battery [Samet and Marshall-Mies, 1987]).

The Towers of London and Hanoi do not measure precisely the same skills (Goel and Grafman, 1995),

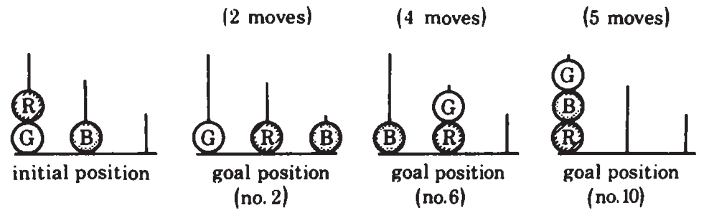

FIGURE 16.4 Tower of London examples. (From Shallice, 1982. Reproduced by permission.)

and correlation between performances on the two tasks is correspondingly low ($r = .37$) (Humes et al., 1997). Goel and Grafman (1995) propose that the Tower of Hanoi does not assess planning as much as it assesses inhibiting a prepotent response (the goal–subgoal conflict). This hypothesis was supported by structural equation modeling data from normal subjects showing that response inhibition contributes to performance (Miyake et al., 2000). Working memory contributes to solutions for medium and hard problems as more subgoal information needs to be kept in mind (Goel, Pullara, and Grafman, 2001; R.G. Morris, Miotto, et al., 1997). Information processing speed also appears to play a role in performances of normal young adults (Bestawros et al., 1999) and patients with multiple sclerosis (Arnett, Rao, Grafman et al., 1997).

At least in the 40- to 79-year age range, neither age nor education appears to affect performance on this task, whether measured by the number of moves required for solution or the number of errors (Glosser and Goodglass, 1990). Participants in their 70s and 80s were significantly impaired compared to those in their 20s and 30s (H.P. Davis and Klebe, 2001). A follow-up of 6.6 years after their first test showed a decline in Tower of Hanoi performance in the elderly group that was not seen on the Rey Auditory Verbal Learning Test. This suggested that for elderly people problem solving declines at a faster rate than some forms of memory. In this same study, patients with anterior lesions tended to do less well than those with posterior lesions. Lateralization differences have also been described. In one study patients with left frontal and right temporal lesions performed worse than control subjects and patients with right frontal and left temporal lesions on four-move problems (R.G. Morris, Miotto, et al., 1997). The left frontal group had larger lesions than the other patient groups, which may have contributed to their poor performance. When Goel and Grafman (1995) compared patients with focal frontal lobe lesions to control subjects they found no differences associated with lesion lateralization. The frontal patients made more errors and appeared to have difficulty choosing a counterintuitive backward move to reach a subgoal. On a simplified version of the Tower of London given to early- and middle-stage Alzheimer patients, along with a lower success rate than their matched control subjects, rule breaking was a prominent feature (Rainville et al., 2002).

The *Tower of Toronto* adds one more layer of complexity—a fourth ring (Saint Cyr and Taylor, 1992). Rather than using rings of different sizes, here the same-size rings have different colors: white, yellow, red, and black. The instructions require the subject to keep lighter colored rings on top of darker ones as they move the set of four blocks from the left one of three pegs to the peg on the right. Saint Cyr and Taylor used this puzzle to examine planning (the development of strategies), learning, and memory for previously developed strategies, by following the initial set of five trials with a second five-trial set $1^1/_2$ hours later. Parkinson patients tended to develop a solution plan slowly, taking and learning an inefficient path that led to a correct solution, and retained that solution on later testing. Amnesic patients performed normally on both learning and retention test trials. Some patients with early stage Huntington's disease also had consistently normal performances, others dealt with the tasks like the Parkinson patients. Late stage Huntington patients' performances were defective on both sets of trials.

Tower of London. Drexel University (TOL$_{DX}$) (Culbertson and Zillmer, 2001)

This formalized test version provides instructions and norms for both children and adults. It uses two boards—one on which the examiner places three colored wooden balls (red, blue, green) in the goal position, and the other containing the three colored wooden balls that the subjects rearrange from a standard "start" position to the examiner's model. Ten problems at each level—child, adult—are given in order of increasing difficulty. Two minutes are allowed for each trial. All ten problems are given. Seven different scores ("indexes") can be obtained for both number of moves and successful completions, and timing aspects. The standardization sample consisted of 264 adults (ages 20–77), of whom 192 were in the 20–29 year old group and only 21 in the 60–77 age range; many of the younger subjects were college students. Theory and interpretation are based on the extensive TOL literature. The format—copying the wooden ball set-up rather than pictures—appears practical; the difficulty levels were essentially defined in prior studies. The instructions are clear and well-detailed as are scoring sheets. Since computerized formats generate data equivalent to older picture-copying formats (Mataix-Cols and Bartres-Faz, 2002), it seems reasonable to expect that this adminstration variation will work well too.

Other tests of planning abilities

Based primarily on theory, several tests in a battery of *Cognitive Processing Tasks* were identified as involving planning (Das and Heemsbergen, 1983; Naglieri and Das, 1988). They include *Visual Search* in which the score is the time to find the target figure (located in a central circle) among many figures (objects, letters) scattered around it; *Trails*, an alternative version of the

Trail Making Test; and *Matching Numbers,* in which subjects have three minutes to find as many pairs of identical numbers as they can on a page with 20 number pairs. Indirect support for the assumption that these are tests of planning comes from factor analyses of this battery in which these three tests cluster together. Also offered as evidence of their sensitivity to planning are correlations with academic achievement tests which are at chance levels for grade 2 but gradually increase to reach the .42 to .55 range at grade 10, presumably reflecting the development of planning ability (Naglieri and Das, 1987).

A more direct approach is to give the patient a task in which planning is a necessary feature. Helm-Estabrooks and her colleagues (1985) played checkers with unilaterally brain damaged patients, recording each move onto "individual checkerboard flow sheets." None of the patients won. Of particular interest were differences between left- and right-lesioned patients as the former made fewer bad moves (losing a checker without taking the opponent's checker in return), appreciated sooner that they would lose, and kept their finger on a moved checker to evaluate the move before committing themselves to it.

Everyday tasks

The abstract nature of many standard tests is different from the planning requirements of ordinary daily activities, such as planning to meet friends, to prepare a meal, or to accomplish a set of errands. These activities present important challenges to many patients with brain disorders. Several methods have been developed to assess the everyday planning skills of patients. Channon and Crawford (1999) devised a series of brief videotapes and stories of everyday awkward situations, such as negotiating a solution with a neighbor about a problem dog. Compared to patients with posterior lesions, anterior patients were more impaired in generating a range of possible solutions to solve the problem and in the quality of the solutions. In another study, patients with focal lesions to the prefrontal cortex and control subjects were asked to plan a response for a hypothetical couple engaged in making real-world financial decisions (Goel, Grafman, Tajik, et al., 1997). The patients with frontal lesions took much longer than control subjects to identify the information that was missing from the problem scenario and less time on the problem-solving phase. They also showed poor judgment regarding the adequacy and completeness of their plans.

Goel and Grafman (2000) examined an architect with a right prefrontal lesion, giving him an architectural task that required him to develop a new design for a lab space. The authors concluded that the patient was impaired in his ability to explore possible alternatives for solutions because of the imprecise and ambiguous characteristics of the design problem. By contrast, he performed well on most standard problem-solving tests, which are more structured with definite rules.

Another patient, this one having sustained a moderate TBI, was asked to devise an emergency management plan in case of weather-related flooding for a hypothetical county (Satish et al., 1999). Using an elaborate interactive computer simulation, a variety of executive skills were assessed. Although the patient was able to plan short-term goals, her decision making and limited use of strategy impaired her overall performance. Her responses on this simulation appeared to explain her postinjury vocational failures and demonstrated specific difficulties that limited her potential.

Script generation (Grafman, Thompson, et al., 1991)

This technique was originally developed to study memory functions but a more recent application investigates the ability to plan a sequence of routine actions. It also appears to have potential value when looking for evidence of breakdown in executive functioning. Applicable "script" topics are those that involve relatively frequent activities undertaken by almost everyone, such as "going to a movie," "eating at a restaurant," or "visiting the doctor." Grafman and his colleagues instructed their probable Alzheimer patients to tell or write "all the things that you do when you get up in the morning until you leave the house or have lunch." Patients' responses were scored for the total number of events in the script, their importance (on a predetermined scale), whether this was a likely event (yes or no), and repetitions (which may or may not be true perseverations).

Dementia patients differed from both depressed elderly patients and normal control subjects in producing many fewer events ($p < .0001$), and more script items given out of order (19% compared to 5% for control subjects with no out of order items for the depressed patients). Dementia patients also made significantly more errors in the other scoring categories. Frontal patients also were impaired on this task. Compared to patients with posterior lesions and control subjects, frontal patients made errors in ordering action in the correct temporal sequence, failed to carry out the script to the stated end point and to remain within the stated boundaries, and made deviant estimates of the importance of specific actions (Sirigu et al., 1995).

More recently these investigators extended the real-world nature of the task by asking subjects to plan actions to be carried out in a "virtual" apartment presented on an interactive computer screen (Zalla et al., 2001). Comparisons between performance on this condition and the verbal script condition found that the realistic context in the "virtual" condition improved the responses of frontal patients. However, in both conditions the frontal group made more errors than control subjects. Frontal patients persisted in carrying out habitual action sequences when inappropriate for the situation at hand and also persisted in including personally relevant but intrusive actions rather than adapting to the test environment.

Allain and his colleagues (2001) not only asked severely injured TBI patients to generate scripts (shop at a supermarket, prepare a salad) following Grafman's model, but then observed them as they engaged in these activities in real life. They found that executive functioning of these patients, all of whom had significant frontal lesions, was impaired, both in script generation and in actual behavior. However, these two aspects of what seemed to be the same task involved different subsets of the executive functions. These patients generated significantly fewer script actions than control subjects and made more script errors, especially sequencing errors. Moreover, when actually performing the tasks, sequencing errors diminished but problems in following regulations, of dependence on help from others, and of distractibility increased. The authors conclude that the cognitive and behavioral responses generated in laboratory studies differ from those elicited by real life.

Purposive Action

The translation of an intention or plan into productive, self-serving activity requires the actor to initiate, maintain, switch, and stop sequences of complex behavior in an orderly and integrated manner. Disturbances in the programming of activity can thwart the carrying out of reasonable plans regardless of motivation, knowledge, or capacity to perform the activity. However, such disturbances are less likely to impede impulsive actions which bypass the planning stages in the action sequence and thereby provide an important distinction between impulsive and consciously deliberate actions. Similarly, Shallice (1982) noted that programming functions are necessary for the successful performance of nonroutine tasks but are not needed when the action sequence is routine. Thus, overlearned, familiar, routine tasks and automatic behaviors can be expected to be much less vulnerable to impaired brain functioning than are nonroutine or novel activities, particularly when the impairment involves the frontal lobes.

Patients who have trouble programming activity may display a marked dissociation between their verbalized intentions and plans and their actions. Hospitalized Korsakoff patients, severely impaired TBI patients who do not always know where they are, and others with profound executive disorders may still talk repeatedly about wanting to leave (to get some money, return to a wife, visit parents, etc.). When informed that they are free to go whenever they wish and even given an explanation of how they might do so, they either quickly forget what they were told, change the subject, or ignore the message. One youthful TBI victim repeatedly announced his very reasonable intention to get a much-needed haircut. Although he knew the way to the barbershop and was physically capable of going there, he never did get his hair cut on his own.

Programming difficulties may affect large-scale purposive activities or the regulation and fine-tuning of discrete intentional acts or complex movements. Patients who have trouble performing discrete actions also tend to have difficulty carrying out broader purposive activities. For example, youthful offenders who displayed an inability to switch ongoing activity by making errors on an untimed trial of the Trail Making Test Part B also tended to be those whose self-report of their criminal activities contained evidence of an inability to make appropriate shifts in the "principle of action" during the commission of the crime (Pontius and Yudowitz, 1980).

Tinkertoy Test (TTT)[1] (Lezak, 1982a)

This construction test gives patients an opportunity—within the necessarily highly structured formal examination—to demonstrate executive capacities. The Tinkertoy Test makes it possible for them to initiate, plan, and structure a potentially complex activity, and to carry it out independently. In the normal course of most neurological or neuropsychological examinations such functions are carried out by the examiner or are made unnecessary (or even unwelcome) by the structured nature of the test material and the restricted number of possible responses in most tests of cognitive functions. Thus, these functions typically remain unexamined, although they are absolutely essential to the maintenance of social independence in a complex society.

The Tinkertoy Test also gives the patient an opportunity to make a "free" construction without the constraints of a model to copy or a predetermined solution. The interplay between executive and constructional functions will more or less limit the extent to which this examination technique tests the constructional capacity of any individual patient. Its usefulness as a constructional test will vary, largely with the patient's productivity.

[1]This combination of 50 wooden pieces comes from "The Classic Tinkertoy Construction Set," Junior size (66 pieces), manufactured by Hasbro, Pautucket, RI 02862. (e-mail: www.tinkertoy.com) It is available in many toy and department stores.

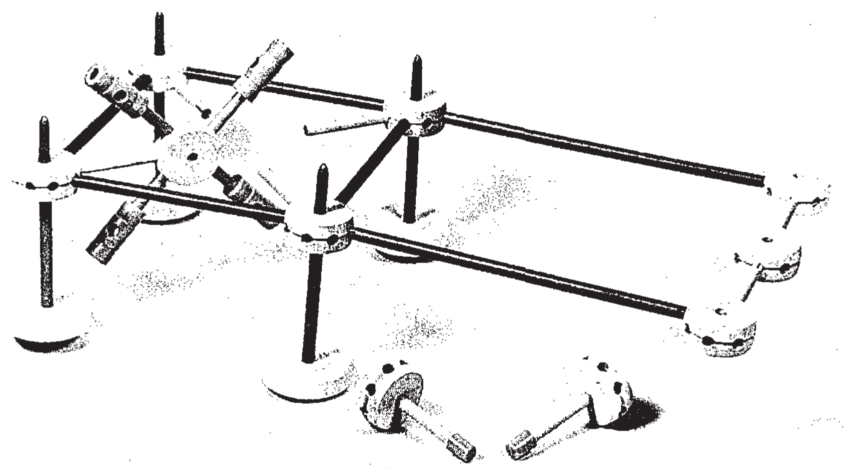

FIGURE 16.5 A 23-year-old craftsman with a high school education made this Tinkertoy "space platform" after he had first tried to construct "a design" and then "a new ride at the fair" (see text).

For example, Figure 16.5 was put together by a youthful TBI patient whose constructional abilities had remained relatively intact (WAIS scaled scores for Block Design = 10, Object Assembly = 14) but whose capacity for integrating complex stimuli was impaired (Picture Arrangement = 6). The ambitiousness, complexity, and relative symmetry of this "space platform" reflect his good constructional skills, although its instability, lack of integration (he could not figure out how to put the two little extra constructions onto the main construction), growth by accretion rather than plan, and the inappropriateness of the name given to it provide concrete evidence of defective executive functioning.

