THE ASSESSMENT OF POSTOPERATIVE MENTAL FUNCTION

G. B. DRUMMOND

The purpose of postoperative assessment of mental function is either to determine the degree of recovery from anaesthesia, by estimation of residual effects, or to detect mental changes caused by the process of anaesthesia and surgery. These changes may be caused by the anaesthetic agents themselves, by physiological changes resulting from anaesthesia, or by the effects of other factors that are present in this period such as mental stress and physical disease.

The aims of assessment may be clinical, as in the assessment of "street fitness" after outpatient general anaesthesia; or assessment may be undertaken experimentally, in the evaluation of drugs and techniques.

In the measurement of recovery, the duration of action of a specific agent, the persistence of changes in mental function caused by anaesthesia using this agent, and the possibility of interaction with other agents are assessed. Mental changes in the postoperative period may be reflected in alterations of psychiatric state, of social integration, or of cognitive powers and skills that are evaluated by means of psychological tests, or by means of physiological tests associated with these changes.

In order to consider postoperative brain function it is useful to examine first some of the central changes that are known to occur during anaesthesia. Following work on isolated brain preparations, using the electroencephalograph to detect states of "wakefulness" and "sleep", a neurophysiological concept of anaesthesia was proposed, based on the hypothesis of a selective susceptibility of the multisynaptic pathways of the reticular system to anaesthetic agents (French, Verzeano and Magoun, 1953). Later work has shown that inhalation anaesthetic agents may be classified by their differing effects on cortical and subcortical electrical activity, and that some of the changes produced are of increased activity (Clark and Rosner, 1973; Rosner and Clark, 1973; Kavan, Julien and Lucero, 1972).

However, anaesthesia is not exclusively related to cortical and subcortical changes. Anaesthesia with agents such as droperidol and fentanyl is associated with minimal electrical changes in cortical function (Sovijärvi and Sainio, 1972), and anaesthetic agents in clinical concentrations have been shown to have effects within the spinal cord (Freund, Martin and Hornbein, 1969) and on the less complex cell systems of the dorsal column (de Jong, Robles and Morikawa, 1969; Taub, Hoffert and Kitahata, 1974).

A great deal of investigation of the changes of brain function in anaesthesia, as for the states of sleep, has been by the study of electrical activity, using the e.e.g., evoked potentials, and intracellular recording. It is difficult to explain, in terms of electrical activity alone, alterations in the proportions of wakefulness, drowsiness, deep sleep, and paradoxical sleep which can persist for several days. Such effects are caused by influences such as sleep deprivation, or by anaesthesia (Yanagida and Yamamura, 1972), and are probably caused by changes in the level of neurohumoral and transmitter substances in the brain (Jouvet, 1969). The effect of drug-induced changes in central catecholamine concentrations on the potency of inhalation anaesthetic agents is discussed by Eger (1974).

Human sleep patterns can be altered by agents that inhibit central catecholamine synthesis (Wyatt et al., 1971), and in animals i.v. anaesthetic agents such as thiopentone cause altered rates of catecholamine synthesis (Sung, Frederickson and Holtzman, 1973). It is possible that reductions in mental function after operation may result from such effects rather than from the direct effects of residual anaesthetic agent, and this may explain the persistence of some of the effects.

For example, James (1969) found that his subjects suffered from tiredness and lack of concentration for up to 6 days after prolonged cyclopropane anaesthesia. Drowsiness is the commonest complaint after outpatient anaesthesia, and may persist for several hours (Fahy and Marshall, 1969). However, even inhaled concentrations of halothane and nitrous oxide in trace amounts (15 and 500 p.p.m.) can increase drowsiness (Bruce, Bach and Arbit, 1974), and corresponding trace blood levels could persist for some time after prolonged anaesthesia.

Postoperative changes in mental function may...
therefore be caused either by the action of residual anaesthetic agent, or by the changes in central activity caused by anaesthesia. A third cause of postoperative changes in mental function is the result of factors present during anaesthesia or after surgery which damage the brain or impair its function.

A reduction in cerebral blood flow caused by hypotension or hypocarbia is such a factor that has been extensively investigated. A recent review (Harp and Wollman, 1973) indicated that although considerable decreases in arterial pressure or arterial carbon dioxide tension may be tolerated in normal subjects without biochemical or other evidence of cerebral hypoxia, the safety margins may be much less in subjects with cerebrovascular disease. In investigations of the cerebral effects of hyperventilation during anaesthesia, Wollman and Orkin (1968) detected a significant increase in reaction time in those patients who had had a PaO2 below 24 mm Hg during anaesthesia, but Blenkarn and co-workers (1972) found no difference between normocarbic and hypocarbic anaesthetic techniques. These workers, using a battery of psychometric tests, measured reaction times, an index of visual perception, a test of short-term memory, and a test of selective auditory attention, in a group of patients who were younger than those studied by Wollman and Orkin. A preliminary study of the effects of hyperventilation on memory and learning ability suggested that hyperventilation in older patients might impair short-term memory (Murrin and Nagarajan, 1974).

