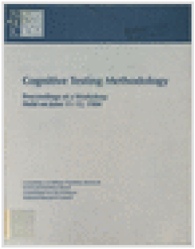


Cognitive Testing Methodology: Proceedings of a Workshop Held on June 11-12, 1984 (1986)

Pages
212

Size
8.5 x 10

ISBN
0309321743

Committee on Military Nutrition Research; Food and Nutrition Board; Commission on Life Sciences; National Research Council

Cognitive Testing Methodology

Proceedings of a Workshop Held on June 11-12, 1984

Committee on Military Nutrition Research
Food and Nutrition Board
Commission on Life Sciences
National Research Council

Order from
National Technical
Information Service,
Springfield, Va.
22161
Order No. PB86-244464

NATIONAL ACADEMY PRESS
Washington, D.C. 1986

NOTICE: The project that is the subject of this report was approved by the Governing Board of the National Research Council, whose members are drawn from the councils of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine. The members of the committee responsible for the report were chosen for their special competencies and with regard for appropriate balance. This report has been reviewed by a group other than the authors according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

The National Research Council was established by the National Academy of Sciences in 1916 to associate the broad community of science and technology with the Academy's purposes of furthering knowledge and of advising the federal government. The Council operates in accordance with general policies determined by the Academy under the authority of its congressional charter of 1863, which establishes the Academy as a private, nonprofit, self-governing membership corporation. The Council has become the principal operating agency of both the National Academy of Sciences and the National Academy of Engineering in the conduct of their services to the government, the public, and the scientific and engineering communities. It is administered jointly by both Academies and the Institute of Medicine. The National Academy of Engineering and the Institute of Medicine were established in 1964 and 1970, respectively, under the charter of the National Academy of Sciences.

This publication is based on a study performed pursuant to Contract DAMD 17-82-C-2306 between the U.S. Army Medical Research and Development Command and the National Academy of Sciences.

The view, opinions, and/or findings contained in this report are those of the authors and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

Order from
National Technical
Information Service,
Springfield, Va.
22161
Order No. PB 86-244464

WORKSHOP ORGANIZERS AND CO-CHAIRMEN

DAVID A. LEVITSKY, Division of Nutritional Sciences, Cornell University, Ithaca, N.Y.

RICHARD C. MOHS, Department of Psychiatry, Mount Sinai School of Medicine, and Psychiatry Service, Veterans' Administration Medical Center, Bronx, N.Y.

COMMITTEE ON MILITARY NUTRITION RESEARCH

ROBERT O. NESHEIM (Chairman), Science and Technology, Cambridge Plan International, Salinas, Calif.

ELLSWORTH R. BUSKIRK, Laboratory for Human Performance Research, Pennsylvania State University, University Park, Penn.

ALLAN L. FORBES, Nutrition and Food Science, Center for Food and Applied Nutrition, Food and Drug Administration, U.S. Department of Health and Human Services, Washington, D.C.

ARTHUR L. HECKER, Medical Nutrition Research, Ross Laboratories, Columbus, Ohio

D. MARK HEGSTED, New England Regional Primate Center, Harvard Medical School, Southborough, Mass.

JAMES R. KIRK, Research and Development, Campbell Soup Company, Camden, N.J.

DAVID A. LEVITSKY, Division of Nutritional Sciences, Cornell University, Ithaca, N.Y.

RICHARD C. MOHS, Department of Psychiatry, Mount Sinai School of Medicine, and Psychiatry Service, Veterans' Administration Medical Center, Bronx, N.Y.

PAUL C. RAMBAUT, Biomedical Research Program, National Aeronautics and Space Administration, Washington, D.C.

CATHERINE E. WOTEKI, Division of Health Examination Statistics, National Center for Health Statistics, U.S. Department of Health and Human Services, Hyattsville, Md.

ALLISON A. YATES, Division of Nutrition, Department of Community Health, Emory University School of Medicine, Atlanta, Ga.