Administration of this test is simple. Fifty pieces of a Tinkertoy set (Table 16.1) are placed on a clean surface in front of the subject, who is told, "Make whatever you want with these. You will have at least five minutes and as much more time as you wish to make something." The necessity for a 5-min minimum time

TABLE 16.1 Items Used in the Tinkertoy Test*

Wooden Dowels	*Rounds*	*Others*
Green (4)	Knobs (10)	Connectors (4)
Orange (4)	Wheels (4)	Caps (4)
Red (4)		Points (4)
Blue (6)		
Yellow (6)		

*Since first used as a test, Tinkertoys have been through several reincarnations and manufacturers. The current sets are colored wood, like the original set. The pieces called for here are the same as those pictured but a little larger.

limit became evident when, without such a limit, bright competitive-minded control subjects did a slapdash job thinking this was a speed test, and poorly motivated or self-deprecating patients gave up easily. Deteriorated patients may stop handling the items after two or three minutes, but should be allowed to sit for several minutes more before being asked whether they have finished with the material. Except for the 5-min minimum, the test is not timed since a pilot study involving both patients and control subjects showed that the amount of time taken may vary without regard to neuropsychological status or with the quality of the performance. Encouragement is given as needed.

Most patients find this test interesting or amusing. Of the 35 subjects with diagnosed neurological disorders who participated in the pilot study, many seemed to enjoy the constructional activity and none raised any objections. Even the one patient who made no construction played with a few pieces, fitting them together and taking them apart, before his attention drifted away. Only blind patients and those sighted patients who cannot manipulate small objects with both hands are unable to take this test.

On completion, the examiner asks what the construction represents (e.g., "What is it?"). If it does represent something (usually a named object), the construction is evaluated for its appropriateness to the indicated name (or concept). In the original scoring system, each of the following criteria earned points, as

TABLE 16.2 Tinkertoy Test: Scoring for Complexity

Variable	*Scoring Criteria*	*Points*
1. *mc*	Any combination of pieces	1
2. *nc*	$n < 20 = 1, < 30 = 2, < 40 = 3, \leq 50 = 4$	1–4
3. *name*	Appropriate = 3; vague/inappropriate = 2; post hoc naming, description = 1; none = 0	0–3
4. *mov*	Mobility = 1, moving parts = 1	0–2
5. *3d*	3-dimensional	1
6. *stand*	Free-standing, stays standing	1
7. *error*	For each error (misfit, incomplete fit, drop and not pick up)	−1
Highest score possible		12
Lowest score possible		−1 or less

noted in Table 16.2 (Lezak, 1982a): (1) whether the patient made any construction(s) (*mc*); (2) total number of pieces used (*np*); (3) whether the construction was given a name appropriate to its appearance and when (*name*); (4a) mobility (wheels that work) and (4b) moving parts (*mov*); (5) whether it has three dimensions (*3d*); (6) whether the construction is freestanding (*stand*); and (7) whether there is a performance error such as misfit in which parts of pieces are forced together that were not made to be combined, incomplete fit in which connections are not properly made, or dropping pieces on the floor without attempting to recover them. The complexity score (*comp*) is based on all of these performance variables (see Table 16.2). A modified complexity score (*mComp*) does not include the number of pieces used. This complexity score (*comp-r*) differs slightly from the one on which the original research is based (*comp-o*). Bayless and his coworkers (1989) suggested the more sensitive 4-point scale for the name category; James L. Mack (personal communication, 1994 [mdl]) identified the poor discriminability of a symmetry score in the original protocol leading to its removal from the scoring system. Regardless of which complexity score is used, findings tend to support the complexity score's sensitivity to impaired executive functions.

An examination of the validity and reliability of the TTT compared the scores from Alzheimer patients and control subjects given by two independent raters (Koss, Patterson, Mack, et al., 1998). Interrater reliability was high. All patient scores were lower than those of control subjects except for *mc* and *error*. Scores also differentiated patients with mild and moderate dementia. They were relatively independent of one another as—for patient performances—only four of 15 possible correlations between scores were significant: *3d* with *np* ($r = .30$), *3d* with *name* ($r = .25$), *3d* with *stand* ($r = .51$), and *np* with *stand* ($r = .30$).

Neuropsychological findings. An initial evaluation of the effectiveness of the Tinkertoy Test in measuring executive capacity was made using the *np* and *comp* scores of 35 unselected patients with cerebral pathology and ten normal control subjects. On the basis of history, records, or family interviews, 18 patients who required total support and supervision were classified as Dependent (D), and 17 were classified as Not Dependent (ND) as the latter managed daily routines on their own and could drive or use public transportation, and five of them were capable of working independently. The two patient groups did not differ in age, education, or scores on Information (WAIS). The control subjects tended to be younger and better educated than the patients.

Both *np* and *comp* scores differentiated the constructions of these three groups (see Table 16.3). All

TABLE 16.3 Comparisons Between Groups on *np* and Complexity Scores

	PATIENT			
Group Measure	*Dependent*	*Nondependent*	*Control*	*F*
np				
Mean ± SD	13.5 ± 9.46	30.24 ± 11.32	42.2 ± 10.03	26.91*
Range	0–42	9–50	23–50	
Complexity				
Mean ± SD	2.22 ± 2.10	5.47 ± 1.77	7.8 ± 1.99	28.27*
Range	−1–8	2–9	5–12	

*One-way ANOVA, $p < .001$.

but one of the Dependent patients used fewer than 23 pieces; those who were Not Dependent used 23 or more. Half of the control group used all 50 pieces but none used fewer than 30. The *np* and *comp* scores of the control subjects and the 19 patients who had age-graded scaled scores of 10 or higher on WAIS Information or Block Design differed significantly. The lower Tinkertoy Test scores of the patients whose cognitive performances were relatively intact suggest that this test taps into more than cognitive abilities. As measured by correlations with the Block Design scaled scores, constructional ability contributes to the complexity of the construction ($r_{\mathrm{comp}\times\mathrm{BD}} = .574$, $p < .01$) but has a weaker association with the number of pieces used ($r_{\mathrm{np}\times\mathrm{BD}} = .379$, $p < .05$).

Other studies also looked at how TTT performances related to tests in common use. For a group of patients with TBI in the mild to moderate range, no relationship appeared between the *comp-r* score and performance on the test of Three-Dimensional Constructional Praxis (Bayless et al., 1989). Among elderly subjects ($M_{\mathrm{age}} = 85.4$ years), of whom half were demented, the TTT performance correlated significantly ($p < .005$) with scores on the Wisconsin Card Sorting Test ($r = .54$) as well as the Trail Making Test ($r = .67$); but correlations between the TTT and tests of visuoperceptual accuracy, psychomotor speed, and vocabulary were in the .21 to .28 range (Mahurin, Flanagan, and Royall, 1993). Differences in levels of correlation between the two sets of tests were interpreted as demonstrating the sensitivity of the TTT as a measure of executive functioning. Mahurin and his colleagues also observed that frail elderly patients whose physical and motivational limitations can preclude most formal testing may still be responsive to the TTT.

A number of executive functions appear to contribute to high scoring constructions, including the abilities to formulate a goal and to plan, initiate, and carry out a complex activity to achieve the goal. In Figure 16.6), "space vehicle," depicts the product of a distinguished neuropsychologist, well known for innovative research. She had never seen Tinkertoys before. Her construction reflects her technical competence, creativity, and well-organized and systematic thinking.

Patients who have difficulty initiating or carrying out purposive activities tend to use relatively few pieces although some make recognizable and appropriately named constructions (e.g., see Fig. 16.7, p. 625), the construction of a 60-year-old, left-handed but right-eyed medically retired plumbing contractor who had had a cerebrovascular accident involving a small area of the left parietal lobe that resulted in transient aphasic symptoms. Those who have an impaired capacity for formulating goals or planning but can initiate activity and are well motivated may use relatively more pieces, but their constructions are more likely to be unnamed or inappropriate for their names and poorly organized (e.g., Fig. 16.5, p. 622). Patients with extensive impairment involving all aspects of the executive functions may pile pieces together or sort them into

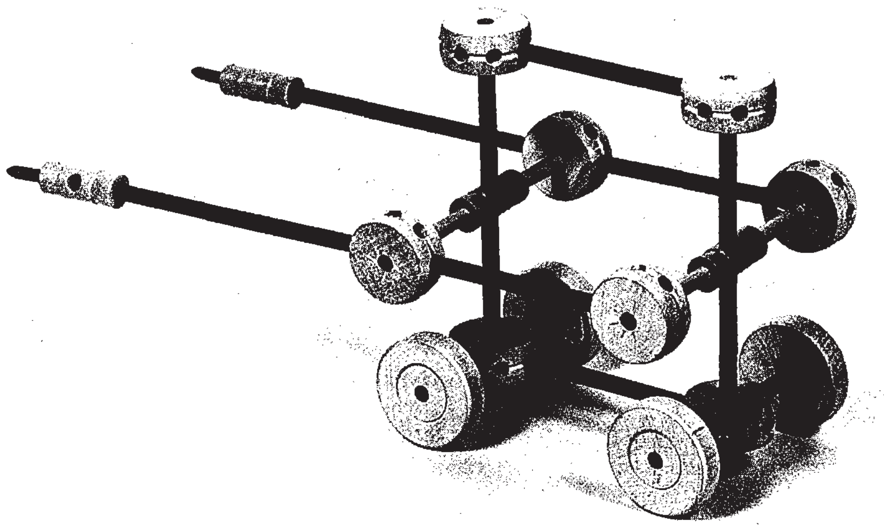

FIGURE 16.6 "Space vehicle" was constructed by a neuropsychologist unfamiliar with Tinkertoys. Although she used only 34 pieces, her complexity score is 11, well above control subjects' mean.

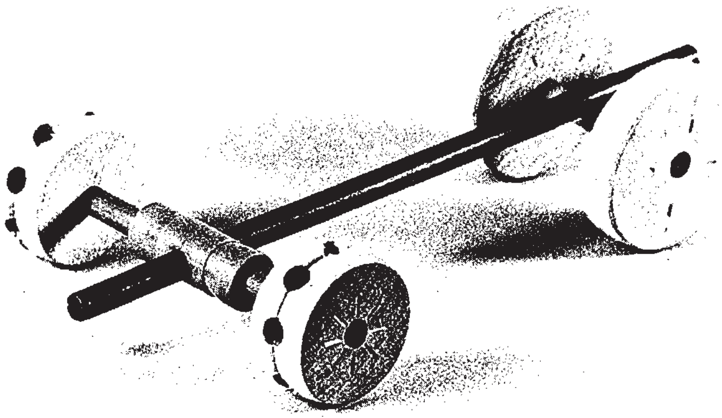

FIGURE 16.7 The creator of this "cannon" achieved WAIS age-graded scaled scores of 16 and 17 on Comprehension and Block Design, respectively (see p. 624).

groups without attempting any constructions, or they use a few pieces to make unnamed and unplanned constructions. For example, Figure 16.8 was the product of a 40-year-old appliance salesman who had suffered a bout of meningitis following a left endarterectomy and thrombectomy undertaken several days after an initial right-sided cerebrovascular accident had resulted in a mild left hemiparesis and slurred speech. Four months after the meningitis had subsided his WAIS Information, Comprehension, and Block Design scaled scores

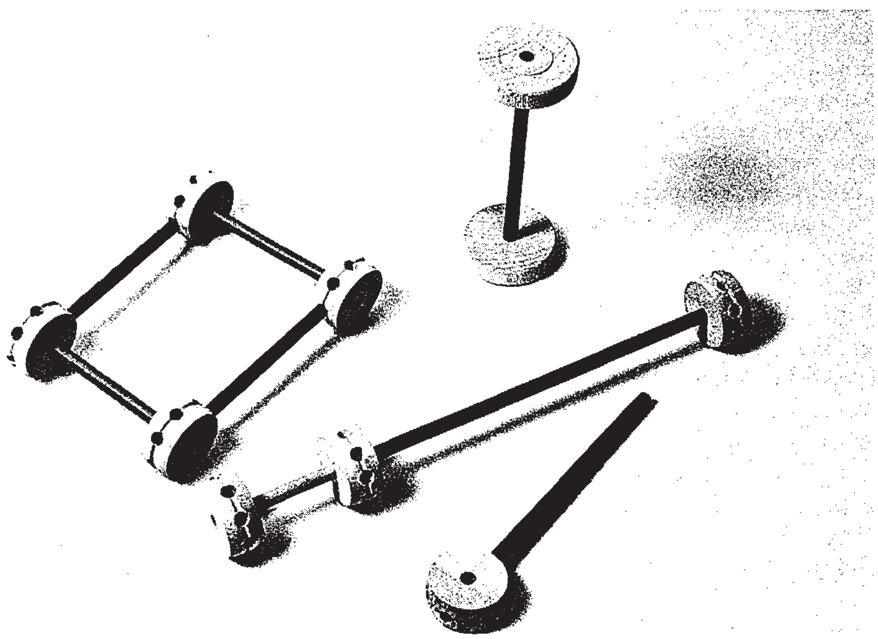

FIGURE 16.8 This patient said he was trying to make "a car" (see text p. 625). He has been totally dependent since the initial illness. Speech is dysfluent; he feeds and toilets himself and walks with a Parkinson-like gait.

were 10, 9, and 6, respectively. Pathologically inert patients, who can usually be coaxed into giving some response to standard test items, are likely to do nothing with as open-ended a task as this.

Studies using the Tinkertoy Test have found the complexity score (original or revised) to be quite sensitive to disorders of executive functions in TBI patients although, for mildly to moderately impaired patients, the score for number of pieces by itself may not be discriminating (Cicerone and DeLuca, 1990). Patients rendered anosmic by TBI typically also sustain orbitofrontal damage with consequent executive function disorders. All 20 such patients examined by Martzke and his colleagues (1991) had psychosocial deficits involving, in most instances, "poor empathy, poor judgment, absent-mindedness," with impaired initiation showing up in many ways. Twelve of them failed this test with *comp-r* scores of 6 or less, although most performed *within normal limits* on other tests purporting to be sensitive to executive functions.

The Tinkertoy Test can be a useful predictor of employability. Only 25 of 50 TBI patients with no physical disabilities were working when examined two or more years after being considered fit to return to work. All but one working patient made scores at or better than the lowest *comp-r* score (7) obtained by 25 normal control subjects; yet 13 of the 25 unemployed patients scored below the lowest control score (Bayless et al., 1989). Tinkertoy Test *comp-o* scores were significantly correlated ($r = .44$, $p < .005$) with postrehabilitation employment status in a study which found that, excepting a correlation of .45 for Trail Making Test-B, the other tests in a representative neuropsychological test battery ran correlations of .35 or less with employment status (Cicerone and DeLuca, 1990). As none of these 87 patients were working or living independently prior to rehabilitation, compared to 38% in supported employment and 40% working competitively afterwards, the Tinkertoy Test and Trail Making Test-B findings suggest that performances on these tests reflect the patients' rehabilitation potential. The Tinkertoy Test also documented improvements made by a patient receiving remediation training for a planning disorder, although scores on Maze tracing and several WAIS-R tests remained essentially unchanged (Cicerone and Wood, 1987).