The possibility of inadequate cerebral perfusion during induced hypotensive anaesthesia has been similarly investigated. In a well-controlled trial, Gruvstad, Kebbon and Lof (1962) found a slightly greater impairment in psychiatric and psychometric tests after hypotensive anaesthesia, but considered these changes of little importance in the patients' everyday life. In a group of younger patients, Eckenhoff et al. (1964) used a battery of tests which they considered relevant to the everyday mental activities of the subjects, and was likely to detect organic brain damage. These were predominantly tests of perception and short-term memory. No differences were found between patients anaesthetized with and without induced hypotension. In elderly patients who had undergone prostatectomy (Rollason et al., 1971), a battery of psychomotor tests of intelligence, perception, memory, and vigilance revealed no difference between patients given a spinal anaesthetic with vasopressor and those given the spinal anaesthetic alone. However, one patient from each group showed deterioration on testing after operation, and two patients with pre-operative signs of organic brain disease demonstrated markedly impaired postoperative performance.

**Psychiatric Investigations**

Patients commonly approach surgical operations with fear of the procedure and possibly of the outcome (Kornfeld, 1969a; Winkelstein, Blacher and Meyer, 1965). Such attitudes may be altered by their previous experience of anaesthesia and surgery. For these reasons, psychic changes are to be expected in the period following surgery, in addition to the effects of anaesthesia, postoperative medication, pain and other discomforts, restricted movement, lack of sleep, and the possibility of organic brain damage.

The main method of investigation of personality and psychic change in this period has been by the use of interviews and personality tests. However, Eckenhoff, Kneale and Dripps (1961), in a prospective study using clinical recovery room observations alone, were able to describe the factors influencing postoperative delirium, in particular the use of narcotic premedicant and the influences of age and pain.

Interviews may be of a free form, by a psychiatrist, intended to discover overt and repressed anxieties, and changes in mood and attitude. Before operation, anxiety caused by real threats such as enforced unconsciousness, the entry of the surgeon into the body, and the possibility of the discovery of unwelcome features of disease, or the possibility of death itself, and by symbolic threats, have been described by this means (Winkelstein, Blacher and Meyer, 1965). Other methods of investigation are by structured interviews to discover specific aspects of the postoperative situation, such as minor discomforts, drowsiness, memory for the operative period, and alterations in personal efficiency and relationships (Fahy and Marshall, 1969; Thomas, 1963; James, 1969; Brice, Hetherington and Utting, 1970), or by questionnaire (Sheffer and Griefenstein, 1960; Fahy, Watson and Marshall, 1969; Cronin, Redfern and Utting, 1973).

A personality test such as the Eysenck inventory (Eysenck and Eysenck, 1964) allows scoring of personality traits on scales of extraversion and neuroticism. It is available as two alternatives which allow pre- and postoperative testing. Such a score
has been shown to be unaffected by operation, but to be related to the incidence and severity of complaints after surgery (Cronin, Redfern and Utting, 1973). The Minnesota Multiphasic Personality Inventory, which allows the scoring of psychiatric symptoms, rather than the description of basic personality traits, has been used to demonstrate reduction in anxiety after open heart surgery and a return of personality profiles towards normal (Gilberstadt and Sako, 1967). Blundell (1967) used a daily rating of personality traits in elderly patients to show a significant reduction in psychological well-being after surgery.

Two studies of the influence of anaesthesia and surgery on the elderly have used standardized questionnaires or interviews recording social integration and competence, in addition to psychological tests, in order to assess overall changes in psychosocial status (Simpson et al., 1961; Blundell, 1967). These workers found that the effects of surgery could improve the psycho-social status of patients, and that a reduction in social integration could be detected without deterioration being detectable by formal psychometric tests.

General anaesthesia appears to reduce the manifestations of anxiety in the period immediately after operation, and these become evident at interview a day later (Winkelstein, Blacher and Meyer, 1965). Repression of emotional responses is more marked in those patients who have not been interviewed before surgery (Sheffer and Greifenstein, 1960). Even though the method of evaluation is less exact than formal psychometric tests, it is evident that the conditions of such investigations must be carefully controlled to allow worthwhile conclusions to be made.