STAFF

SAMUEL B. McKEE, Staff Officer, National Research Council, Washington, D.C. (Until June, 1985)

SUSAN E. BERKOW, Staff Officer, National Research Council, Washington, D.C. (July 1985-December 1985)

FRANCES M. PETER, Editor

CHRISTINE C. JANSON, Consultant Editor

FOOD AND NUTRITION BOARD

RONALD C. SHANK, Department of Community and Environmental Medicine and Department of Pharmacology, University of California, Irvine, Calif.

ROBERT H. WASSERMAN, Department/Section of Physiology, New York State College of Veterinary Medicine, Cornell University, Ithaca, N.Y.

MYRON WINICK, Institute of Human Nutrition, Columbia University College of Physicians and Surgeons, New York, N.Y.

SUSHMA PALMER, Executive Director, Food and Nutrition Board, National Research Council, Washington, D.C.

OVERVIEW: COGNITIVE TESTING WORKSHOP

Many investigators do not agree about whether, or the extent to which, nutritional factors affect the cognitive performance of military personnel. Although modern warfare increases the need for optimal cognitive functioning, military personnel must frequently perform under conditions of less than optimal nutritional support. To define the relationship between nutrition and cognitive performance more clearly, the Food and Nutrition Board's Committee on Military Nutrition Research within the National Research Council's Commission on Life Sciences sponsored a workshop at which experts on military performance, cognitive testing, and nutrition reviewed relevant data, identified key issues, and developed strategies for future research.

The first goal of the workshop was to ascertain the current state of knowledge concerning cognitive performance testing. Toward this end, participants were asked to present papers describing the variety of cognitive tests and test batteries that have been developed in recent years to assess specific aspects of human cognitive performance. The speakers also discussed the applications of these tests in various settings. It was the consensus of the workshop participants that tests now in use provide measures of important aspects of cognition likely to be affected by nutritional variables and that research programs designed to measure these effects could begin by conducting these tests and assessing the results. The participants also agreed that because human cognitive abilities are extremely complex, no single test or test battery can be designed to measure all aspects of the effects that nutritional factors can exert on cognitive performance.

A second goal of the workshop was to develop guidelines for selecting appropriate cognitive tests for use by the Department of Defense in its nutrition research program. After considerable discussion, participants recommended that three criteria be given primary consideration:

- Tests should be sensitive to the effects of nutritional variables. It would also be useful to determine the tests' sensitivity to other variables, such as sleep deprivation and drugs, which can also affect brain and body chemistry.

- Only tests of known reliability should be conducted. A reliable test was defined as one for which the major variables affecting performance are known and therefore can be controlled experimentally or treated as covariables.

- The tests should be relevant to military tasks. Although there are very few data correlating cognitive test scores with performance in combat or in other military environments, it was nevertheless agreed that measures of certain cognitive functions are likely to correlate with military performance during tasks with high cognitive demands.

The third goal of the workshop was to identify research topics with promise of improving our ability to assess the effects of nutrition on the performance of military personnel. Workshop participants identified two such subjects:

- There is a need to determine the sensitivity of existing tests to measure cognitive functions such as attention, memory and learning, mood, reading and comprehension, and self-monitoring of cognitive performance. The investigations should initially focus on nutritional variables that are likely to be encountered by military personnel and suspected of affecting cognition on the basis of previous observations such as those concerning acute water and caloric deprivation.

- The predictive power of laboratory tests of cognition needs to be determined for military personnel in situations involving heavy cognitive demands. Relevant data could be obtained by correlating the performance of military personnel on specific cognitive batteries with their performance under such demanding field conditions.

SESSION I

INTRODUCTION--CURRENT METHODS OF ASSESSING COGNITIVE PERFORMANCE

The purpose of this session was to review and discuss the current methods of assessing cognitive performance. Four areas of cognitive research are drawn upon to illustrate the kinds of tests that are available to the military for studying the effect of nutritional deficits on cognitive performance. These include tests of attention, memory, neuropsychology, and complex problem solving. Each of these areas is discussed in relation to the development of sensitive and reliable tests of cognitive function.