Tinkertoy constructions may be useful for differentiating between dementia types as 18 patients with multi-infarct dementia achieved a lower *comp-o* score than 18 patients with probable Alzheimer's disease. On most structured tasks, both patient groups performed at the same level, much lower than that of intact elderly subjects (Mendez and Ashla-Mendez, 1991). Their performances differed qualitatively as well: the Alzheimer patients used most pieces but in separate combinations made up of a few pieces, while the multi-infarct patients' constructions were single, simple, and had the fewest pieces. This test is also sensitive to severity of dementia (Koss, Patterson, et al., 1998): mildly impaired Alzheimer patients obtained significantly higher *3d* and *comp* scores than moderately impaired ones.

Self-Regulation

Assessment of self-regulation: 1. productivity

Reduced or erratic productivity can be due to a dissociation between intention and action as well as to weak or absent development of intentions or to a planning defect. This productivity—or inactivity—problem becomes readily apparent in patients who "talk a good game," may even give the details of what needs to be done, but do not carry out what they verbally acknowledge or propose. Patients who do one thing while saying or intending another also display this kind of dissociation. The initiation of an activity may be slow or may require a series of preparatory motions before the patient can make a full response. These patients may make stuttering sounds preparatory to speaking, for example, or agitate the body part that will be undertaking the intended activity before it becomes fully activated. This too is not an intention defect but one of translation from thought to action.

Defective productivity, like many other executive disorders, can usually be observed in the course of an interview or tests of other functions. This requires the examiner to be alert to qualitative aspects of behavior, such as stuttering that heralds the onset of speech, or comments about an error without correction.

Use of standard examination procedures

Slowed responding is probably the most common cause of low productivity in people with brain disorders. It can occur on almost any kind of test, in response latencies and/or performances that are slowed generally or only when certain kinds of functions or activities are called upon. Slowing can and should be documented as it may provide cues to the nature of a disorder which are not apparent in the patient's responses in themselves.

An example of the kind of documentation that provides valuable information about slowing involves responses to a picture shown to elicit a story, the Cookie Theft Picture. Typically responses are evaluated for their linguistic attributes, but timing the rate of responding (words per minute) demonstrated significant differences between patients with multi-infarct dementia, those with probable Alzheimer's disease, and elderly controls (Mendez and Ashla-Mendez, 1991). Response sluggishness also shows up in correct but overtime responses on timed tasks (e.g., Picture Completion, Picture Arrangement, Block Design, and others of the WIS-A batteries.

Patients who are slow to develop a set but whose cognitive functions are intact may achieve quite respectable test scores. Their problem appears only in the first one or two items of a new test, after which they perform well and rapidly. It is typical of these patients when given tests from the WIS-A battery to be slow to solve the easy items of Block Design, to have long latencies on the first few items of Picture Completion or Picture Arrangement, and to give only a few words on the first trial of a word fluency task but perform other trials well. Patients slow to form a set are likely to have a relatively limited recall on the first trial of either the Auditory–Verbal Learning Test or the California Verbal Learning Test word learning tests, but to do well on the interference list since by this time they are familiar with the format.

Another pattern of slowing appears in dwindling responses. The patient begins performing tasks at a rapid enough rate but loses speed and may ultimately stop responding altogether in the course of a trial or set of trials. Tests which require many similar responses to be given rapidly for a minute or more, such as verbal fluency or symbol substitution tasks, are best suited to bring out this production defect.

Assessment of self-regulation: 2. Flexibility and the capacity to shift

The ability to regulate one's own behavior can be demonstrated on tests of flexibility that require the subject to shift a course of thought or action according to the demands of the situation. The capacity for flexibility in behavior extends through perceptual, cognitive, and response dimensions. Defects in mental flexibility show up perceptually in defective scanning and inability to change perceptual set easily. Conceptual inflexibility appears in concrete or rigid approaches to understanding and problem solving, and also as stimulus-bound behavior in which these patients cannot dissociate their responses or pull their attention away from whatever is in their perceptual field or current thoughts (e.g., see Lhermitte, 1983). It may appear as inability to shift perceptual organization, train of thought, or ongoing behavior to meet the varying needs of the moment.

Inflexibility of response results in perseverative, stereotyped, nonadaptive behavior and difficulties in regulating and modulating motor acts. Each of these problems is characterized by an inability to shift behavior readily, to conform behavior to rapidly changing demands on the person. This disturbance in the programming of behavior appears in many different contexts and forms and is typically associated with lesions in the frontal lobes (Luria, 1966; Truelle, Le Gall, et al., 1995). Its particular manifestation depends at least in part on the site of the lesion.

When evaluating performances in which the same response occurs more than once, it is important to distinguish between *perseveration* and *repetitions* due to attentional deficits. As an "involuntary continuation or recurrence of ideas, experiences, or both without the appropriate stimulation" (M.L. Albert, 1989), perseveration involves a "stickiness" in thinking or response due to a breakdown in automatic regulatory mechanisms. Perseverations result from an inability to terminate an activity or switch into another activity (E. Goldberg, 1986). Repetitions made by patients whose abilities for mental and motor flexibility are intact but who have difficulty keeping track of immediately previous or ongoing actions—as for example patients with diffusely impaired brain functioning whose ability to do or think of more than one thing at a time is limited—are not perseverations and should not be labeled as such. This kind of repetition occurs in formal testing, most commonly on word generation tasks: tests of semantic memory (word fluency) or learning ability (word list learning). These patients repeat a word when they have forgotten (lost out of short-term storage or lost to working memory) that they said it 10 or 20 sec before, or they cannot perform a mental task and keep track of what they are doing at the same time. Repetitions will typically differ qualitatively from perseverations as the latter appear in repeated repeating of one word or several, or repeated use of the same word or action with stimuli similar to those that initially elicited the word or action.

By and large, techniques that tend to bring out defects in self-regulation do not have scoring systems or even standardized formats. Neither is necessary or especially desirable. Once perseveration or inability to shift smoothly through a movement, drawing, or speaking sequence shows up, that is evidence enough that the patient is having difficulty with self-regulation. The examiner may then wish to explore the dimensions of the problem: how frequently it occurs, how long it lasts, whether the patient can self-recover (for instance, when perseverating on a word or movement, or when an alternating sequence breaks down), and what conditions are most likely to bring out the dysfunctional response (kind of task, laterality differences [e.g., design copying vs. writing], stress, fatigue, etc.). An efficient examination should be different for each patient as the examiner follows up on the unique set of dysfunctional responses displayed at each step in the course of the examination. When a subtle defect is suspected, for example, the examiner may give a series of tasks of increasing length or complexity. When a broad, very general defect is suspected, it may be unnecessary to give very long or complex tasks but, rather, for planning and rehabilitation purposes, it may be more useful to expose the patient to a wide range of tasks.

At the conceptual level, mental inflexibility can be difficult to identify, shading into personality rigidity on the one hand and stupidity on the other. Tests of abstraction that emphasize shifts in concept formation touch upon mental flexibility (see pp. 574–577, 584–592).

Uses of Objects and Alternate Uses Test (AUT)[1]

This is another kind of test that assesses inflexibility in thinking and has also served to identify creativity in bright children (Getzels and Jackson, 1962; see also Guilford et al., 1978). The printed instructions for the Uses of Objects test ask subjects to write as many uses as they can for five common objects: brick, pencil, paper clip, toothpick, sheet of paper. Two examples are given for each object, such as "Brick—build houses, doorstop," or "Pencil—write, bookmark," with space on the answer sheet for a dozen or more uses to be written in for each object. The Alternate Uses Test version of Uses of Objects provides two sets of three objects each: shoe, button, key; pencil, automobile tire, eyeglasses. One AUT format allows the subject four minutes in which to tell about as many uncommon uses for the three objects in a set as come to mind (Grattan and Eslinger, 1989). Acceptable responses must be conceivable uses that are different from each other and from the common use. Another format allows one minute for each of the six target objects and evaluates performance on the basis of the sum of acceptable responses using the Guilford group's (1978) criteria (R.W. Butler, Rorsman, et al., 1993). Following these scoring rules 17 control subjects ($M_{age} = 40 \pm 8$, $M_{educ} = 14.5 \pm 2$) gave an average of 22 ± 9.5 responses.

The tendency to give obvious, conventional responses such as for *Brick* "to build a wall," or "to line a garden path," reflects a search for the "right" or logical solution, which is called *convergent* thinking. In *divergent* thinking, on the other hand, the subject generates many different and often unique and daring ideas without evident concern for satisfying preconceived notions of what is correct or logical. The divergent thinker, for example, might recommend using a brick as a bedwarmer or for short people to stand on at a parade. Divergent thinking is a sign of cognitive flexibility. Age-related decline in number of uses has been observed in a comparison of younger and older adults (mean ages 48 and 72, respectively) (Parkin and Lawrence, 1994).

Neuropsychological findings. In recommending Uses of Objects to evaluate mental inflexibility, Zangwill (1966) noted that "frontal lobe patients tend to embroider on the main or conventional use of an object, often failing to think up other, less probable uses. This is somewhat reminiscent of the inability to switch from one principle of classification to another" (p. 397).

A 28-year-old man awaiting trial on murder charges had a history of several TBIs in car accidents, untreated and occasionally out-of-control Type 1 diabetes since his teen years, and heavy alcohol and street drug use. Despite only ten years of formal education he achieved scaled scores of 9 and 10 on WAIS-III Information and Comprehension, scores of 12, 11, and 10 on Picture Completion, Picture Arrangement, and Block Design, respectively. His responses to Alternate Uses for *Shoe* were: "*play catch,* look at it, admire it, make footprints, can't think of other things;" for *Button,* responses were "throw it up and down—play catch, magic tricks to make it disappear, collect them, can't think of others." Among other *defective* performances were his bicycle drawing (no spokes, no chain), *Identification of Common Objects* (concrete and premature responses), and *Design Fluency* (seven scorable designs—he named two others "lamp").

None of the Alternate Uses scores achieved by patients with frontal lobe tumors reached the mean of control subjects, and the patients' produced only about half as many acceptable responses as the control subjects ($p < .001$) (R.W. Butler, Rorsman, et al., 1993). Yet 10 of 17 patients in this study performed *within normal limits* on a verbal fluency task (FAS) but the other seven gave far fewer responses ($p < .02$) than control subjects. In a comparison of patients with focal lesions, 89% of control subjects' responses to Uses of Objects were acceptable, patients with posterior cortical lesions gave 68% acceptable responses, and for those with basal ganglia lesions the acceptable response rate was 60% (Eslinger and Grattan, 1993). The worst performers were patients with frontal lesions who gave only 12% acceptable responses. Scores on Alternate Uses correlated significantly ($r = .61$) with a measure of empathy, which was interpreted as demonstrating a relationship between empathy and cognitive flexibility in persons with brain lesions (Grattan and Eslinger, 1989). Productivity in this kind of test tends to decrease with anxiety (Kovacs and Pléh, 1987).

Most studies have reported large standard deviations for group scores. For example, despite large mean differences on this test, between 20 Parkinson patients ($M = 2.9 \pm 9.55$) and their 20 control subjects ($M = 11.3 \pm 10.76$) (Raskin et al., 1992), the even larger

[1]This is one of Guilford's *Measures of Creativity,* which can be ordered from Consulting Psychologists' Press.

standard deviations appear to have obscured some real differences that nonparametric techniques might have documented.

In another set of fluency tasks, *Possible Jobs,* subjects are asked to name jobs associated with pictured objects (e.g., safety pin) or designs (e.g., setting sun) (R.W. Butler, Rorsman, et al., 1993). Another task in this set asks for descriptions of the consequences of unusual situations (e.g., if food were not needed to sustain life). Another calls for drawing elaborations, i.e., adding lines to copies of a figure to make as many different recognizable objects as possible. These tasks, which were identified as "complex" in comparison to the "simple" fluency tasks (Controlled Oral Word Association Test, Design Fluency), proved to be more sensitive to the presence of a frontal lobe tumor than the more traditional "simple" tests of fluency.

Homophone Meaning Generation Test (Warrington, 2000)

This test of flexibility of thinking asks the subject to generate different meanings for common words. Each of the eight words (*form, slip, tick, tip, bear, cent, right,* and *bored*) has at least three distinct meanings. The generation of multiple meanings of these words requires switching among dissimilar verbal concepts. For example, the word "tick" could mean a clock sound or a small insect. The score is the total number of correct meanings produced. The normative sample consisted of 170 participants ages 19 to 74 with a minimum of ten years of education. The total number of words generated ranged from 10 to 35, $M = 23.7$. The test has satisfactory reliability (Crawford and Warrington, 2002). Patients with anterior lesions performed worse than those with posterior lesions but no significant laterality effects appeared (Warrington, 2000). These data are consistent with data from fluency tests in showing that patients with frontal lesions have deficits in generation of concepts and in cognitive flexibility.

Cognitive Bias Task (CBT) (E. Goldberg, Harner, et al., 1994)

On each trial the subject must look at the top one of three vertically placed designs and then choose one of the two designs below it that "you like best." The designs differ on the basis of five binary dimensions: shape (circle/square), color (red/blue), number (one/two identical components), size (large/small), and contour (outline/filled with a homogeneous color). The two lower designs always have different degrees of similarity to the target. Low or high similarity scores indicate that subjects allowed the target to influence their choice (context dependent response), while middle range scores are less clear in revealing subjects' response patterns. There are no right or wrong answers as this procedure asks the subject to indicate a preference.

Despite subjects' freedom to choose any response criterion, a study of 15 healthy subjects found high test–retest reliability ($r = .88$) (E. Goldberg, Harner, et al., 1994). Right-handed male control subjects tended to be more "target-driven" (context dependent) than right-handed women. Response bias in nonright-handed women is comparable with that of right-handed men, i.e., context dependent. Data from small samples of male patients showed that right frontal lesions produced target-driven responses while left frontal lesions gave responses based on perceptual preferences independent of the target. Response bias in female patients was independent of lesion laterality: both right and left frontal lesioned women produced more target-driven responses than their control subjects (E. Goldberg and Podell, 2000). Target-driven behavior may represent a form of stimulus-boundedness.