Delirium in the postoperative period may occur during recovery, when pain, age, and the preoperative mental state of the patient have been shown to affect its incidence (Eckenhoff, Kneale and Dripps, 1961), or it may persist for several days. It is likely that preoperative fear, and the postoperative environment, are important factors (Kornfeld, 1969b). Other factors that are known to result from the action of drugs, such as impairment of memory, altered time sense, and the déjà vu phenomenon (Robson, Burns and Welt, 1960; Henrie, Parkhouse and Bickford, 1961; Lambrecht and Parkhouse, 1961) may contribute to disorientation after operation. Clinical observation of toxic psychosis after surgery in the elderly suggests that the most important factor is a failing mental status before the operation, and that immobilization and anaemia may be further contributory factors (Scott, 1960).

Delirium after cardiac surgery has been closely studied, although the relative importance of organic and psychological factors in the causation is still not clear (Baxter, 1974).

Gilman (1965) studied the clinical neurological features of patients after open heart surgery, and found that patients with signs of cerebral disturbance of agnostic type also had signs of neurological deficit such as an extensor plantar response. He considered that such patients had suffered specific cerebral injury, and associated this with operative hypotension, prolonged cardiopulmonary bypass, and a more elderly state.

**Psychological Tests**

Psychological tests are used to measure mental function in a controlled and quantitative fashion, and to allow comparison of mental function in different subjects or differing situations. Although many tests have been devised to measure a limited component of mental function, almost all are dependent upon certain basic mental processes, and the overlap of the aspects measured by each test is considerable.

For example, although a test such as the visual reaction time appears to be a straightforward measurement of the "speed" of a simple mental and motor process, from stimulus to response, "stimulus" and "response" are only convenient labels for complex processes (Fitts and Posner, 1967). Despite the fact that the perception of the sudden illumination of a light seems to be almost immediate, incoming data from the eyes have to be organized to allow the subject to see a coherent object with features such as outline, dimension, brightness, position, and so on (Hubel and Wiesel, 1962). This process may be enhanced by other data provided by other senses, and related to what is seen. Further organization of this visual information allows the abstraction of particular qualities from these outlines, with the help of a fund of past experience, and leads to the identification of the object as a light bulb, to its placing in its correct context, for example as the stimulus in a reaction time test, and to its placing in time, in terms of its previous state. In this example the process may be simplified by the past experience of observing the light bulb in this position, and by the expectation of a change in intensity (Welford, 1958).
The final step in perception in this example lies in the question: "Has the light just undergone an increase in intensity?" This depends on the criteria that the subject applies. In this instance, the decision is clear-cut, but in many other cases the decision is determined not only by the stimulus itself but also by the criteria used. For these latter cases, the changes may be infrequent, less obvious, and the attention of the subject may wander. Changes may be missed, or false changes detected when in fact no real change has occurred. Criteria for signal detection depend upon the implication (or "cost to the subject") of on the one hand, the missed signals, and on the other the false alarms. The criteria may be altered by habituation to signal events, and the inhibition of unimportant stimuli (Mackworth, 1970).

Hence, although certain tests may be thought of as testing either "learning" or "motor skill" or "apprehension", processes and factors such as vigilance, perception and the abstraction and integration of information, are intrinsic to them all. Another important factor is short-term memory, particularly in the more complex tests of skill that involve the "translation" of information from one mode to another (Welford, 1962); and also in tests of perception such as dichotic listening when two channels of auditory information are presented simultaneously. In a similar way Fleishman (1958) showed that several tests of motor skill involve common components such as fine control sensitivity, co-ordination of several limbs in complex movements, and the ability to respond correctly to a given stimulus.

Psychometric tests were developed to allow the evaluation and scoring of intelligence, either in terms of developmental age, or on a point scale. They contain a number of subsets consisting of intellectual tasks, in which the subject may be asked to define words, remember and reproduce facts, solve arithmetic problems, and recognize likenesses and differences. He may also be asked to perform tests involving coordination and judgement, such as dotting the centre of a series of circles; or tests involving the memorizing and retention of words, pictures, and numbers for a period of time. These tasks are scored for correct responses, time taken for completion, and errors, according to the format. An implicit assumption in such "batteries" of tasks whose scores are considered as a total, is that the tests are of functional equivalence; and that the use of a number of tests removes the dependence upon specific knowledge, and allows general reasoning ability to be estimated.

A test of this type is the Wechsler Adult Intelligence Scale. It has been realized for some time that this test contains some elements whose results correlate poorly with the overall score for the test, for example the memory span for digits. As mental ability declines with age, this test, and other elements that may involve a greater degree of memory, such as the digit-symbol test, show a more marked reduction than tests such as information and vocabulary. Organic disease likewise is associated with a selective deterioration in these same tests. Steinberg (1954) assembled a series of tests which he considered to be of increasing complexity, from arithmetic and fluency to tests of relational ability using verbal and visual analogues. Testing during the administration of 30% nitrous oxide to student volunteers showed that the more complex tasks were more affected.