EVALUATING COMPLEX-PROBLEM-SOLVING ABILITY

Earl Hunt

The human claim to be king of the beasts rests on our ability to use reason rather than to fight, run, or breed. Because our power rests on thought, it behooves us to be aware of anything in our environment, including nutrition, that influences thinking ability. We can only be aware of how things influence our thought if we can measure how well we are thinking.

Measurement procedures are based on theories about the variable to be measured. Physical health could be measured by recording performance in athletic games, but few physicians would advocate such a procedure. A better picture of physical health can be obtained by breaking down physical performance by specific physiological systems. It is possible to take somewhat the same approach in analyzing human cognition. Mental functioning can be analyzed by isolating and studying the elementary information-processing capacities of the brain. This is the approach that justifies studies of memory, sensory discrimination, and the concentration of attention as legitimate parts of the study of cognition.

Although such analyses are useful, they represent indirect approaches to the measurement problem. No rational person is interested in the relation between, say, the amount of processed carbohydrates in a person's diet and that person's ability to retain lists of arbitrary letter-number pairs. The ability to learn arbitrary associations is only of interest to the extent that it is an indicator of a basic ability to learn to make associations that are useful in solving real-world problems. If this is the case, then testing learning by using lists of paired associates is analogous to testing cardiac function by using a blood pressure cuff. We see here the importance of theory in measurement. If a person's mental competence is an inevitable result of his or her competence in elementary information processing, then general competence can be established by measuring the various elementary capacities in isolation. Procedures can be devised to do this, and there is some truth in this approach. A compendium of an individual's basic abilities does provide some information about general cognitive ability, in particular about negative ability. People who are markedly deficient in a variety of basic information processes are unlikely to be capable (by human standards) of very complex thought, just as the classic 98-pound weakling is unlikely to be able to play professional football. However, the possession of adequate basic information-processing capacities permits but does not guarantee adequate or outstanding problem-solving ability. When people think, they synthesize their elementary information-processing abilities to solve problems. The quality of the synthesis may often be the major determiner of the quality of thought (1).

These rather abstract remarks can be given content by considering something that everyone would agree is an example of higher-order cognition, problems in elementary physics. Physics problem solving has been studied in detail, and several principles of problem solving have been identified in the course of the research (Chi _et al_., 1981; Larkin _et al_., 1980).

Physics problems are verbal statements describing a specific physical situation. For example:

> A block of weight w is placed on a plane of length _m_, inclined at an angle _a_ to a horizontal floor. If the coefficient of friction between the block and the plane is _c_, what will the block's speed be at the moment at which it strikes the horizontal floor?

When expert physicists see such problems, they first classify them in terms of the abstract physical principle involved (Chi _et al_., 1981). The problem given above involves the balance of forces. The label assigned to the problem cues the procedures for calculating the quantities important in solving the problem (Larkin _et al_., 1980). The answer to a problem may be derived during these calculations or it may be easily derivable from their results. If this is not so, the expert is likely to reconsider the problem, trying new classifications and accompanying procedures until the answer becomes apparent (Larkin _et al_., 1980). Some general principles can be extracted from this example. Both experts and novices call on several different information-processing capacities in problem solving. They must recognize previously seen stimuli, both at the physical level (words) and at the more abstract level (concepts, such as balance of force). An internal mental representation of the situation must be built up from a verbal description. This implies a capacity for comprehension of language and an ability to hold information in a temporary workspace, which will be called working memory. Most people will either imagine a diagram or draw one explicitly, implying an ability to manipulate visual-spatial images. Eventually, equations must be solved. This process requires the extraction of rules from long-term memory and their exercise in working memory.

A psychologist of the 19th century would say that physics problem solving draws on several faculties of the mind. Since the term faculty is somewhat archaic (2), the term latent trait will be used instead. Latent traits are abstract constructs, representing mental capacities that are drawn upon in producing overt, observable problem solving. Problem solving in physics is in no way unique in its drawing on several traits. Virtually all problem-solving situations depend on more than one mental ability. This is also true of intelligence tests. Most such tests are intentionally compendiums of subtests requiring different types of mental behavior, i.e., calling upon different traits (Wechsler, 1975).