Item generation

This technique is an effective means of exploring executive functioning. While productivity and the ability to vary one's responses rapidly are essential to success on these tests, other aspects of executive functioning also contribute to good performances, such as self-monitoring (to avoid repeating a response), remembering and following rules, use of strategies, and—of course—creative imagination. The major difference between the two main approaches to this kind of task lies in the degree of structure provided for the subject. See pp. 519–523 for verbal fluency tests.

Design Fluency (Jones-Gotman and Milner, 1977)

This test was developed as a nonverbal counterpart of Thurstone's Word Fluency Test. In the first—free condition—trial, the subject is asked to "invent drawings" that represent neither actual objects nor nameable abstract forms (e.g., geometric shapes) and that are not merely scribbles. After being shown examples of acceptable and unacceptable (e.g., a star is nameable, a scribble or an amoeboid shape requires no thought) drawings made by the examiner, subjects are given five minutes in which to make up as many different kinds of drawings as they can, "many" and "different" being emphasized in the instructions. The first of each type of unacceptable drawing or too similar a drawing is pointed out as is a drawing so elaborate as to decrease the quantity produced. The second, four-minute trial is the fixed (four-line) condition in which acceptable drawings are limited to four lines, straight or curved. Again the subject is shown acceptable and unacceptable examples and the instructions place emphasis on the subject's making as many different drawings as possible. The control subjects' average output on the free five-minute condition was 16.2 designs and on the fixed (four-minute) condition it was 19.7. Approximately 10% of the responses were judged perseverative. Jones-Gotman reported that the free condition is more sensitive than the fixed one (unpublished ms., no date).

Each condition is scored separately but following essentially the same rules. First, all perseverative responses are identified and subtracted from the total. These "include rotations or mirror-imaging versions of previous drawings, variations on a theme, complicated drawings that differ . . . (in) small details, and scribbles. (They) must be scored harshly" (Jones-Gotman, no date). All nameable drawings (in examiner's judgment or named by subject) and four-line condition drawings with more or fewer lines are also removed. The novel output score is then the number of remaining drawings. A perseveration score can be computed by subtracting all other erroneous responses from the total and determining the percentage of perseverative responses out of the remaining subtotal. Reported reliability correlations for interjudge scoring are in the .74 to .87 range (Jones-Gotman, 1991a). Examining the performance of college students, S.L. Carter and her colleagues (1998) reported interrater reliabilities from .66 to .99 except for lower coefficients for nameable errors and the incorrect number of lines in the fixed condition. Good to excellent interrater reliability coefficients also have been reported for a sample of older subjects (ages 51.5 to 89.6 years) (Woodard, Axelrod, and Henry, 1992). However, with some severe impairments, scoring is not only not necessary but not possible (see Fig. 16.9).

A 62-year-old man, born and raised in an Asian country but living in the West for the last 25 years, displayed a radical personality change after a heavy flower pot had fallen on his head from a display shelf. Before the accident he had been a lively, cheerful man, enjoying retirement with other elderly men from his country who frequently fished together and visited with one another. He had been interested in his family, politics, and current events. Since the accident he has been morose, withdrawn, and lacking in spontaneity and interests.

Significant in the history of this graduate engineer was two years' imprisonment after his country's government was overturned. The adjacent cell had been occupied by an army officer who was tortured so severely that he committed suicide by hanging in that cell.

When brought to a psychiatrist for treatment the patient was diagnosed as depressed and put into a six-month outpatient program (in English, in which he was not fluent!) for depressed elderly patients which proved to be ineffectual. In planning the neuropsychological assessment, frontal damage was suspected and Design Fluency included in the battery prepared for him (see Fig. 16.9). He was examined in his second language in which he was fully fluent. After producing the first design (upper left) the instructions were repeated with emphasis on drawing designs that could not be named. His next drawing was to the right of the first. When asked what it was, he said, "bombe." After the third try the test was discontinued: the patient had demonstrated an inability to make up a design due to impaired inventiveness and mental fluency and—even more important for understanding his condition and his obvious misery—loss of the ability to repress his painful, and now obsessive, memories. His personality change seemed best understood as reflecting a compromised capacity for repression plus diminished spontaneity due to a frontal lobe injury. His morbid depression was a symptom, not the cause of this change.

Test characteristics. Several studies used the free condition to examine aging effects on functions associated with the frontal lobes. For the novel output score, Daigneault and her colleagues (1992) found no age effects for subjects in the 15 to 65-year range, but the later study reported a significant tendency ($p = .038$) for perseverative responses to increase with age. With an age range extending up to 75 years, productivity of another group of healthy subjects diminished signifi-

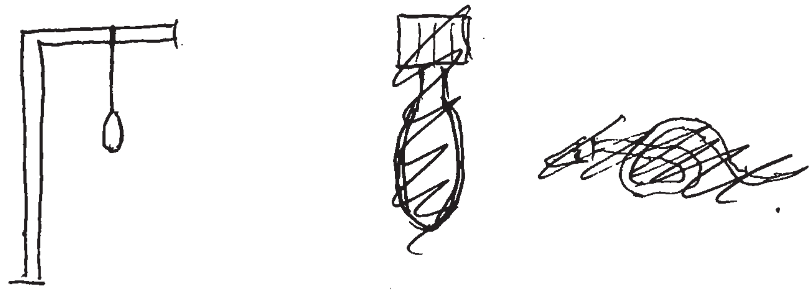

FIGURE 16.9 *Figural Fluency Test* responses by 62-year-old man described on p. 630.

cantly with age (Mittenberg, Seidenberg, et al., 1989), a change interpreted as reflecting a decline in prefrontal functioning.

Neuropsychological findings. TBI patients produced fewer novel designs in the free condition compared to control subjects (Varney, Roberts, et al., 1996). A small sample of TBI patients with frontal lesions made many more nonperseverative errors (rule-breaking) than nonfrontal patients and normal control subjects on the free condition and many more perseverative errors on the fixed condition (H.S. Levin, Goldstein, Williams, and Eisenberg, 1991). Frontal lobe patients tended to have reduced output on both free and fixed conditions relative to normal subjects and patients with posterior lesions. Patients with right-sided lesions generally tended to have lower productivity (except right posterior patients on the free condition), and those with right frontal lesions were least productive. Patients with frontal—particularly right frontal—and right central lesions showed the greatest tendency to perseveration relative to the control group on both free and fixed conditions (Jones-Gotman, 1991a). Studies using the fixed four-line condition found no differences in either novel output score or perseverations for patients with right or left (aphasic) hemisphere disease (M.L. Albert and Sandson, 1986) or Parkinson's disease (Sandson and Albert, 1987). Only a production lag by aphasic patients compared with normal control subjects proved significant at the 5% level.

Five-point Test (Regard, Strauss, and Knapp, 1982)

The use of a structured background for examining response fluency was introduced in the Five-point Test, which consists of a page on which are printed 40 contiguous squares in a 5×8 array, each square containing five symmetrically and identically arranged dots (see example I, Fig. 16.10). The examiner asks the subject to make "as many different figures as possible within 5 minutes by connecting [any number of] the dots with straight lines" without repeating any figure. That the figures should be *different* is emphasized in the instructions.

Age but not sex differences appeared in the 6- to 12-year age range for total production and rotated figures

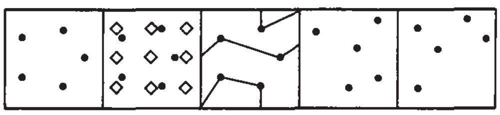

FIGURE 16.10 Ruff Figural Fluency Test (Parts I–V). (From R.W. Evans, Ruff, and Gualtieri, 1985. Reproduced by permission.)

(an indicator of strategy), but from age 10 production levels were in the adult score range. Self-monitoring and self-correcting first appeared among the 10- and 12-year-olds. Patients with psychiatric disorders produced more designs and made fewer perseverative errors than patients with brain disorders on a 3-minute version (G.P. Lee, Strauss, et al., 1997). The number of designs produced did not distinguish frontal from nonfrontal lesioned patients.

Ruff Figural Fluency Test (RFFT)[1] (R.W. Evans et al., 1985; Ruff, Light, and Evans, 1987)

This expanded version of Regard's Five-point Test consists of five sheets of paper, each containing 40 squares. The first is identical in appearance to the Five-point Test sheet. Of the other four, II and III retain dots in the original position but contain interference patterns; the dots on trials IV and V are asymmetrically positioned, with all squares alike on each page (see Fig. 16.10). The instructions are essentially the same as those of the Five-point Test except that the RFFT provides a three-square practice page for each trial, and the allotted time is one minute. In instructing patients, I [mdl] always ask for "patterns" rather than "designs," as used by Ruff and his colleagues, as many people think a design requires artistic talent and may feel unequal to the task. Also, my instructions include emphasizing that the subject can connect *any two or more* dots, as many persons assume that they need to connect all five dots—which precludes development of strategies and slows productivity. Patients who tend to be concrete in their thinking and/or sluggish in altering a response set may continue to connect all five dots in each frame through the first and even the second and third set, despite continuing repetition of "any two or more dots." Performances are scored for number of unique patterns and for number of repetitions of a pattern. Unlike the Five-point Test, rotations are not scored but should be noted, for a series of orderly, nonrepeating rotations is the hallmark of a strategic approach.

Test characteristics. For adults, no sex differences have appeared. However, both age and education affected productivity to a significant degree ($p < .001$) but not accuracy (Ruff, Light, and Evans, 1987). Interrater reliability correlations for unique designs and repetitions are high, .93 and .74, respectively (Berning et al., 1998). Motor skill (as measured by the Finger Tapping Test) may also contribute to a higher productivity rate on the RFFT (R.W. Evans et al., 1985).

[1]This test can be ordered from Psychological Assessment Resources.

A practice effect can be expected on repeat administrations (M.R. Basso, Bornstein, and Lang, 1999).

Neuropsychological findings. The RFFT production score discriminated mild from severe head trauma patients and both groups from normal control subjects (Ruff, Evans, and Marshall, 1986). Inspection of the data shows that the number of repeated patterns increased from control subjects (5.8 ± 7.3) to mildly injured (8.8 ± 14.9) to severely injured (10.1 ± 12.5) patients but intragroup variability was too large for these differences to reach significance using a parametric evaluation. Patients may have difficulty complying with the two key requirements at once: to be productive and to avoid repetitions (which call for continuous self-monitoring). If they are conscientious they produce either many patterns with a few repetition errors or relatively fewer but nonrepeating patterns; if they are not conscientious they go as fast as they can with frequent repetitions and occasional omissions. Generally, the greatest productivity with fewest perseverations is achieved by persons who quickly develop and then maintain a strategy so that each square no longer calls for a unique solution but rather, the pattern for a long series of squares has been predetermined by the strategy. This test also allows the examiner to see concretely the development and/or the disintegration of strategy. Both Alzheimer patients with mild to moderate dementia and Parkinson patients without dementia have reduced Figural Fluency scores (Fama, Sullivan, Shear, et al., 1998).

Design Fluency Test (Delis, Kaplan, and Kramer, 2001)

This design fluency test has three conditions, each allowing one minute for completion (see also p. 637). The first consists of squares with five asymmetrically positioned filled dots. The second condition introduces interference with five additional unfilled dots to be ignored. The third is a switch condition in which the patient is shown squares with five filled and five unfilled dots and asked to draw four straight lines alternating between filled and unfilled dots. Patients with frontal lobe lesions produced fewer designs than control subjects (Baldo et al., 2001). All participants produced fewer designs in the switch condition in which frontal patients did not show a disproportionate cost. No performance differences in side of lesion appeared for frontal patients.

Assessment of perseveration

Perseveration is one of the hallmarks of impaired capacity to shift responses easily and appropriately. Specific types of perseveration tend to appear within one response modality or kind of examination technique but may not show up in a different kind of examination or with a patient whose problems do not involve the modality in question (M.L. Albert and Sandson, 1986; E. Goldberg, 1986; E. Goldberg and Costa, 1986). For example, aphasic patients often produce word substitution errors that are perseverations of a previous response (N. Martin et al., 1998). When perseveration is suspected, or has been observed but needs concrete documentation, many tests can be useful, especially those which involve impaired response modalities.

To test for perseveration, the patient can be asked to copy and maintain alternating letters or patterns (see Fig. 16.11) or repetitive sequential patterns of hand movements with separate trials for each hand to determine whether there are laterality differences in hand control (e.g., see A.-L. Christensen, 1979; Luria, 1966, pp. 677–678). Luria (1966) gave patients a sheet of paper with several word series typed in rows such as "circle, circle, circle, cross, circle" or "square, cross, circle, cross, cross," with instructions to draw the indicated figure below each word as fast as possible (see E. Goldberg, 1986; E. Goldberg and Costa, 1986). Similar chains of verbal commands may also elicit perseverative tendencies. A variety of figures can be named in this manner, including the simple geometric forms, letters, and numbers (e.g., see Sandson and Albert's, 1987, *Stuck in Set Test*). E. Goldberg and Bilder (1987) described seven parameters of graphic figures that can enhance susceptibility to perseveration in subsequent copies of the figures: e.g., *closed/openness* refers to the

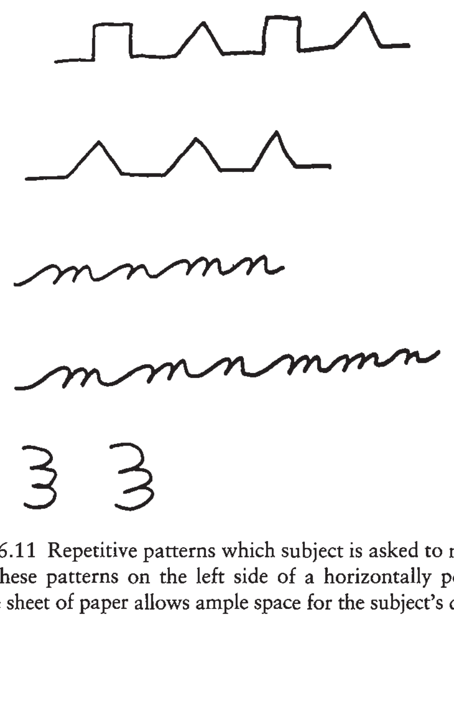

FIGURE 16.11 Repetitive patterns which subject is asked to maintain. Placing these patterns on the left side of a horizontally positioned letter-size sheet of paper allows ample space for the subject's drawings.

tendency to close an open figure (such as a cross) if drawn after a closed one (such as a circle); *straightness/curvedness,* a straight figure (a cross) drawn after a curved one (crescent moon) may be given curved features. Goldberg and Bilder also described four types of perseveration that can occur in simple drawing responses to these kinds of chained verbal commands: (1) *Hyperkinetic-like motor perseveration* refers to inability to terminate an elementary movement that continues in multiple overdrawings of single elements or continuation until stopped by the edge of the page. (2) In *perseveration of elements,* the patient can reproduce discrete elements but introduces elements of previously drawn figures into subsequent ones. (3) *Perseveration of features* involves the perpetuation of some characteristic of a previously drawn figure, such as "openness." (4) In *perseveration of activities,* different categories of stimuli, for example, words and numbers, mathematical and geometrical symbols, become confounded. These authors considered only type 1 to be a true motor perseveration.