Intelligence tests can be biased either toward the concept of intelligence as the ability to perceive logical relationships and to manipulate symbols, when verbal and arithmetic tests, and tests of abstract reasoning are used; or they may be biased towards the ability to handle practical situations, when the tests will measure skills and manipulative abilities. These tests are used not only for the measurement of mental development and ability, but also for the detection of mental disturbance. For example, elements of the Wechsler Adult Intelligence Scale test the ability to perceive logical relationships and to use symbols. These elements of the test may be sensitive to changes in ideation caused by mental disturbance, particularly if the test is administered by a trained psychologist who can interpret the responses.

The choice of psychometric tests for pre- and postoperative use is dictated by practical factors such as the advantage of duplicate, standardized tests, the time necessary for the test administration, objectivity, and ease of scoring. Workers have differed on theoretical grounds in their choice of tests. Wollman and Orkin (1968) believed that although some tests may be representative of the complex functions of the brain, they may not be applicable in the detection or quantitative evaluation of subtle changes in brain function. Simpson et al. (1961) who tested 250 elderly patients before and after routine surgery, did not use either the Wechsler or Raven Matrices tests, partly because they did not include tests of long-term memory and manual
dexterity. Other workers used memory tests such as paired associate word learning and the Benton test of visual retention because they considered that such tests were relevant to everyday mental activities, and were relatively sensitive to brain damage (Eckenhoff et al., 1964; Rollason et al., 1971).

Such tests have been used in the investigation of the effects of experimental anaesthesia, anaesthesia in the elderly, hypotensive and hypocapnic anaesthesia, and of cardiopulmonary bypass (James, 1969; Simpson et al., 1961; Blundell, 1967; Eckenhoff et al., 1964; Rollason et al., 1971; Blenkarn et al., 1972; Gilberstadt and Sako, 1967). Although in young experimental subjects, memory dependent tests seemed to be more affected (James, 1969), or could be more sensitive (Blenkarn et al., 1972), in elderly patients Blundell (1967) found that the major effects were seen in mental organization rather than in memory.

In a study of the effects of trace concentrations of halothane and nitrous oxide, with a battery of psychometric tests, Bruce, Bach and Arbit (1974) found that memory and digit span showed significant deterioration, whereas other elements of the Wechsler intelligence and memory scales did not.

Vigilance.

Vigilance tasks are concerned with the reduction in attention which a subject shows when presented with a monotonous and uninteresting task. This reduction in attention is estimated by the decrement in the detection of signal events, by the change in the number of false alarms, and the change in the speed of response. These changes may be caused either by habituation of physiological responses, or by changes in the criteria used by the subject in detecting a signal, or by both these mechanisms (Mackworth, 1970).

A test of this type that has been used with patients after surgery is the letter scoring test, in which the subject has to delete a particular letter out of a large block of lines of random letters, as rapidly as possible. Scoring is by counting the deletions made correctly, the errors made, and the number of lines completed (Rollason et al., 1971; Dixon and Thornton, 1973).

An interesting test of perception and attention, used in the estimation of recovery, is the Bender face-hand test (Jaffe and Bender, 1951). This is based upon the tendency of the incompletely recovered subject to neglect or incorrectly localize one of two simultaneously administered cutaneous stimuli. A touch on the face is more readily perceived and reported than a simultaneous touch, of equal duration and intensity, applied to the hand. The time that elapses before both stimuli are felt and accurately located, in a recovering subject with eyes closed, is surprisingly long. After anaesthesia with nitrous oxide and pethidine supplements, the mean time to recovery using this test was 43 min, whereas after thiopentone supplements the time was 107 min (Widdowson, Teresita and Virtue, 1955). Gale (1957) found this test to be more sensitive than formal clinical assessment of recovery.

Reaction time.

Tests of reaction time can be simple, when a single event is used to signal a single response; or disjunctive, when a choice of response has to be made according to a variable signal. Reaction time increases in a non-linear way as the choice increases. Analysis of the relationship suggests that for the simple reaction time, a choice still exists—the decision as to whether the signal has appeared or not (Welford, 1960).

In postoperative use, Blenkarn et al. (1972) considered that this test could be affected by residual muscle paralysis, as they had found that both simple and disjunctive reaction times increased by similar amounts. The possibility that the motor element of this test may be considered separately is supported by the analysis of a large series of motor skills carried out by Fleishman (1958). Statistical analysis of the tests suggested that the features of the reaction time were "reaction time" and "speed of arm movement", and these varied according to the type of test considered.

An increase in reaction time after premedication was attributed by Hendry, Norris and Nisbet (1963) to a loss of coordination, and a reduced level of arousal. It is probable that in these circumstances, the test may also be one of vigilance. Bruce, Bach and Arbit (1974) used a complex test of reaction time, designed to simulate the "vigilance task" of watching an e.c.g. monitor and listening to an audible signal of heart rate. A choice of response was given, according to the combination of changes in audible and visual signal. They found that performance of this task was significantly impaired after exposure of the subject to trace levels of halothane and nitrous oxide.
Motor skills.