All mental traits are eventually determined by biological status, because the exercise of a trait requires information processing, and information processing is a function of the brain. When performance is based on knowledge, the relationship between the brain, as an information processor, and the mind, as a problem solver, is greatly attenuated. This is an important point for two reasons. Knowledge-driven behavior is very important in human life, especially once the early schooling years have passed. Most of us earn our place in society by doing what we know how to do, not by being able to solve the esoteric problems presented on intelligence tests. Furthermore, knowledge is not limited to declarative knowledge about facts. People also acquire procedural knowledge about how to solve problems. Procedural knowledge may or may not be tied to a domain. Continuing with the physics illustration, the equation relating instantaneous velocity to accelerating force is a specific fact of limited applicability. The concept of examining the relationship between quantities to be conserved can be applied more widely, e.g., to problems in balance of forces, thermodynamics, and the analysis of electrical circuits. Treating a specific problem as a special case of a generalized form of reasoning, such as arithmetic or logic, is a very powerful technique that works outside of physics. It is also a learned technique that is apparently associated with exposure to Western schooling (Scribner, 1978).

In general, the more domain-specific a strategy is, the more efficiently it uses the brain's basic information-processing capacities. This is particularly true of information-processing demands on working-memory capacities. The most efficient way to solve a problem is to know the answer already, making no demands on problem-solving mechanisms or the physiology behind them. To complicate the issue, learning itself is a physiologically based process, and obviously the ability to learn is a necessary, though not sufficient, condition for the acquisition of knowledge.

To sum up, virtually any act of cognition forces the thinker to draw upon several latent traits. This has a strong implication for studies that wish to treat cognition as a dependent variable manipulated by noncognitive factors. We do not study the relationship between, say, cadmium levels in hair and performance on an intelligence test because we think cadmium affects test performance but because we believe that cadmium indicates a physiological latent trait that in some way controls an intellectual latent trait. The intellectual trait, in turn, is indicated by but is not identical to test performance. For some physiological variables the use of indicators does not unduly complicate the logic of the study, because the indicators are believed to be determined almost entirely by the latent trait of interest. This assumption is almost never justified in the analysis of cognitive behavior. Furthermore, it is almost never justified when sociological variables are introduced into the picture. (Does anyone really believe that the cultural milieu of a child's home is adequately indexed by family income?)

In the following section these general points are made concrete by examining three studies of the relation between consumatory behavior and cognition. The studies were chosen to illustrate measurement problems rather than to affirm or deny the authors' substantive conclusions. The first study uses a fairly simplistic approach in which inferences are drawn from the experimental manipulation of a single indicator. The second study uses a multiple-regression approach to make its point, but does not explicitly consider latent traits and their relationship. The third study does use a latent-trait approach.

EXPERIMENTAL MANIPULATION OF COGNITION

Many studies analyze the effects of an experimental manipulation on a single dependent measure of cognition. A good example is Pollit *et al.*'s (1983) study of the effects of short-term fasting on the cognitive performance of children. A within-subject design was used. Children aged 9 to 11 years were given the matching familiar figures (MFF) test (Kagan *et al.*, 1964) twice, once on days when they had eaten breakfast and once on days when they had not. The children were also given a standardized intelligence test once to use intelligence as a covariate in the design. It was found that more errors were made on the MFF test on the no-breakfast than on the breakfast day. The effect was found for children with scores both above and below the median on the intelligence test. The authors concluded (p. 173) that "MFF test findings in the past and present study are relevant to the educational setting because errors on this test correlate negatively with educational achievement measures." Two studies are cited indicating correlations of about -0.4 between errors on the MFF and scores on tests of basic school skills.