Copying and drawing. Tasks that contain within them repeated elements tend to bring out perseverative tendencies; e.g., petals of a flower or many randomly placed lines (as in M.L. Albert's Test of Visual Neglect, p. 379; Sandson and Albert, 1984), rows of dots or circles (Bender-Gestalt Test, particularly cards 1, 2, and 6), or geometric figures (Benton Visual Retention Test). Also sensitive to perseveration are tasks involving writing to command or copying letters, numbers, or words. Perseverative patients often have difficulty in just writing the alphabet, a number series, or their address. Perseveration is least likely to show up in signatures as they are so overpracticed as to be automatic for almost all but the most impaired patients (see Fig. 16.12).

Perseveration in response set. The careful examiner will review all behavior samples, including responses to True-False and multiple-choice tests such as many personality inventories. Frankle (1996) identified a response set, "acquiescent perseveration," that characterized patients with cerebral damage as they tend to have more runs of four or more True responses, including runs of nine or more, than do intact subjects or psychiatric patients without known brain disease. These latter groups rarely even reach, much less exceed, runs of nine.

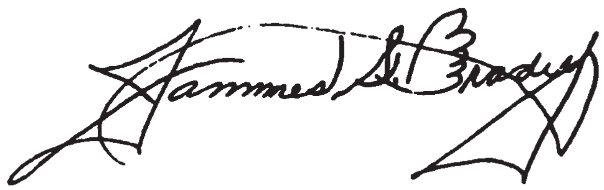

FIGURE 16.12 Signature of middle-aged man who had sustained a gunshot wound to the right frontal lobe.

Examining motor regulation

Luria techniques. Many of the examination techniques in use today were described by Luria (1966), reported by A.-L. Christensen (1979), and incorporated into test protocols (e.g., Grigsby and Kaye, 1996 [see below]; Truelle, Le Gall, et al., 1988). When giving tasks designed to examine the capacity for motor regulation, the examiner must continue them long enough for defective responses to show up. Frequently, patients can maintain the correct response set for the first few sequences and only become confused or slip into a perseverative pattern after that. For example, Malloy, Webster, and Russell (1985) found that when giving Luria's alternating response tests, more than two-thirds of the errors occurred on the last five trials. They caution against giving few trials. If the patient's response deteriorates, the examiner should ask the patient to recall the instructions as patients with frontal lobe damage may be able to repeat the instructions accurately while continuing to respond incorrectly, thus demonstrating a dissociation between comprehension and action.

Of Luria's motor examination tasks, rapid finger sequencing (piano playing) and hand sequencing (fist–edge–palm,[1] see A.L. Christensen, 1979, p. 44) and successive oral movements (e.g., show teeth, stick out tongue, place tongue between lower teeth and lip; ibid., p. 46) were the most sensitive to frontal damage (Truelle, Le Gall, et al., 1995). Copying a series of rapidly presented, paced (by a metronome) hand movements (palm down, one finger out; palm up, all fingers out; fist with hand resting on side) was also sensitive to frontal damage and to temporal lesions (Jason, 1986). Patients with left hemisphere lesions due to stroke tended to have difficulty controlling hand postures, moving rapidly through a repetitive or mixed-movement sequence, and ordering mixed movements sequentially as they were prone to error and perseverations (Harrington and Haaland, 1991a).

Impaired regulation of motor responses can be elicited by tests in which the patient must make converse responses to the examiner's alternating signals (A.-L. Christenson, 1979; Luria, 1966; Luria and Homskaya, 1964). For example, if the examiner taps once, the patient must tap twice and vice versa; or if the ex-

[1]A.S. Kaufman and Kaufman (1983) provide graded difficulty levels from age $2^1/_2$ by varying position presentation and length of sequences (from 2 to 5).

aminer presses a buzzer to give a long signal, the patient must press for a short signal. Patients with self-regulation problems may irresistibly follow the examiner's response pattern.

Withholding responses (the "Go/no-go" paradigm) also examines motor regulation. In these formats the subject must respond to only one of two signals (e.g., squeeze the examiner's hand at the word "red," with instructions to not react when the examiner says "green") (A.-L. Christensen, 1979). This technique and one requiring converse responses (Competing Programs) brought to light motor regulation deficits in Parkinson patients (Raskin, Borod, and Tweedy, 1992).

Behavioral Dyscontrol Scale (BDS)[1] (Grigsby, Kaye, and Robbins, 1992; Grigsby and Kaye, 1996)

This brief bedside examination was designed to measure ability to regulate purposeful behavior, based on the work of Luria. Seven of the nine items measure motor control (including motor sequencing and a go/no-go challenge), one item consists of alphanumeric sequencing, and one rates insight. With the exception of the 4-point scale for the item rating insight, each item is scored on a 3-point scale. The manual provides detailed administration and scoring descriptions. Normative data are based on 1310 adults with a mean age of 74 years. Test–retest reliability, though high, was computed on a sample too small to carry much weight. Interrater reliabilities (obtained on several small samples) ranged from .84 to .98 for individual items, .98 for the total score. Of a large-scale sample (1,313 persons in the 60 to 99 age range), correlations for age and education were −.36 and .52, respectively (Grigsby, Kaye, Shetterly, et al., 2002).

Neuropsychological findings. Although originally designed to assess deficits that trouble many elderly persons, these brief tasks have demonstrated problems in younger patients with chronic progressive multiple sclerosis (Grigsby, Kravcisin, et al., 1993). Studies have related BDS scores to the everyday functional competence of elderly persons. BDS scores correlated significantly with poststroke patients' activities of daily living (ADLs) (e.g., −.38 Using the toilet; −.37 Bathing; −.36 Bed and chair transfers) (Grigsby, Kaye, Kowalsky, and Kramer, 2002). Similar findings (with somewhat higher correlations) are reported for elderly medical and surgical patients (Grigsby, Kaye, Eilertsen, and Kramer, 2000). The BDS also predicted functional independence for a group of 68 mostly geriatric patients who had either orthopedic or neurological disorders three months after hospital discharge (Suchy et al., 1997).

Executive Control Battery (ECB) (E. Goldberg, Podell, Bilder, and Jaeger, 2000)

The four tests in this battery derive from the work of Luria. The *Graphical Sequences Test* involves drawing sequences of simple figures following verbal commands. The items include circle–square–triangle–cross, numerals, and other overlearned objects such as a flower, a car, or a house. Subjects might be asked to "write a number" or "write a number in words." Productions are scored for the presence of four types of perseveration: hyperkinetic motor perseveration and perseveration of elements, features, and activities. The *Manual Postures Test* assesses ability to imitate various asymmetrical hand postures. Errors are scored according to whether they are *echopraxic* (reproducing what is seen rather than making the called-for left–right switch to duplicate the examiner's posture) or visuospatial distortions. The *Motor Sequence Test* requires rapid alteration of both simple and complex unimanual and bimanual motor sequences. Performance is judged for motor perseverations, stereotypies, and other deficits of sequential motor organization. The *Competing Programs Test* examines ability to respond to commands under conflict conditions. Like other tests in the battery, responses are scored for echopraxia, behavioral sterotypies, and disinhibition.

A shortened version of the Graphical Sequences Test has been used to study perseverative behavior in patients with mild dementia from Alzheimer's disease and vascular disease (Lamar et al., 1997). While both groups produced more perseverations than control subjects, the group with vascular disease made more perseverations than the Alzheimer patients.

Frontal Assessment Battery (FAB) (Dubois, Slachevsky, et al., 2000)

This brief set of tests consists of a few items examining each of the following: conceptualization (similarities on three items: *banana–orange, table–chair, tulip–rose–daisy*), item generation (letter fluency: *S*), motor sequencing (Luria's "fist–edge–palm"), sensitivity to interference (conflict task), inhibitory control (go/no-go task), and environmental autonomy (i.e., testing for imitation or utilization behavior). Tasks are described in an appendix to the article. The estimated examination time is 10 minutes. The authors reported good interrater reliability and internal consistency. Total score differentiated patients with frontal lobe dis-

[1]The manual can be ordered from Jim Grigsby, Ph.D., University of Colorado Health Sciences Center, 1355 S. Colorado Blvd, #306, Denver, CO 80222. (e-mail: jim.grigsby©uchsc.edu tasks)

orders from control subjects with 89% accuracy. No comparison was made between patients with anterior and posterior lesions, so the "frontal" sensitivity of this test has not been examined.

Perseverance

Problems with perseverance may also compromise any kind of mental or motor activity. Inability to persevere can result from distractibility, or it may reflect impaired self-control usually associated with frontal lobe damage. In the former case, ongoing behavior is interrupted by some external disturbance; in the latter, dissolution of ongoing activity seems to come from within as the patient loses interest, slows down, or gives up. *Motor impersistence,* the inability to sustain discrete voluntary motor acts on command, tends to occur in those patients with right hemisphere or bilateral cortical damage who display fairly severe mental impairment, although some patients with left hemisphere lesions may also display the phenomenon (Joynt et al., 1962; Kertesz, Nicholson, et al., 1985; Pimental and Kingsbury, 1989). Motor impersistence of Alzheimer patients did not correlate with any specific cognitive domain, such as attention, memory, language, visuoperception, or visuoconstructional abilities (O.L. Lopez, Becker, and Boller, 1991).

The *Motor Impersistence* battery contains eight brief tests with origins in the neurological examination (Joynt et al., 1962; Benton, Siran, Hamsher, et al., 1994): (1) keeping eyes closed; (2) protruding tongue, blindfolded; (3) protruding tongue, eyes open; (4) fixating gaze in lateral visual fields; (5) keeping mouth open; (6) fixating on examiner's nose (during confrontation testing of visual fields); (7) sustaining "ah" sound; and (8) maintaining grip. Motor impersistence may also show up when patients are asked to hold their breath. Of course, not all impersistent patients will fail all eight tests. Only three patients of 24 left hemiplegic patients failed tongue protrusion while 20 could not maintain a fixated gaze (Ben-Yishay, Diller, Gerstman, and Haas, 1968). The proportion of patients failing these tasks increased in the same order as task difficulty, as determined by Joynt and his coworkers (1962). Such an orderly progression suggests a common underlying deficit that occurs with varying degrees of severity. The number of tests failed also reflected severity of impairment as documented by measurements of cognitive abilities, visuomotor efficiency, and functional competence (Ben-Yishay, Diller, Gerstman, and Haas, 1968). Limb impersistence, demonstrated when the patient cannot maintain arm extension for 20 secs., may be lateralized (Heilman and Watson, 1991). When impersistence is pronounced in patients with right hemisphere dysfunction, rehabilitation prospects are poor (Joynt and Goldstein, 1975).

Effective Performance

A performance is as effective as the performer's ability to monitor, self-correct, and regulate the intensity, tempo, and other qualitative aspects of delivery. Patients with brain disorders often perform erratically and unsuccessfully since abilities for self-correction and self-monitoring are vulnerable to many different kinds of brain damage. Some patients cannot correct their mistakes because they do not perceive them. Patients with pathological inertia may perceive their errors, even identify them, and yet do nothing to correct them. Defective self-monitoring can spoil any kind of performance, showing up in such diverse ways as unmowed patches in an otherwise manicured lawn, one or two missed numbers in an account book, or shoelaces that snapped and buttons that popped from too much pressure.

Testing performance effectiveness

While few examination techniques have been developed for the express purpose of studying self-monitoring or self-correcting behavior, all test performances provide information about how the subject responds. The nature of the patient's errors, attitudes (including awareness and judgment of errors), idiosyncratic distortions, and compensatory efforts will often give more practical information about the patient than test scores that can mask either defects or compensatory strengths. In a neuropsychological examination, self-monitoring defects may appear in cramped writing that leaves little or no space between words or veers off the horizontal; in missed or slipped (e.g., answers to item 9 on line 10) responses on paper-and-pencil tests; in speech that comes in quick little bursts or a monotonic, unpunctuated delivery; and in incomplete sentences and thoughts that trail off or are disconnected or easily disrupted by internal or external distractions. Tests in which subjects can check their written responses for accuracy as they are working on them, such as arithmetic calculations, symbol substitution tests, and drawing fluency tasks, will readily expose poor self-monitoring.

Random Generation Task (Baddeley, 1966, 1986)

The subject is asked to generate a sequence of 100 letters of the alphabet in random order. Initial studies used four rates of generation: 0.5, 1.0, 2.0, and 4.0 sec. Although the task may sound easy, even normal subjects find it difficult to avoid either stereotyped sequences (e.g., *X–Y–Z*) or common combinations (e.g., *F–B–I*) or omitting responses. Output was measured in three ways: the frequency of each letter, which detected redundancy in the output; the frequency of letter pairs (*digrams*), again looking for redundancy; and the frequency of letter pairs in alphabetical sequence (*stereotyped digrams*). A small group of control subjects consistently increased randomness of output as generation rates slowed. As in fluency tests which forbid repetition, self-monitoring is necessary for success on this task; when self-monitoring fails, this task will be failed.

Other random generation tasks include random generation of numbers based on the work of F.J. Evans (1978). In the initial version, subjects are instructed to say numbers from 1 to 10 in random order at the same rate as a metronome's beat (1/sec). The sequence of 100 numbers is recorded. A significant practice effect has been reported (Marisi and Travlos, 1992). Since most subjects develop a strategy after a brief experience with this task, the authors recommend discarding first trial data. Another variation is the *Mental Dice Task* in which subjects are asked to "call out digits (from 1 to 6) in a sequence as random as possible" such as might appear when rolling a die repeatedly (Brugger, Monsch, et al., 1996). Even moderately demented patients were able to understand this instruction. As with letter generation, the task requires suppression of habitual and stereotyped responses plus response monitoring (Miyake et al., 2000).

Various random generation tasks have been used to examine patients with brain disorders. Clinical studies suggest that patients with frontal lobe dysfunction are impaired on these tasks as they demonstrate a strong sequential response bias (Brugger, 1997). Deficits have been observed in patients with anterior communicating artery aneurysms (Leclercq et al., 2000), Korsakoff's syndrome (Pollux et al., 1995), and degenerative diseases such as Alzheimer's (Brugger, Monsch, et al., 1996) and Parkinson's (R.G. Brown, Soliveri, and Jahanshahi, 1998). Transcranial magnetic stimulation over the left dorsolateral prefrontal cortex interfered with randomness on a letter generation task (Jahanshahi and Dirnberger, 1999).

Executive Functions: Wide Range Assessment

Some techniques for examining executive functions involve so many of them that they defy classification under any one of the subdivisions. Of course, naturalistic observation is chief among these; but since few examiners have the time and resources to spend the hours or days needed to know the status of their patients' executive behavior, these clinical methods may serve as useful substitutes for the real thing.