Tests of motor skill may be of a simple form, such as the transferring of pegs from one row of holes to another in a pegboard (Vickers, 1965; Blundell, 1967) or of a complex type such as simulated driving (Egbert, Oech and Eckenhoff, 1959; Elliott et al., 1962; Green et al., 1963). They are assessed by rate of completion, by the number of errors made, and a measure of the accuracy. The inclusion of all three of these features is necessary for evaluation as they are to some extent interdependent; errors will increase and accuracy suffer if the speed of performance of the task is increased (Fitts and Posner, 1967).

However, as emphasized above, even the simplest of such tasks may involve a whole series of steps of perception, decision, and execution. Execution itself involves the coordination of visual and proprioceptive sensation with effector activity, by a variety of “feedback” loops (Paillard, 1960). Steinberg (1954) showed that 30% nitrous oxide inhalation caused more deterioration of performance in a task of manual dexterity, than in what were considered complex tasks involving abstract cerebral function.

Clinical criteria of recovery such as ataxia, steadiness of gait, and evenness of speech, are estimates of motor skill, as is the measurement of extraocular muscle control by means of the Maddox Wing (Hannington-Kiff, 1970).

Memory.

Memory and learning are thought to consist of short-term and long-term processes. Short-term memory involves the transient retention of information. This information is probably encoded by the temporal and spatial patterns of neuronal activity within the cerebral cortex ( Eccles, 1973). This activity can be interfered with by drug effects and by hypoxia, either by failure of acquisition (Robson, Burns and Welt, 1960; Summerfield, 1964; Crow and Kelman, 1971), or an increase in the rate of decay (Grove-White and Kelman, 1971b; Brown, 1958).

After about 1 min, consolidation of the memory trace seems to occur, and thenceforth recall is possible even after temporary interruption of mental activity ( Eccles, 1973; Cherkin and Harroun, 1971). Both long- and short-term memory seem to be sensitive to adverse effects, but it is possible that the process of long-term retention is more susceptible to influences such as hypoxia and anaesthesia ( Crow and Kelman, 1971; James, 1969).

Short-term memory may be assessed by recall of six random single digits or letters, after intervals up to 30 sec or so after presentation; the interval is usually filled with a second task to prevent rehearsal. The omissions and incorrect recollections are recorded as errors, and the error rate is used for analysis. Short-term memory has been investigated in the study of effects of hypocapnia in anaesthesia by Blenkarn et al. (1972), and Murrin and Nagarajan (1974). No significant change was detected, using the digit or forward letter span tests. Bruce, Bach and Arbit (1974) found that a 4 hr exposure to trace concentrations of halothane and nitrous oxide, or nitrous oxide alone, reduced the digit span.

Long-term memory is commonly assessed by the time taken to learn groups or pairs of items, using one member of the pair as a “prompt” in testing. An example is the paired associate word learning test, used by Eckenhoff et al. (1964) and Rollason et al. (1971) in patients after hypotensive anaesthesia. Other tests such as the Hunt Minnesota test, where six nonsense syllables have to be learnt, or such as the memorizing of picture/word associations, have been used in studies of drug effects (Henrie, Parkhouse and Bickford, 1961; Osborn et al., 1967). Blundell (1967) included a test of memory for personal details and current events in a wide range of psychosocial tests for postoperative assessment.

Amnesia for postoperative events can be more common in the elderly (Lambrechts and Parkhouse, 1961), and after hyoscine premedication. Gruber and Reed (1968) found that general anaesthesia reduced recall, whereas spinal anaesthesia did not, but were unable to demonstrate a greater reduction in recall in older patients. However, 18 of 20 patients interviewed within 1 hr of anaesthesia were able to recall the interview the next day (Winkelstein, Blacher and Meyer, 1965). Recall of preoperative events is greater if the patients are informed that a study of recall is being conducted (Authier et al., 1974).

Physiological Tests

The measurement of the frequency at which a flickering test disc appears to become a steady light (the critical fusion frequency) allows assessment of cortical function. The methods of this test are reviewed by Simonson and Brozek (1952). The results of the test show little variation with practice, or from day to day. Sleep deprivation has little effect. The fusion frequency falls with age,
hypoxia, and persistent exposure to the test disc. Drugs such as alcohol and barbiturates cause a marked reduction in the fusion frequency.

This test has been used to investigate the effects of hypotension and hyperventilation (Berg, Nilsson and Vinnars, 1957; Allen and Morris, 1962; Whitwam et al., 1966), and the timecourse of the action of sedative agents (Grove-White and Kelman, 1971a).