The authors' conclusions may or may not be true. The MFF test presents a cartoonlike target figure, such as a teddy bear sitting in a chair, and an array of six similar cartoons, one of which is exactly identical to the target. The others differ from the target in minor ways, e.g., in the arrangement of a ribbon on the bear. The task is to identify the exactly similar item. Passing the test is supposed to be an index of the ability to take an analytical perspective in examining pictures. Intuitively, the test seems to involve some but not all of the traits that would lead to success in school and to have some unique features of its own. This suspicion is reinforced by analyses presented by the test developers, who pointed out that errors on the MFF test could be predicted by the combined measures of intelligence and characteristic speed of responding. The combined predictors did better than either predictor alone (Kagan et al., 1964, p. 27).

Figure 1 presents an analysis of the Pollitt et al. (1983) results and summarizes the following points:

o The two observable dependent measures, MFF scores and school achievement scores, are produced by three latent traits, one common to the two tests and two measure-unique traits. This conclusion is based on the fact that the absolute value of the correlation between the two measures is well below the value it would have if they were both measures of the same trait (3).

o Short-term fasting influences one of the latent traits underlying the MFF test, but which one is not known.

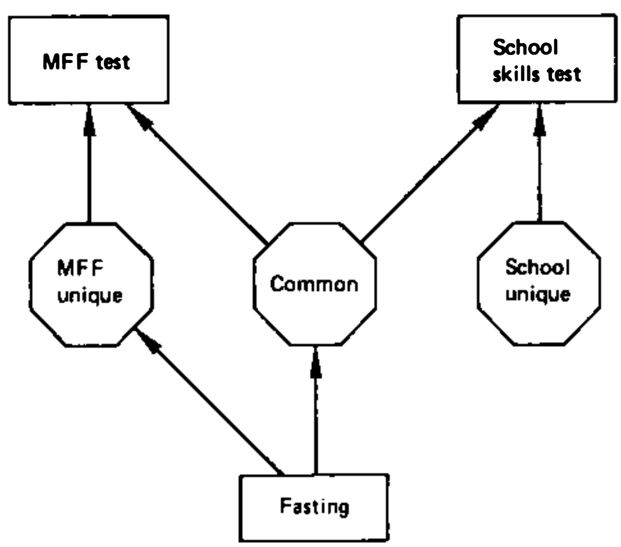


FIGURE 1. Different possible explanations of the results of Pollit et al. (1983). Fasting could influence factors common to school achievement and behavior on the MFF test or some factor unique to that test. In this and subsequent figures, observable variables are shown in rectangles and latent variables are shown in hexagons. Causality is indicated by a single-headed arrow. Correlation without implication of causality is shown by a double-headed arrow.

Note that the criticisms directed at this study have very little to do with the particular choice of the MFF score as a dependent variable. The same criticisms could be directed at any study that used a single dependent variable that was a less than perfect indicator of the variable of interest (4).

A MULTIVARIATE STUDY WITHOUT LATENT VARIABLES

Multiple-regression studies isolate a single dependent variable and study its relation to several independent variables. Experimental designs such as that used in the Pollit *et al*. (1983) study can be analyzed by multiple regression (Searle, 1971), so it is not surprising that multiple-regression studies share some of the problems of univariate experimental designs. This will be seen in the study to be discussed, a report by Lester *et al*. (1982) on the relation between refined carbohydrate intake and cognitive functioning in children.

Lester *et al*.'s study was correlational in the sense that it made use of natural covariation in variables. There were two questions of interest: does intellectual capacity covary (negatively) with the amount of refined carbohydrates in the diet, and if it does, is the linkage associated with a reduced ability to rid the body of metal pollutants? School children aged 5 to 16 years in rural Maryland were examined. Full-scale, verbal, and performance IQ scores were estimated by using a form of the Wechsler Intelligence Scale for Children (Wechsler, 1974). The wide-range achievement test (WRAT) (Jastak and Jastak, 1978) was also administered. This is a test of accomplishment in reading, spelling, and arithmetic. The independent variables were the amount of processed carbohydrates in the diet, estimated from self-report, and the amount of lead and cadmium in the children's hair. Age, sex, race, and a coding for family socioeconomic status (SES) were also recorded for use as covariates.