Behavioural Assessment of the Dysexecutive Syndrome (BADS) (B.A. Wilson, Alderman, et al., 1996)

This set of tests was developed to examine performance on a wide range of real-world tasks. All but one of the six tests are question-and-answer or paper-and-pencil tests; much of the content will be familiar to most subjects. The *Rule Shift Cards* test assesses flexibility by having subjects view playing cards and respond under two "rule" conditions: in the first condition the subject is instructed to say "yes" if a presented card is red and "no" if it is black. After a series of cards has been shown, the instructions change such that, for each card, the subject should say "yes" if it is the same color as the previous card and "no" if it is not. The *Action Program Test* instructs subjects to figure a way to get a cork out of a tube with a variety of objects at their disposal. Subjects must develop a plan and then try to get the cork out by manipulating materials such as water in a beaker and a metal hook. The *Key Search Test* asks the subject to draw a plan for finding a lost key in a square-shaped area (see also *Plan of Search* [Terman and Merrill, 1973]). The *Temporal Judgement Test* asks four questions concerning estimations on how long activities take. Planning an effective route through a zoo in order to visit certain sites is assessed with the *Zoo Map Test.* Planning and priority setting are also assessed with the *Modified Six Elements Test* in which subjects are instructed to complete as many paper-and-pencil tasks (e.g., simple calculations, naming pictures) as possible in a brief time while attempting at least something from each of the test's six parts. Performance is judged by how well subjects organize their time. The *Dysexecutive Questionnaire* (*DEX*) is a 20-item symptom checklist which both the patient and a collateral source complete. The manual reports very high interrater reliability (ranging from .88 on Temporal Judgement to 1.00 on the number of tasks passed independently). Test–retest reliability was best for the Action Program, Key Search, and Temporal Judgement tests (r = .64–.71), lower for the other tests.

Neuropsychological findings. Comparing means and standard deviations of a group of patients with brain disorders (primarily TBI) with a group of control subjects, all of the tests showed group differences although the Key Search difference was only marginal. Norris and Tate (2000) also compared a patient group with control subjects using nonparametric analysis because the variables were not normally distributed. Group differences appeared for Action Program, Zoo Map, Modified Six Elements, and the total profile. The remaining three tests did not show group differences. These groups differed on only two of the commonly used tests purporting to examine executive skills: the Porteus Mazes and Controlled Oral Word Association Test. They showed no differences on the Wisconsin Card Sorting Test (WCST), Trail Making Test, Complex Figure Test, and Cognitive Estimation Test. The BADS total score correctly classified 84% of control subjects and 64% of patients, which compared favorably with the commonly used tests (81% and 64%, respectively). Concurrent validity was adequate for the Rule Shift and Action Plan tests and the total profile score. Schizophrenic patients also are impaired on the BADS (Krabbendam et al., 1999; B.A. Wilson, Evans, Emslie, et al., 1998).

Delis-Kaplan Executive Function System (D-KEFS) (Delis, Kaplan, and Kramer, 2001)

The D-KEFS is a set of nine tests, each intended to stand alone. There is no composite score. The authors state that tests were selected to be sensitive to many of the types of executive impairment seen in patients with brain disorders. However, no theoretical rationale, other than inclusion of both verbal and nonverbal tests, is provided for their selection. D-KEFS's *Trail Making Test, Verbal Fluency* (letter and category), *Design Fluency, Color-Word Interference Test* (Stroop conflict task), *Sorting Test, Twenty Questions Test, Tower Test,* and *Proverb Test* are, for the most part, variations on the most commonly used tests purporting to examine executive function. The *Word Context Test* was developed by Dr. Kaplan in the 1940s to test children's understanding of words. Card-Sorting, Twenty Questions, Proverb, and Word Context tests are discussed in Chapter 15; Design Fluency is reviewed in this chapter. The authors state that one advantage of the D-KEFS tests is that they are co-normed on 1750 participants ranging in age from 8 to 89. An alternate form with normative data from a sample of 295 subjects is available for Verbal Fluency, Sorting, and Twenty Questions. Raw scores can be converted to standard scores or, in some cases, cumulative percentile ranks. Scoring software is available for $163.

Many of the standard tests have been lengthened in this version, with additional easy and difficult items to avoid ceiling and floor effects, and with subtests that break down performances into the fundamental components required for success on these complex tasks. One can compute an astounding number of scores for many of the tests. The Trail Making Test alone has 12 primary measures and 12 optional ones. The Sorting Test has five primary measures and 29 additional ones. Whether the many additional features are worth the patient's extra effort and the examiner's extra time for administration and scoring has not been established.

Whereas the principal scores generally have acceptable reliability, the additonal D-KEFS scores often have low reliability which varies across age groups. For example, the manual reports that the newly added switching condition of the Design Fluency test has test–retest reliabilities varying with age groups from .13 to .58. Scores on the second testing were higher and appear to represent a practice effect. Although scores from the more standard conditions for this test also showed a practice effect, test–retest correlations were considerably higher (ranging from .43 to .73). For some of the nine tests, standard deviations of test–retest scores were larger for the second testing, which suggests weak reliability. Internal consistency scores, calculated for some tests, varied by age group. Some examples of the median internal consistency scores for the entire normative group are: Verbal Fluency Test, Category Switching condition = .54; Sorting Test, Free Sorting Confirmed condition = .78; Twenty Questions Test, Total Weighted Achievement score = .45; Tower Test, Total Achievement score = .61; and Proverb Test, Total Achievement score, Free Inquiry condition = .78. Overall, the internal consistency reliability coefficients are lower than might be desired. Correlations between performances of a small sample on the D-KEFS tests and the Wisconson Card Sorting Test produced coefficients ranging from .30 to .60 for WCST "categories achieved" and .20 to .71 for WCST "perseverative responses." The strongest correlation was between the Proverb Test Total Achievement score and the WCST perseverative responses.

The clinical usefulness of these modifications of familiar tests is largely unknown. The manual presents data for nine Alzheimer and nine Huntington patients. In a study using the verbal and design fluency tests, participants were tested in the standard manner and for their ability to switch back and forth between different sets (e.g., alternating between naming exemplars of fruit and furniture for the verbal fluency tests) (Baldo, Shimamura, Delis, et al., 2001). Patients with focal frontal lobe lesions did not have more difficulty than control subjects with the new switching condition for either fluency test. By far the most studied of the tests is the Card Sorting Test; a similar version—the California Card Sorting Test—has been in use since 1992 (Delis, Squire, et al., 1992; see also pp. 585–586). Compared to control subjects, multiple sclerosis patients made fewer attempted sorts and fewer correct sorts (Beatty and Monson, 1996). In this study, Correct Free Sorts showed a modest correlation (.64) with categories achieved on the WCST. Perseverative responding on the two tests did not appear to be related ($r = .15$). Patients with frontal lesions produced fewer attempted sorts, correct sorts, and correct sort descriptions than control subjects, while Parkinson patients differed from control subjects only in making more perseverative sorts (Dimitrov et al., 1999).

Standard versions of the Trail Making Test, verbal fluency, twenty questions, and proverb interpretation are in the public domain.

Executive Function Route-Finding Task (EFRT) (Boyd and Sautter, 1993)

To accomplish this task subjects must find their way from a starting point to a predetermined destination within the building complex in which the examination is given (see also Sohlberg and Mateer, 2001). For a practical level of difficulty, the final destination must be a minimum of five choice points and one change in floor level away from the starting place. Ideally, there will be signs giving directions for the destination.

For example, my [mdl] patients begin on the third floor of the clinic building and have as their goal the cafeteria in University Hospital South across a street (by way of the street or an enclosed bridge) and at least five choice points away (first corridor: right, left, or straight; if right, elevator, stairs, or corridor, etc.) and, while also on a third floor, this cafeteria is seven floors below the clinic third floor (the hospital and clinics are built on a hill). The clinic building has numerous signs indicating the direction to Hospital South; Hospital South has signs for the cafeteria. Halls and elevators are full of both visitors and medical center personnel providing ample opportunities for the patient to ask directions.

While accompanying the patient the examiner records the path taken and how the patient gets there. The examiner also answers questions and gives encouragement and advice as needed, noting these too. After reaching the destination, the examiner may need to question the patient further to clarify whether moves were made by chance, what cues the patient used to find the way, etc. Performances are rated on a 4-point scale to measure the degree to which the patient was dependent on the examiner for (1) understanding the task; (2) seeking information; (3) remembering instructions; (4) detecting errors; (5) correcting errors; and (6) ability to stick with the task (on-task behavior).

Two examiners participated in the feasibility study with high ($r = .94$) interrater reliability indicating that this is a very scorable task. Scores obtained by 31 rehabilitation patients with varying degrees of TBI severity correlated well ($p < .01$) with both the Verbal Comprehension and Perceptual Organization factor scores of the WAIS-R and a shortened form of the Booklet Category Test. In general, these patients were mostly dependent on nonspecific executive cues (e.g., examiner questioning guided patients on how they might begin or what information they needed next) but they also required directed cueing.

Spikman and her colleagues (2000) used this task to study planning by patients with chronic TBI of at least moderate severity. The two scores they obtained were the number of times patients needed cues and a score combining the adequacy of information seeking, error detection, and error correction. Like the original study, they found high interrater reliability ($r > .90$). The EFRT was the only one of a number of executive tasks on which the patients performed significantly worse than control subjects. Patients with documented frontal lesions had even more difficulty than those without frontal damage. The authors attributed this test's sensitivity to its lack of structure and the need for participants to seek information and to detect and correct errors.

Behavioral Assessment for Vocational Skills (BAVS): Wheelbarrow Test (R.W. Butler, Anderson, et al., 1989)

This ingenious and quite naturalistic examination technique requires the subject to assemble the parts of a mail-order wheelbarrow within a 45 minute period. The clinicians who rate the performance also play the role of job supervisor and, although they offer as little structure as possible, they can become more directive if the subject's limitations require help to stay on task or complete it. Distractibility problems are elicited by interjecting a "brief alternate task" and then redirecting the subject's attention back to the wheelbarrow. A rater/supervisor also gives one constructive criticism in response to an error to see how the subject deals with criticism. Performances are rated on a 5-point scale for 16 vocationally relevant aspects, such as following directions, problem solving, emotional control, judgment, and dependability.

Ratings on this task did not correlate significantly with visuospatial test scores, visual tracking (Trail Making Test), or the Wisconsin Card Sorting Test. However, they did predict the levels of three categories of work performances by 20 TBI patients in volunteer trial work settings: work quantity ($r = .74$), work quality ($r = .75$), and work-related behavior ($r = .64$) (all correlations were significant at $p < .01$).

MOTOR PERFORMANCE[1]

Distinctions between disturbances of motor behavior resulting from a supramodal executive dysfunction and specific disorders of motor functions are clearer in the telling than in fact. A defective sequence of alternating

[1]We are grateful to Kathleen Y. Haaland for her helpful contributions to this section.

hand movements, for example, may occur—with a cortical lesion—as a specific disability of motor coordination or it may reflect perseveration or inability to sustain a motor pattern; or it may be a symptom of subcortical rather than cortical pathology (Heilman and Rothi, 2003). Some diagnostic discriminations can be made from observations of the defective movement, but the classification of a particular disability may also depend on whether the pattern of associated symptoms implicates a cerebellar or a frontal lesion, whether the disorder appears bilaterally or involves one side only, or whether it may reflect a sensory deficit or motor weakness rather than a disorder of movement per se. Many motor disorders that accompany cerebral brain damage cannot, by themselves, necessarily be linked with particular anatomic areas.

Examining for Apraxia

The examination for apraxia reviews a variety of learned movements of the face, the limbs, and—less often—the body (Goodglass, Kaplan, and Barresi, 2000; Heilman and Rothi, 2003; Strub and Black, 2000). The integrity of learned movements of the face and limbs, particularly the hands, is typically examined under two conditions: *imitation* of the examiner (*a*) making symbolic or communicative movements, such as familiar gestures; (*b*) using actual objects; or (*c*) pantomiming their use without objects; and *to command* for each of these three kinds of activity. A tactile modality can be introduced by blindfolding patients and handing them such familiar objects as a glass, a screwdriver, a key, or a comb, with instructions to "show me how you would use it" (De Renzi, Faglioni, and Sorgato, 1982). Table 16.4 lists activities that have been used in examinations for apraxia. The examiner may demonstrate each activity for imitation or direct its performance, asking the subject to "do what you see me doing" or "show me how you . . . . " Some of these activities should involve the use of objects, either with the object or in pantomime . The examiner should be alert to those patients who are not apraxic but, when pantomiming to command, use their hand as if it were the tool (e.g., hammering with their fists, cutting with fingers opening and closing like scissors blades). The concreteness of their response reflects their concreteness of thought. This use of a body part as object occurs more commonly among brain damaged patients without regard to lesion laterality than in neurologically intact persons (Mozaz et al., 1993).

The difficulty in knowing just what to score and how to score it probably explains why no scoring system has achieved general acceptance. Five different systems give some idea of the range of scoring possibilities:

1. Haaland and Flaherty (1984) developed a scoring system for a 15-item battery of movements to imitation: five transitive movements (e.g., brush teeth), five intransitive movements (e.g., salute), and five meaningless movements (e.g., index finger to ear lobe). They recorded errors in hand position, arm position, and target. Patients are designated "apraxic" if they make four or more errors on this 15-item battery (i.e., 2 SD below control subjects' mean) (Haaland, Harrington, and Knight, 2000). Normative data for 75 control subjects are available.

2. A 14-category scoring system which takes into account errors of content, of timing (including sequencing), of a spatial nature (e.g., change in amplitude of movements, body-part-as-object), and of "other" error (including no response) brought out six error types occurring most typically with left cortical lesions: they involved spatial distortions—including body-part-as-object; incorrect spatial relationships between the hand and fingers; incorrect spatial relationships between the hand and the imagined object; incorrect movement with the imagined object; changes in number of movements normally called for; and correct response to the wrong target

TABLE 16.4 Activities for Examining Practic Functions

	Use of Objects	*Symbolic Gestures*	*Other*
Face (Buccofacial)	Blow out match	Stick out tongue	Whistle
	Suck on straw	Blow a kiss	Show teeth
Upper Limb	Use toothbrush	Salute	Snap fingers
	Hammer nail	Hitchhike	Touch ear with index finger
	Cut paper	"OK" sign	Hold up thumb and little finger
	Flip coin	"Stop" sign	Make a fist
Lower Limb	Kick ball		
	Put out cigarette		
Whole Body	Swing baseball bat	Bow	Stand (or sit)
	Sweep with broom	Stand like boxer	Turn around
Serial Acts (can be done in pantomine or with real objects)	Preparing a letter for mailing (fold letter, put in envelope, seal and stamp envelope)		

(e.g., combing movements for "hairbrush") (Rothi, Mack, Verfaellie, et al., 1988).