The electroencephalograph measures the sum of the electrical activity of the cerebral cortex, and therefore provides a continuous indication of its action. However, in situations where no focal lesion is present, a marked alteration in cerebral state has to occur before recognizable changes in the pattern can be detected (Bromage, 1953; Cohen et al., 1967). Subjective changes and altered memory caused by nitrous oxide could not be correlated with changes in e.e.g. frequency by Henrie, Parkhouse and Bickford (1961).

Doenicke and his co-workers (1966) used the electroencephalograph to assess periods of drowsiness after use of i.v. induction agents. They were able to detect marked differences between methohexitone and propanidid. The value of the electroencephalograph would appear to lie in its ability to estimate the proportions of time spent in the drowsy or sleeping state. An increased tendency to sleep will not be detected by psychometric tests, unless the longer tests of vigilance are used, and such a tendency is clearly of clinical importance.

FACTORS INFLUENCING ASSESSMENT

Many factors may be present in the postoperative situation which are likely to affect the results of psychometric tests. Probably the most important of these is a change in the level of arousal. The optimal level of arousal for the performance of a particular task becomes lower for a more complex task (Colquhoun and Corcoran, 1964).

The level of arousal is influenced by personality, sleep loss, and depressant drugs (Mackworth, 1970). However, Blundell (1967) did not find that intelligence or the presurgical psychiatric state influenced the results of a series of psychometric tests. Motivation of the subject is important in determining the level of arousal. It is known, for example, that subjects able to describe the purpose of the investigation in which they are taking part, are more likely to show the expected result; and that volunteers are highly motivated, and will perform boring tasks for long periods without deterioration in performance (Orne, 1962). Knowledge of test results causes an increase in performance, even if the details of performance which are given to the subject are false (Mackworth, 1970). Environmental features such as noise and temperature alter arousal, and thus the results of tests of vigilance (Broadbent, 1963).

Performance in certain tests may be influenced by social class (Williams, 1960). Organic disturbance such as pyrexia (Blundell, 1967) and physical discomfort (Rollason et al., 1971) can reduce test scores or make the test appear more of an effort. It is clear that control groups have to be carefully examined to make sure that groups are comparable in respect to these factors.

Several investigations have demonstrated differential impairment of mental function. This implies that the result of a particular test will depend upon the extent to which it reflects this specific mental function, and the impairment of the specific function which is present at the time of testing. If analgesia or sedation reduces arousal (Hendry, Norris and Nisbet, 1963), alters memory processes (Summerfield, 1964; Osborn et al., 1967), impairs time estimation (Robson, Burns and Welt, 1960), or interferes with proprioception (Legge, 1965), then test results will be modified.

The effects of different sedative agents may be distinguished by the use of tests which have to be performed within a time limit (paced tasks), and tests in which the subject performs at his own speed (unpaced tasks) (Orzack and Kornetsky, 1963). Thus, the element of pacing must also be controlled.

Analysis and interpretation of these tests may be difficult. It may not be valid to assume that separate elements of a battery of tests are equivalent, and therefore the comparison of score totals can be misleading. The interpretation of speed and performance and error rate may be difficult, as these factors have an inverse relationship. Finally, the use of tests of significance that are distribution sensitive, such as the t test, may be unjustified (Eckenhoff et al., 1964). If possible, control groups, controlled for all the factors mentioned, should be studied in addition to testing subjects before and after operation, to allow estimation of effects such as rehearsal (Blundell, 1967).

INITIAL RECOVERY FROM ANAESTHESIA

Changes in cerebral state during initial recovery from anaesthesia are not well described. Stoelting,
Longnecker and Eger (1970) have measured the alveolar concentration of anaesthetic which lies between the concentration that just allows the patient to open his eyes on request, and the concentration that just prevents this response, and described this as the "MAC awake". They found that this was relatively constant at about 0.6 of the actual MAC for that agent.

Scott (1972), in a review of reports of consciousness during anaesthesia, has described initial return of hearing, before discomfort or vision. Evidence of selective attention or awareness of "threatening" auditory input, as reported by Levinson (1965) on the basis of e.g. evidence, and recall using hypnosis, has not been supported by later investigations of possible selective auditory sensibility (Terrell et al., 1969; Lewis, Jenkinson and Wilson, 1973).

A period of increased muscle tone during recovery from halothane anaesthesia has been described, and the use of a central stimulant, methylphenidate, was associated with its abolition (Brichard and Johnstone, 1970). Later work showed that very frequently a period of spasticity with increased muscle tone, ankle clonus, and extensor plantar responses becomes evident when a recovering patient is responsive to pain, and ceases when the patient responds to verbal stimuli. This was seen during recovery from halothane, methoxyflurane, and cyclopropane, and less commonly after nitrous oxide/pethidine anaesthesia (Soliman and Gillies, 1972).