Full-scale IQ scores covaried negatively with the proportion of refined carbohydrates in the diet ($r = -0.25$). The reported-diet measure was statistically independent of (or, more precisely, not reliably related to) race, sex, age, and SES. The cadmium and lead content of hair both reliably related to IQ score (lead, $r = -0.34$; cadmium, $r = -0.31$). Cadmium, but not lead, was reliably related to carbohydrates in the diet ($r = 0.25$).

Lester *et al*. then performed an interesting and informative analysis. They computed a series of hierarchical multiple regressions using the three intelligence measures (full-scale, verbal, and performance IQ) and the three achievement measures as dependent variables. In each analysis predictions were first made with race, sex, age, SES, and hair cadmium level as predictors. The residual portion of the dependent variable was then related to the portion of processed carbohydrates in the diet. There was a reliable relationship (5) for every dependent variable except spelling. On the basis of this, Lester *et* al. concluded that the consumption of processed carbohydrates had a negative effect on mental competence beyond that due to its statistical (and possibly causative) association with high levels of cadmium in the body.

This study has two noticeable advantages over such studies as that by Pollit et al. (1983). Multiple measures of complex reasoning were used, and these measures could be regarded as being important in themselves rather than as indicators for some other variable. Certain technical statistical issues could be raised, but these probably do not affect the thrust of the results (6). A more relevant criticism is that Lester et al. might have presented a much clearer picture of their results if they had used a latent-trait analysis instead of restricting their attention to parallel regression analyses of individual variables. Let us consider why.

It can be argued that mental competencies fall into two broad areas, verbal and nonverbal reasoning. Note that the stress here is on competency and not the achievement-aptitude distinction implicit in separating the IQ and WRAT measures. The various dependent variables in the Lester et al. study could be looked on as indicators of these two latent traits, with each competency variously involved in each test. Similarly, the independent variables can be looked on as indicators of three classes of causal agent: metal pollutants, carbohydrates (which may be a partial cause of the variation in metal pollution), and family effects not associated with either carbohydrates or metal pollution. The causal diagram is shown in Fig. 2. If Lester et al. had attempted to trace this diagram, a more economic picture could have been obtained of the locus of both carbohydrate and metal pollution effects.

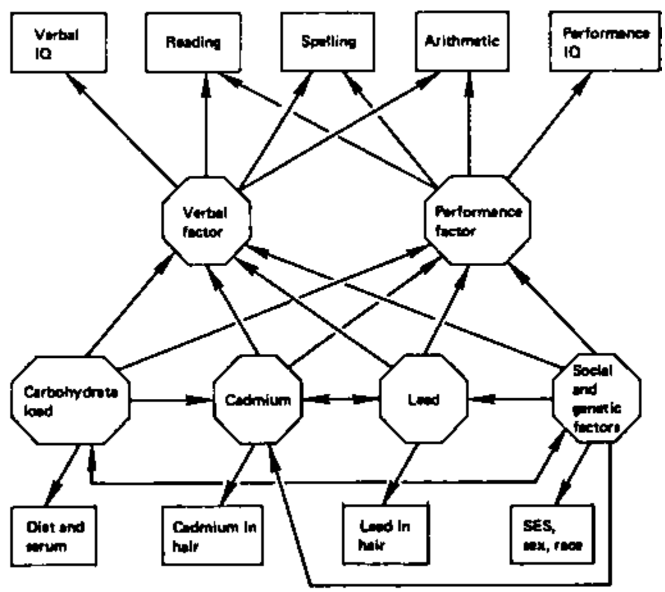


FIGURE 2. A possible causal model for the Lester et al. (1982) results. The symbols are explained in the legend to Fig. 1. Causative factors (carbohydrates, metal pollutants, social and genetic factors) may act selectively on verbal and nonverbal aspects of thought. These factors, in turn, may be differentially involved in various reasoning tasks.

I am arguing here for the statistical technique known as causal path analysis (Hiese, 1975). The reason for the name is clear. Figure 2 goes beyond Lester et al.'s descriptive results by proposing a detailed model of the manner in which metal pollutants and carbohydrates influence mental test scores. It is not clear that the statistical methods of causal analysis could have been applied to the Lester et al. study, which was not designed with this statistical technique in mind. In interpreting their results, however, Lester et al. introduced an implicit causal analysis. (Consider, in particular, their remarks on the independent effect of carbohydrates and pollutants and their comments on the relation of dietary and pollutant influences to learning disorders.)