3. Poeck (1986) offered a five-part assessment scheme based on a qualitative analysis of errors for a lengthy series of movements: correct execution, augmentation phenomena, fragmentary movement, perseveration, and other types of errors. The number of perseverations is not scored as they tend to occur as intrusive motor elements of the perseverated movement rather than in the original complete form of the movement. (This observation may account for Rothi, Mack, and their colleagues' report that perseveration errors occurred too rarely for consideration, since they did not provide a scoring category for partial perseverations.)

4. Another scoring system gives 3, 2, or 1 points to a correct imitation made on a first, second, or third trial, respectively, and no points when the patient does not achieve the correct movement within three trials (De Renzi, Motti, and Nichelli, 1980). Thus, with a 24-item protocol, the maximum possible score is 72.

5. Based on good interrater agreement, and most practical for clinical work, Goodglass, Kaplan, and Baressi (2000) offer a 3-point judgment of "normal," "partially adequate," and "failed" which can be expanded to four points: "perfect," "adequate," "partially adequate," and "inadequate" (Borod, Fitzpatrick, et al., 1989).

Task characteristics. Age tends to have some effect on the quality of pantomimed movements as a substantial portion of over-60 healthy subjects may make body-part-as-object responses (L. Willis et al., 1998). The number of these and other, less frequent errors varies with the task (Ska and Nespoulos, 1987, 1988b); for example, some movements (e.g., scratching one's back) become more difficult as flexibility diminishes (Ska and Nespoulos, 1986). That R.J. Duffy and Duffy (1989) found no difference in the frequency of body-part-as-object responses between patients with right and those with left lateralized brain lesions and normal control subjects, all compared in groups in which the average age was over 60, suggests that age may be more of a determinant in the appearance of this error type than lesion presence or lateralization. The range of activities tested enables the examiner to assess the extent and severity of the disorder. Apraxia is more common for transitive movements (object use) than other movements (intransitive, meaningless), which may relate to the complexity of these movements or their dependence on object use (Haaland and Flaherty, 1984).

Neuropsychological findings. Apraxia may occur as the result of focal lesions or degenerative diseases. Among patients with unilateral lesions, most apraxias of use and gesture affect both sides of the body but typically occur with lesions in the left cerebral cortex (De Renzi, 1990; Schnider et al., 1997). Studying stroke patients with lesions in anterior or posterior regions, Haaland, Harrington, and Knight (2000) found that those with *ideomotor limb apraxia* (inability to make correct gestures on command) "had damage lateralized to a left hemispheric network involving the middle frontal gyrus and intraparietal sulcus region." This finding supports the importance of the frontoparietal circuits in reaching and grasping movements (Heilman and Rothi, 2003). In an earlier study, Haaland and Yeo (1989) noted that patients with left parietal lesions are impaired on hand posture tasks, whether single or sequenced, but impaired performances by frontal patients were reported only in some studies.

Degenerative disorders such as Alzheimer's disease, Parkinson's disease, Huntington's disease, and corticobasal degeneration may also produce apraxia (Hamilton et al., 2003; D.H. Jacobs et al., 1999; Leiguarda et al., 1997; R.L. Schwartz, 2000; L. Willis et al., 1998).

Apraxia may occur in only one or two modalities, usually with visual (imitation) or verbal (command) presentation; rarely will apraxia be purely tactile (De Renzi, Faglioni, and Sorgato, 1982). While failure is more likely in the command than the imitation condition (Goodglass and Kaplan, 1983), the opposite can occur (Rothi, Mack, and Heilman, 1986; Rothi, Ochipa, and Heilman, 1991). Patients exhibiting apraxia on a test will also tend to have reduced recourse to gestural communication (Borod, Fitzpatrick, et al., 1989). Testing for movement imitation and oral apraxia over periods greater than two years from onset, A. Basso and her colleagues (2000) reported that all but one of 14 patients improved significantly during the first year after onset. Little further improvement occurred and six worsened after the first year.

Florida Apraxia Screening Test-Revised (FAST-R) (Rothi, Raymer, and Heilman, 1997)

This revision of the original test consists of 30 verbal commands to demonstrate gestures. Twenty items involve object use (transitive) and ten require meaningful, tool-free gestures (intransitive) such as , "Show me how you salute." All items can be completed with one arm/hand; usually, the dominant hand is examined. A practice trial shows the patient the expected degree of precision and elaboration of movement. Productions are scored for content, temporal features, and spatial features. The score is the number of items performed correctly.

Florida Action Recall Test (FLART) (R.L. Schwartz et al., 2000)

In some cases apraxia may represent a loss of knowledge about the action necessary to use an object. The FLART was designed to assess this type of conceptual apraxia. It consists of 45 drawings of objects placed in scenes implying an action, such as a slice of toast with a pad of just melting butter on top. Instructions include asking subjects to imagine what tool would be needed to act upon the object and to pantomime the action associated with that tool in relation to the drawing. Patients are instructed to pantomime tool use and told that using a hand to complete the action without the assistance of a tool (hand error) is unacceptable. The total score is the number of items for which the pantomime was interpretable and deemed correct. Interrater reliability was very good (*Kappa* = .97). Patients with mild to moderate Alzheimer's disease scored significantly worse than control subjects. With no time limit, the control group's time to completion was approximately 12 min; for patients with mild to moderate Alzheimer's disease, time to completion ranged from 10 to 43 min. Using 32/45 as a cut-off score, nine of the 12 Alzheimer patients were impaired while none of the 21 control subjects performed below this score. Conceptual apraxia has been found in other studies of Alzheimer patients using different tasks (Dumont et al., 2000).

Test for Apraxia (van Heugten, Dekker, Deelman, et al., 1999)

This test is based on the seminal work by De Renzi in evaluating patients with apraxia. Nine objects are used in testing the ability to pantomime their use on verbal command: first with objects absent and then with objects present, plus demonstration of the actual use of objects. Also included are six items asking for imitation of the examiner's gestures, oral (e.g., blowing out a candle) and hand (e.g., making a fist) gestures as well as closing eyes. A study of 44 stroke patients with apraxia, 35 stroke patients without apraxia, and 50 control subjects demonstrated good construct validity for this test. Its sensitivity and specificity in detecting apraxia were greater than 80%. Assessing object use was more sensitive than imitation of gestures.

Neuropsychological Assessment of Motor Functions

The motor dysfunctions within the purview of neuropsychology are those that can occur despite intact capacity for normal movement. They also have an intentional component that makes them psychological data, unlike reflex jerks, for example, or the random flailing of a delirious patient.

Motor tasks have long been used as indicators of lesion lateralization (G. Goldstein, 1974; Reitan, 1966). On speed or strength tests, it has been assumed that pronounced deviation below a 10% advantage for the dominant hand reflects lateralized brain damage on the side contralateral to the dominant hand, while a much larger dominant hand advantage may implicate a brain lesion contralateral to the nondominant hand (Jarvis and Barth, 1994; Reitan and Wolfson, 1993; see pp. 642, 645). A recommendation that lateralized brain damage is likely when either the nonpreferred hand performance exceeds that of the preferred hand or the preferred hand performance exceeds that of the nonpreferred hand by 20% (e.g., Golden, 1978a) has been seriously questioned. L.L. Thompson, Heaton, Matthews, and Grant (1987) found that this rule would misclassify as having lateralized hemisphere dysfunction up to 18% of left-handed normal subjects on the Finger Tapping Test, as many as 36% of this group on the Grooved Pegboard, and almost 50% of them on the Hand Dynamometer (Grip Strength) test. While misclassifications were greatest for left-handed subjects, around 20% of intact right-handed subjects would be labeled as having "Dominant hemisphere dysfunction" on the basis of the Hand Dynamometer and Grooved Pegboard scores, and 18% would fall into the "Nondominant hemisphere dysfunction" category (see also pp. 305–306). Moreover, mean variations for 26 normal subject groups (*n*s from 10 to 1,128) on the Hand Dynamometer have ranged from dominant > nondominant by 16.5% to nondominant > dominant by 3.3% (for five groups, mean nondominant strength exceeded dominant), leading Bohannon (2003) to conclude that, "Available information may be insufficient to justify using between-side comparisons to make judgments about grip-strength impairment" (p. 728).

Thus findings on speed and strength tests have to be interpreted with caution. Bornstein (1986b,c) found that 25% to 30% of right-handed normal subjects had intermanual discrepancies that exceeded these expectations on at least one speed or strength test; 26% of the normal males and 34% of the females showed no difference or a nondominant hand advantage, again on at least one test; but virtually none of the control subjects had significantly discrepant performances on two or three different motor tests. Right-handed patients with lateralized lesions also displayed considerable variability: those with right brain damage generally conformed to discrepancy expectations (i.e., slowed left hand) more consistently than those with left lateralized lesions, and more than half of the right-damaged patients displayed the intermanual discrepancies expected with lateralized lesions on at least two of the three tests. These findings suggest that more than one motor skill test is required for generating hypotheses about lateralization; and when left hemisphere disease is suspected, the examiner must look to "other nonmotor tasks" (Bornstein, 1986b; see also Spreen and Strauss, 1998).

Further complicating the issue is R.[F.] Lewis and Kupke's (1992) report that patients with nonlateralized lesions tend to perform relatively less well with their nondominant hand because of sluggishness of that hand to adapt to a new task. Moreover, Bornstein (1986c) found sex differences in patterns of performance variability. And on the other hand—literally—Grafman, Smutok, and their colleagues (1985) reported that left-handers who had missile wounds to the brain displayed few residual motor skill deficits long after the injury, a finding that may reflect a less stringent pattern of functional lateralization which allows for greater functional plasticity.

Manual dexterity and strength

Many neuropsychologists include tests of manipulative agility in their examination batteries. These are timed speed tests[1] that either have an apparatus with a counting device or elicit a countable performance. These tests may aid in the detection of a lateralized disability, as may strength testing.

Finger Tapping Test (FTT) (Halstead, 1947; Reitan and Wolfson, 1993; Spreen and Strauss, 1998)

Probably the most widely used test of manual dexterity, this was originally—and by some is still—called the *Finger Oscillation Test.* It is one of the tests Halstead chose for his battery, and its score contributes to the Impairment Index. It consists of a tapping key with a device for recording the number of taps. Each hand makes five 10-sec trials with brief rest periods between trials. The score for each hand is the average for each set of five trials although some examiners give fewer or more trials (Mitrushina, Boone, and D'Elia, 1999; W.G. Snow, 1987b). Reitan and Wolfson (1993) recommend the average of five consecutive trials within a five tap range which may require more than five trials and even "as many as 10 trials in cases of extreme variability" (Jarvis and Barth, 1994). With normal control subjects, Gill and his colleagues (1986) found no fatigue effects on 10-trial administrations but did observe a small but significant increment for men—but not women—retested weekly for ten weeks.

Rosenstein and Van Sickle (1991) called attention to variations in finger tapping instruments which can result in significant performance differences. For example, the manually recording instrument sold with the Halstead-Reitan Battery (HRB) differs from the electronic tapper offered by Western Psychological Services (WPS) in that both the distance the tapper moves and the force required are greater for the former than the latter so that tapping rates run higher for the electronic model (Brandon et al., 1986). Moreover, the lever on the HRB tapper is to the right of the counting box, forcing the left hand into a relatively awkward posture compared with the right hand position. As a result, a right–left hand discrepancy shows up for left-handed persons who do not display the expected left-hand advantage with the HRB instrument (see also L.L. Thompson, Heaton, Matthews, and Grant, 1987), but do show it with the electronic tapper. Like the electronic tapper, a finger tapping program for computers (Loong, 1988) generated somewhat higher tapping scores than the HRB tapper (Whitfield and Newcombe, 1992).

Test characteristics. The 28 subjects who comprised Halstead's control group (see p. 673) averaged 50 taps per 10-second period for their right hand and 45 taps for their left. They provided the cut-off score standard (impaired ranges: ≤ 50 for the dominant hand, ≤ 44 for the nondominant hand) for more than a generation of HRB examinations. Some normative studies vary widely from these scores (see Mitrushina, Boone, and D'Elia, 1999), perhaps in part because different instruments were used, but also because demographic variables influence finger tapping speed significantly. Although faster tapping with the preferred hand is expected, Bornstein (1986c) found that 30% of males and 20% of females from the general population had faster nonpreferred to preferred hand tapping.

Both age and sex exert powerful effects on tapping speed: men consistently tap faster than women (Heaton, Grant, and Matthews, 1991;[2] Heaton, Ryan, et al., 1996; Mitrushina, Boone, and D'Elia, 1999). Slowing with age becomes prominent from about the fifth decade with greatly increasing decrements through subsequent decades. (See Ruff and Parker (1993[2]) for age $\times$ sex norms for four age groups from 16–24 to 55–70.) When applied to normal populations over age 60, the traditional cutting scores correctly identified as normal only 2% to 12% of women and 8% to 10% of men among healthy subjects in the 55 to 70 age range (Bornstein, Paniak, and O'Brien, 1987[2]) and produced similar proportions of false positive classifications—increasing with age and weighing heavily against women—in another large-scale normative study (Trahan, Patterson, et al., 1987[2]). Bornstein and his colleagues (1987) recommended lowering the cut-off scores to ≤ 33 and ≤ 32 for men's dominant and non-

[1] I [mdl] do not give speed-dependent motor tests to motorically slowed patients as I know in advance that they will do poorly and prefer to use our time for more informative testing—and thus also avoid frustrating or embarrassing these patients unnecessarily.

[2] Data reproduced in Mitrushina, Boone, and D'Elia, 1999.

dominant hands, respectively, and to ≤20 and ≤25 for women. These cutting scores would minimize false positive classifications greatly but also increase false negative cases.

Education effects are small, with privileged groups tending to perform a little better on average than is usually reported (Fromm-Auch and Yeudall, 1983[1]), while low levels of schooling are associated with slower tapping performances (Bernard, 1989; Bornstein, 1985[1]; Bornstein and Suga, 1988[1]; Heaton, Grant, and Matthews, 1991[1]; Heaton, Ryan, et al., 1996). Bornstein's (1985) age × education norms are also reported in Spreen and Strauss (1998). Higher education was associated with faster tapping in the oldest but not the youngest group (Bornstein, 1985). Averaged WAIS scores were not related to tapping speed in a large sample of normal subjects ranging in age from 19 to 71 years (Horton, 1999).