The period of recovery from anaesthesia may be defined as the time from the end of the operation to the end of the effects of the anaesthetic agents. However, as all of the effects of the anaesthetic agents may be impossible to define or measure, in practice recovery is held to have taken place when a specific criterion of recovery has been fulfilled, using either a single test or a battery of tests. Strictly speaking, it is only true to say that patients who do not fulfil such criteria are not recovered, rather than to assume that the patients who do satisfy the criteria are in fact recovered.

Choice of the test is influenced by the balance between discrimination and sensitivity. The discrimination of the test is the ability to distinguish recovered patients from those who have not recovered, with as little misclassification as possible. This ability will be greatest if a less subtle indication of recovery is used, so that a more distinct difference between recovered and non-recovered patients is present. For example, Dixon and Thorn-
ted with an increase in symptoms, but anaesthetic agents, personality, and age were not. Data from a postal questionnaire indicated that symptoms were even more common (Fahy, Watson and Marshall, 1969).

These workers concluded that postanaesthetic questionnaires were a sensitive method of estimation of postanaesthetic morbidity.

Bruce, Bach and Arbit (1974) studied the effects of exposure to traces of halothane vapour (15 p.p.m.) and nitrous oxide (500 p.p.m.) for a 4-hr period. They found significant reductions in vigilance, reaction time, and short-term memory. During the period of exposure, 6 of the 20 subjects fell asleep for some time, compared with 1 out of 40 control subjects.

After 5 hr of cyclopropane anaesthesia, James (1969) found that volunteer subjects did not feel fully recovered for an average of 3 days. Support for such subjective evidence is given by the electroencephalographic studies of Doenicke and colleagues, who found drowsiness after use of barbiturate induction agents (1967) and potentiation of the effects of alcohol (1966) for up to 12 hr.

Other subjective methods used are the patient’s opinion of his state of recovery (Green et al., 1963; Goldman and Kennedy, 1964) or the recording of symptoms such as euphoria, sleepiness, or drunkenness (O’Mullane, 1957).

Relatively insensitive but widely used and accepted clinical criteria of recovery have been used, either singly, together, or along with more complex tests. These include rousability, return of the lash reflex, opening the eyes either spontaneously or to command, answering questions, the ability to talk rationally, or to sit up when asked (Simmons and Blanshard, 1957; Burheim, Terry and Herron, 1966; Foley et al., 1972; Jolly, 1960; Gale, 1957; Wyant, Dobkin and Aasheim, 1957; Hutchinson, 1963). Smith, Allen and Perrin (1967) showed that the lash reflex and the response to auditory stimuli returned at the same time, which was significantly earlier than the flicker fusion frequency returned to normal.

Proprioception and muscle coordination may be assessed subjectively, clinically, or by physiological and psychological tests. In the testing of posture control, learning effects will be minimal, but in psychomotor tests such as track tracing or the steadiness-precision test (Fleishman, 1958) a learning effect may be present. Wyant, Dobkin and Aasheim (1957) used the symptom of diplopia in the clinical estimation of recovery. Hannington-Kiff measured and compared recovery from anaesthesia, using the Maddox wing to estimate the degree of extraocular muscle imbalance (1970, 1972). This test seems to be sensitive, and highly reliable, but calls for a degree of patient cooperation.

Ataxia and swaying have been used as clinical signs (Wyant, Dobkin and Aasheim, 1957; Green et al., 1963; Goldman and Kennedy, 1964; Baird and Flowerdew, 1970) and swaying has been measured more formally by Vickers (1965) in a study of tests of recovery, when it was found to be of only poor predictive value.

Psychological testing of recovery is more commonly carried out for investigative than for clinical purposes, although Dixon and Thornton (1973) have developed a letter scoring test that may be applied in a clinical situation. These tests are usually more sensitive than clinical assessment in detecting abnormalities (Steinberg, 1954), but show learning effects and greater variability. Matched control groups may be necessary to allow the detection of changes. Doenicke et al. (1966) argued that as these tests may be influenced by fatigue, they may be less useful. However, if this were so, such a sensitivity would make these tests able to detect subtle drug effects. Grove-White and Kelman (1971b) found no fatigue effect in responses to a test of short-term memory, even though their subjects reported feeling exhausted.

Psychological tests have been roughly divided into those assessing higher mental functions alone, and those testing a combination of these factors and motor performance. Such a division may not be always appropriate, for the content of most motor tests is complex, as has been discussed above.