AN EXPLICIT PATH ANALYSIS

A study conducted in my own laboratory (7) will illustrate the explicit analysis of latent variables. The purpose of the study was to explore the relation between intellectual performance and the extent of social drinking in an initially homogeneous group of nonalcoholic adults. Eighty University of Washington alumni (aged 33 to 63 years) participated in the study. Each person answered an extensive questionnaire concerning social drinking habits. They also took five tests that were chosen to tap two broad areas of mental competence, verbal and nonverbal reasoning. The three verbal tests were a vocabulary test, a task that required rapid verification of simple sentences describing equally simple pictures (Clark and Chase, 1972), and a test that required the subjects to hold information in memory while simultaneously comprehending rather complex sentences (Daneman and Carpenter, 1980). The two nonverbal tests were Raven's progressive matrices test (Raven, 1965) and the Wisconsin card-sorting test (Tarter and Parsons, 1971). In the matrix test the examinee must induce a pattern that links visual figures together. In the card-sorting test the examinee must induce the rule being used to sort geometric patterns into groups. In addition to taking these tests, the alumni gave their ages and allowed us to record their first- and second-year college grade point averages.

The verbal and nonverbal tests were treated as (imperfect) indexes of the respective underlying (latent) reasoning traits. The drinking questionnaire had three key questions: how often do you drink, how much do you drink on those occasions when you drink, and what was the greatest amount drunk on one occasion in the past year? For brevity, these will be referred to as the how often, how much, and most questions. The questions were asked about both present and past practices. Rather than use this information to derive an index of estimated weekly consumption, we treated the answer to each question as an imperfect indicator of a general use-of-alcohol latent trait. Two such traits were defined, one for current practices and one for the cumulative effect of drinking. Finally, the freshman and sophomore grades were regarded as indicators of a past-achievement trait, and reported age was treated as a perfect indicator of an age (latent) variable.

To summarize, the observable variables were assumed to define four underlying independent variables, age, past achievement, total drinking, and present drinking, and two dependent variables, verbal and nonverbal reasoning (Fig. 3). The model under investigation allows correlations to be made within the sets of independent and dependent latent traits and causal regression analysis to be done between independent and dependent traits.

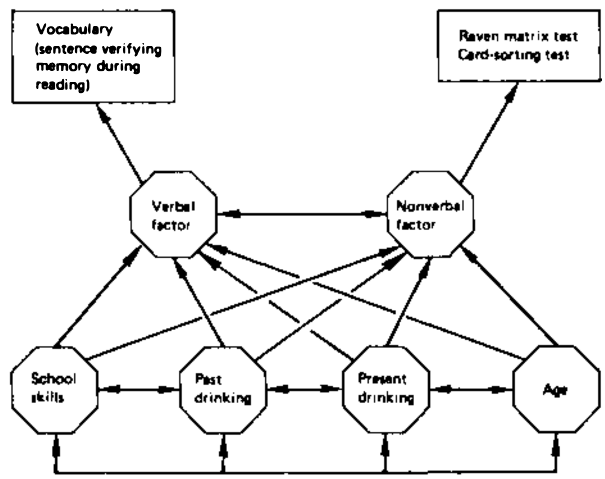


FIGURE 3. The causal model tested in the study of social drinking. The symbols are explained in the legend to Fig. 1. Drinking practices, age, and school achievement may be correlated and may vary in their influence on verbal and nonverbal factors in reasoning.