Reliability reports vary from study to study. Test–retest study data demonstrate this variability (McCaffery, Duff, and Westervelt, 2000b). The FTT appeared to be highly reliable ($r = .94$ for men, .86 for women) for a small sample of normal subjects retested in ten weekly sessions (Gill et al., 1986). Good but less impressive reliabilities were found for more than 60 healthy adults retested after six months ($r = .71$, .76, for preferred and nonpreferred hands, respectively) (Ruff and Parker, 1993) and for 384 healthy adults retested after two to 12 months ($r = .77$, $-.78$) (Dikmen, Heaton, et al., 1999). Retesting clinical samples (alcohol/trauma, schizophrenia, vascular disorder) at an average of two years between tests (interval range was 4 to 469 weeks) found reliability coefficients in the .64 to .87 range, with the higher coefficients for the nondominant hand (G. Goldstein and Watson, 1989). Four retests of epilepsy patients over six to 12 month intervals found the lowest correlation between the first two tests ($r = .77$, dominant hand), with correlations between retests 2, 3, and 4 all .90 or higher, suggesting a practice effect (Dodrill and Troupin, 1975). Even with Alzheimer patients, a small (ultimately, a 12% increase) but consistently growing practice effect appeared over five assessments at weekly intervals (Teng, Wimer, et al., 1989).

Neuropsychological findings. Brain disorders often, but not necessarily, tend to have a slowing effect on finger tapping rate (Haaland, Cleeland, and Carr, 1977; Reitan and Wolfson, 1996b; Stuss, Ely, et al., 1985). Lateralized lesions usually slow the tapping rate of the contralateral hand (G.G. Brown, Spicer, Robertson, et al., 1989; Haaland and Delaney, 1981; Reitan and Wolfson, 1994). However, these effects do not appear consistently because patients with posterior lesions may not show slowing. Because of inconsistent finger tapping findings, this test cannot be used as a sole indicator of the side of lesion.

Diffuse brain injury impedes rate of tapping, even one year after injury (Haaland, Temkin, Randahl, and Dikmen, 1994). In contrast, grip strength recovered to the normal range in this group of TBI patients with varying injury severity. These findings were interpreted as indicating that slowed processing, frequently reported with TBI, underlies slow finger tapping in these patients. Some TBI patients also have difficulty inhibiting movement of other fingers while tapping, this problem increasing in frequency with severity (Prigatano and Borgaro, 2003). Epilepsy patients generally perform poorly on this test (Dodrill, 1978b), but in evaluating their performances the slowing effects of some anticonvulsive medications must be taken into account. Diseases that involve the spinal cord as well as the brain, such as multiple sclerosis, have a significant slowing effect on FTT scores (Heaton, Nelson, et al., 1985). Some alcoholics may tap more slowly than normal control subjects, but from almost half to 75% of the reported studies showed no group differences between alcoholics and normal subjects (Leckliter and Matarazzo, 1989; Parsons and Farr, 1981). In evaluating this material, one should keep in mind that the studies reviewed in these reports on alcoholics used the original cutting scores which tend to have the high false positive rates discussed above.

Purdue Pegboard Test (Purdue Research Foundation, 1948; Tiffin, 1968)

This neuropsychologically sensitive test was developed to assess manual dexterity for employment selection. It has been applied to questions of lateralization of lesions (L.D. Costa, Vaughan, et al., 1963) and motor dexterity (Diller, Ben-Yishay, Gerstman, et al., 1974) among brain-damaged patients. Following the standard instructions, the patient places the pegs first with the preferred hand, then the other hand, and then both hands simultaneously (see Fig. 16.13). A practice trial for each condition is recommended. Each condition lasts for 30 sec so that the total actual testing time is 90 sec. Although the standard instructions call for only one trial for each condition, when examining patients with known or suspected brain damage, three cycles are recommended. The score is the number of pegs placed correctly.

Average scores of normative groups, consisting of production workers and applicants for production work jobs, ranged from 15 to 19 for the right hand, from 14.5 to 18 for the left hand, from 12 to 15.5 for both hands, and from 43 to 50 for the sum of the first

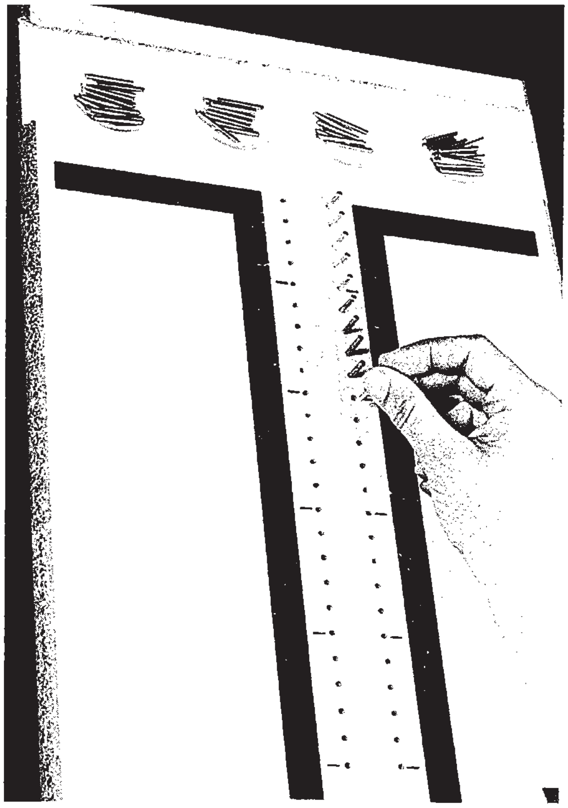

FIGURE 16.13 The Purdue Pegboard Test. (Courtesy of the Lafayette Instrument Co.)

three scores (Tiffin, 1968). As would be expected, handedness significantly affects performance. A study of 30 left-handers and 30 right-handers found that both groups placed approximately four more pegs with the preferred hand than the nonpreferred hand during three trials (Triggs et al., 2000). (For normative data, see Agnew and her coworkers, 1988; Spreen and Strauss, 1998; and Yeudall, Fromm, et al., 1986.)

Test characteristics. Averages for groups of women tend to run one-half to two or more points above the averages for men's groups (Spreen and Strauss, 1998). Scores drop with advancing age, at a slightly greater rate for men than for women (Agnew et al., 1988; Spreen and Strauss, 1998). Agnew and her colleagues also reported that the disparity between hands tends to increase with age as the nondominant hand shows greater slowing over time, a finding that appears to be supported even in the relatively small data sample presented by Spreen and Strauss. Five repeated weekly testings for right hand, left hand, and both hands trials correlated on the average in the .63 to .81 range, but correlations as low as .35 and as high as .93 were recorded (Reddon, Gill, et al., 1988). Practice effects occurred as performances became faster from week to week, with the 12 men in the Reddon study showing a significant increase in speed for trials with each hand (but not both hands); increases in speed shown by the 14 women did not reach significance (see also McCaffery, Duff, and Westervelt, 2000b). Reliability is better for the three-trial than the one-trial administration (Buddenberg and Davis, 2000).

Neuropsychological findings. In a study of the efficiency of the Purdue Pegboard Test in making diagnostic discriminations, the accuracy of cutting scores was 70% in predicting a lateralized lesion in the validation sample, 60% in predicting lateralization in the cross-validation sample, and 89% in predicting brain damage in general for both samples (L.D. Costa, Vaughan, et al., 1963). Since the base rate of brain damaged patients in this population was 73%, the Pegboard accuracy score represented a significant ($p < .05$) prediction gain over the base rate even without taking sex into account. Two separate sets of cutting scores were developed for older and younger age groups. Further, for patients of all ages, a brain lesion is likely to be present whenever the left (or nonpreferred) hand score exceeds that of the right (preferred) hand, or the right (preferred) hand score exceeds that of the left (or nonpreferred hand) by 3 points or more. One-sided slowing suggests a lesion on the contralateral hemisphere; bilateral slowing occurs with diffuse or bilateral brain damage. However, ratio scores comparing the two hands are so unreliable that even large lateralized differences may only have diagnostic value when similar differences show up on other tests (Reddon, Gill, et al., 1988).

Grooved Pegboard[1] (Kløve, 1963)

This test adds a dimension of complex coordination to the pegboard task. It consists of a small board containing a 5×5 set of slotted holes angled in different directions. Each peg has a ridge along one side, requiring it to be rotated into position for correct insertion. It is part of the Repeatable Cognitive-Perceptual-Motor Battery (R. Lewis and Kupke, 1992) and the Wisconsin Neuropsychological Test Battery (Harley, Leuthold, et al., 1980). The score is time to completion. For most clinical purposes both hands should be tested, but one hand may suffice for studying changes in motor speed per se, as can occur with medication (e.g., R.F. Lewis and Rennick, 1979). Mitrushina, Boone, and D'Elia (1999) provide normative data from 16 studies.

[1]Lafayette Instruments offers the lowest price.

Test characteristics. Bornstein (1985[1]); Heaton, Grant, and Matthews (1991[1]); Heaton, Ryan, et al., (1996); and Ruff and Parker (1993[1]) examined demographic variables. Age effects appeared in all of these studies with slowing increasing with advancing age. Bornstein (1985) reported education differences for the dominant hand only, but the Heaton group found no education effects. Bornstein and Suga (1988[1]) attributed the discrepancy in education findings to differences in sample composition as their subjects' education levels were lower than the levels in the Heaton, Grant, and Matthews samples. Bornstein (1985) also reported small but significant sex differences for both hands with considerable overlap between groups. For a large sample ages 16 to 70 (180 of each sex), using the dominant hand, men took on the average 5 sec longer to complete the test than women, with considerable overlap (70.2 ± 13.2 sec, 65.2 ± 12.3 sec, respectively); nondominant hand mean time differences were a bit smaller with greater overlap (76.3 ± 15.3 sec, 72.0 ± 15.1 sec, respectively) (Ruff and Parker, 1993). Similar sex differences were found in a study of 102 young adults: women outperformed men in the dominant hand condition but differences between groups were not significant with the nondominant hand (S.L Schmidt et al., 2000). Substantial test–retest reliabilities have been found ($r \geq .82$) (Dikmen, Heaton, et al., 1999), although practice effects do not appear consistently when control subjects are retested (McCaffery, Duff, and Westervelt, 2000b). When each hand has three trials, performance improves significantly over trials (S.L. Schmidt et al., 2000).

Bornstein, Paniak, and O'Brien (1987) showed that previously established cutting scores misclassified 66% of dominant hand performances and 72% of nondominant ones by intact subjects, although virtually no brain damaged patients were misclassified. They recommended a new set of cutting scores (≥92 dominant, ≥99 nondominant) which misclassified only 11% and 9%, respectively, of normal subjects but more patients (27% and 40%, respectively).

Neuropsychological findings. This test can aid in identifying lateralized impairment (Haaland, Cleeland, and Carr, 1977). Bornstein (1986b) suggested that a right/left hand score ratio greater than 1.0 suggests right hemisphere disease, and a ratio less than 1.0 may be indicative of damage involving the left hemisphere; but he cautioned that these ratios are too variable to rely on without supporting data from other tests. However, deficits on the ipsilateral hand trial after stroke or tumor to either hemisphere suggest that both hemispheres are equally important for performance on this task (Haaland and Delaney, 1981). Its complexity makes this a sensitive instrument for measuring general slowing whether due to medication (R.F. Lewis and Rennick, 1979; C.G. Matthews and Harley, 1975), diffuse brain dysfunction (Nathan et al., 2001), or progression of disease processes such as parkinsonism (Matthews and Haaland, 1979) or HIV infection (E.N. Miller, Selnes, et al., 1990; Y. Stern, McDermott, Albert, et al., 2001). Slowing on this test may also appear with toxic effects of microorganism excretions (Grattan, Oldach, et al., 1998) and environmental lead (Bleecker, Lindgren, and Ford, 1997).

[1]Data reproduced in Mitrushina, Boone, and D'Elia, 1999.

Hand Dynamometer or Grip Strength Test (Reitan and Wolfson, 1993; Spreen and Strauss, 1998)

This technique detects differences in hand strength under the assumption that lateralized brain damage may affect strength of the contralateral hand. The standard neuropsychological administration calls for two trials for each hand alternating between hands. The score is the force exerted in kilograms for each hand averaged for the two trials. A testing protocol for occupational therapy evaluations recommended three trials for each hand and found the average score to be more reliable than the best score (Mathiowetz et al., 1984). James L. Mack pointed out that this test requires effort and that the degree of voluntary effort a subject puts forth may vary for any number of reasons (personal communication [mdl], September 1991). He therefore recommended that the standard administration be compared with a second one in which attention is diverted from the task by performing a little sensory test, such as two-point discrimination, on the other arm. A number of workers have developed norms for this task (Mitrushina, Boone, and D'Elia, 1999). Differences between the reported mean scores are all within a one to two kilogram range.

Test characteristics. Sex differences are unequivocal (Dodrill, 1979); the sexes differ further in that men show a greater intermanual discrepancy than do women (Bornstein, 1986c[1]). Significant age effects appear (Bornstein, 1986c; Ernst et al., 1987), but men and women do not show them in the same way or in all studies. In one, men's scores held up until age 40 and then decreased (Fromm-Auch and Yeudall, 1983); however, they did not begin dropping until after age 60 in the 1986 Heaton, Grant, and Matthews[1] study. Fromm-Auch and Yeudall's data do not show a corresponding pattern of weakening with age for women, although Koffler and Zehler (1985) documented lower scores from age 40. Data on education effects are equivocal: Bornstein (1985[1]) found that education contributed significantly to grip strength scores, but Ernst (1988) did not for an elderly sample, nor did Heaton and his colleagues. What education effects have been reported for grip strength tend to be relatively slight (Leckliter and Matarazzo, 1989) and may be more related to such other variables as healthful nutrition and/or good working conditions than to each other.

This is a highly reliable technique. With ten trials, some fatigue effects occur, but not on the first two trials (Reddon, Stefanyk, et al., 1985). Using a two-minute rest between trials, Dunwoody and coworkers (1996) found that performance actually improved over the first three trials, presumably as the subjects became more familiar with the task and, perhaps, as their muscles warmed up. Over ten weeks of weekly retesting, some increase in strength appeared, but not within the first three weeks. For both hands, the Reddon group found good average test–retest reliability for men (r = .91) and women (r = .94). R. Lewis and Kupke (1992) reported almost perfect test–retest reliability (r = .98). A comparison of women's test–retest scores on the two-trial condition showed that reliability correlations for right hand performances were somewhat lower than for the left (r = .79, r = .86, respectively) (Mathiowetz et al., 1984). (For other test–retest data, see McCaffery, Duff, and Westerveld, 2000b.)

Neuropsychological findings. As in other tests of manual abilities, strength between hands varies widely in patients with lateralized brain disorders as well as in normal control subjects (Bornstein, 1986b; Dodrill, 1978a). Using a classification criterion of −2 SD, Koffler and Zehler (1985) found 27% of normal subjects misclassified as brain damaged when dominant hand strength exceeded that of the nondominant hand by 5 kg: 21% were called "brain damaged" because the strength of the dominant hand was not greater than that of the nondominant hand. Like finger tapping, similar contralateral deficits were seen after damage in a variety of locations to the left or right hemisphere (Haaland and Delaney, 1981).

[1]Data reproduced in Mitrushina, Boone, and D'Elia, 1999.