Tests involving perception, vigilance, and mental organization include the Bender face–hand test (Jaffe and Bender, 1951; Widdowson, Teresita and Virtue, 1955; Gale, 1957); the labyrinth test of Chapuis, the countdown test, the Duker’s test of repetitive calculations (all used by Doenicke, Kugler and Laub, 1967); recognition of objects in the picture vocabulary of the Stamford–Binet test (Burheim, Terry and Herron, 1966); and a letter scoring test (Rollason et al., 1971; Dixon and Thornton, 1973). Tests which also involve motor performance include the simple visual reaction time (Hendry, Norris and Nisbet, 1963; Vickers, 1965; Doenicke, Kugler and Laub, 1967), the track tracing test.
ASSESSMENT OF POSTOPERATIVE MENTAL FUNCTION

(Doenicke, Kugler and Laub, 1967), bead threading (Burheim, Terry and Herron, 1966), the pegboard test (Vickers, 1965) and the driving simulator (Egbert, Oech and Eckenhoff, 1959; Elliott et al., 1962; Green et al., 1963).

Although it is possible that temporary arousal for a short period may enable a subject to attain a high level of performance in these tests which involve less of higher mental function (Dixon and Thornton, 1973), sensorimotor coordination is in fact greatly impaired by subanaesthetic concentrations of nitrous oxide, as are more “complex” tasks (Steinberg, 1954). If the postanaesthetic pattern of wakefulness and drowsiness is similar to that found after sleep deprivation, then temporary improvement of all mental functions after arousal is possible, interspersed with periods of inattention.

Lindsley, Hobika and Etsten (1961) tested the continuous response of patients to an auditory signal which was reduced in intensity by the frequency with which the patient made a simple pinching movement. Recovery of this “operant response” to successive proportions of the preoperative level was measured, and was possible before the patients became verbally accessible. A combination of perceptive and motor responses was assessed by Newman and co-workers (1969, 1970) in a modification of the Bender Gestalt test. A Gestalt is a response to a group of stimuli that is perceived as a whole, and the test involves the copying of a geometric figure.

Grove-White and Kelman (1971b) followed the recovery from methohexitone, diazepam, and 4-hydroxybutyrate using a test of short-term memory.

Using the driving simulator as a complex test of performance, Elliott et al. (1962) noted that volunteers who were still influenced clinically by barbiturates did not all show a deterioration in the accuracy of their driving, and the same workers (Green et al., 1963) found no deterioration in driving performance after clinical recovery. Such findings are contrary to those of Drew, Colquhoun and Long (1958), who were able to correlate venous blood alcohol concentration with a reduced accuracy of steering and an increase in the amount of steering-wheel movement, in a similar test involving driving simulation. They also found that changes in performance were related to the personality of the subjects, extraverts being more affected than introverts.

Despite these findings with alcohol, estimation of venous blood concentrations of other agents seems to be of little value in assessing their duration of action. Hollister and Clyde (1968) found that the venous blood concentrations of pentobarbitone in volunteers did not correlate with assessments of subjective effects. The venous blood concentration present after awakening from thiopentone anaesthesia has been found to vary according to the dose given (Dundee, Price and Dripps, 1956). As the time course of changes in concentration of these agents in blood differs from the time course in other tissues, this discrepancy is not inconsistent with a dose-effect relationship in the brain.

The electroencephalograph has been used to detect cerebral effects of drugs, although different drugs vary in their effects on the e.e.g. (Gruber et al., 1957), and well-defined changes are only associated with full clinical anaesthesia. E.g. patterns cannot be correlated with symptoms or changes in memory caused by inhalation of subanaesthetic concentrations of nitrous oxide (Henrie, Parkhouse and Bickford, 1961). However, the electroencephalograph allows continuous monitoring of the level of consciousness and the recognition of disturbed sleep patterns or increased drowsiness for considerable periods after barbiturate anaesthesia (Doenicke et al., 1966).

The measurement of flicker fusion frequency has been used to assess recovery (Vickers, 1965; Smith, Allen and Perrin, 1967; Grove-White and Kelman, 1971a). Although the test requires carefully controlled conditions and subject cooperation, it is accurate if carefully carried out, the variation in response is small, and learning effects are minimal. In a study of the time course of action of diazepam and methohexitone, Grove-White and Kelman (1971a) considered that the test was too sensitive for the assessment of recovery.

CONCLUSION

As cerebral function is the most complex of body processes, it is unlikely that a “simple test” of impaired function can be devised. However, of the tests discussed, those that most nearly approach the criterion of a simple, reproducible, quantitative, sensitive and discriminating test seem to be short-term memory tests and the flicker fusion test.

As mental processes are interdependent, the clinical or social “relevance” of the test chosen may be less important than these other criteria.

In many studies it is possible that temporary arousal during the performance of a test could have
concealed impaired cerebral function. Such impairment will be unmasked if the subject is exposed to a situation where vigilance is necessary for a prolonged period. The reported incidence of drowsiness after "outpatient" anaesthesia suggests that this may be such a situation.

REFERENCES