The analysis of covariance structures technique (Joreskog and Sorboom, 1979) was used to estimate the extent to which each observable measure served as an index of the appropriate latent trait and then to estimate the correlation coefficients between the various latent traits. For our purposes the most interesting of these coefficients is the set of causal regression coefficients relating independent and dependent variables (Table 1). There was a substantial positive relation between past (collegiate) achievement and current verbal performance. This finding repeats many similar findings and serves chiefly to indicate that our tests were appropriately sensitive to variation in ability. Second, and crucial to the study, there was no indication of a negative relation between current drinking practices and either verbal or nonverbal reasoning ability. This null finding is important since it contradicts some but not all of the reports in the literature on the effects of social drinking; however, discussing this issue would take us too far afield, for our interest centers on the methodological points.

The strength of the multivariate latent-trait approach is that it provides a method for exploring relationships between latent traits

TABLE 1. Correlation between questionnaire response and intelligence scores

Intelligence test type	Correlation (r)			
	Grades	Drinking		Age
		Current	Total	
Verbal	0.63 ± 0.27	0.32 ± 0.19	0.01 ± 0.17	0.09 ± 0.13
Nonverbal	0.32 ± 0.19	0.16 ± 0.18	0.20 ± 0.18	-0.34 ± 0.21

that are not exclusively bound to a single testing procedure. The analytic method may be used to contrast models that assume an independent (latent) variable is influencing a dependent variable directly against models that assume the independent variable is a correlate of a second variable that influences the dependent latent variable. This situation arises frequently in the study of cognition.

The strength of the latent-variable approach is also its weakness. Latent-variable analysis makes it possible to test specific models of the relationship between theoretical variables. However, latent-variable analysis does not lend itself to an apparently atheoretical, descriptive approach in the way that regression analysis does. As a result, the investigator can study fairly explicit theoretical models, but cannot conduct a very useful exploratory study.

There are also a number of difficult technical statistical issues that have been glossed over here. In the preceding paragraph it was stated that the approach may be used to contrast different causal models. Not all models can be so contrasted, even though the contrast makes sense at a nonmathematical level. They should not be overlooked in practice. Our study probably used the smallest number of subjects, 80, that could provide a reliable evaluation of a latent-variable model. Not all reasonable models can be analyzed at all. It is not easy to determine in advance of the study whether a particular model can be evaluated. Technical problems such as these may prove to be the limiting factor in the use of latent-variable analysis.

CONCLUDING COMMENTS

Complex problem solving is just that, complex. Problems are solved by executing procedures that use elementary information-processing capacities. The sythesis expressed in the procedure is crucial.

Therefore, a test of an external variable's influence on an elementary information-processing activity, in isolation, may fail to portray the external variable's influence on problem solving itself.

Problem-solving procedures are closely tied to the content of the problem being solved. There are two levels at which the term content can be understood. Content could refer to a specific domain of knowledge, such as physics, art history, or baseball. A more psychological distinction is that between verbal and nonverbal (visual-spatial and pure reasoning) problems. It is virtually certain that different information-processing actions are applied to internal representations in these different domains. Indeed, there is considerable evidence that different parts of the brain are used to manipulate visual and verbal representations. It would hardly be surprising, then, to find that environmental variables may have different influences on visual and verbal processes. On the other hand, in actual problem solving, visual and verbal representations are closely intertwined. Omar Khayyam's poetry invites the construction of a visual representation by comprehension of a verbal message. To understand the process of comprehension, it is necessary to study comprehension itself, not its bits and pieces.

The distinction between verbal and nonverbal problem solving has been used as an illustration of how the components of problem solving can be studied *in situ* with a multivariate approach. This is only one of several possible ways in which complex problem solving can be analyzed. Another frequently used analysis stresses the differential involvement of short-term (working) and long-term memory processes. The same principles of analysis apply: to study how the components work in problem solving, it is necessary to examine influences at the latent-trait level rather than assuming there is a single measurement procedure. A study by Gieselman *et al*. (1982) provides an excellent example of this approach.

The statistical procedures that are usually used to analyze latent traits were derived from factor analysis and are heavily influenced by linear regression concepts. The illustrations presented here relied on such methods. The concept of defining latent traits by multivariate analysis, however, is more general than the statistical procedures used to illustrate the approach. In particular, statistical analyses based on linear regression assume that the influence of the independent variable on the dependent variable does not depend on the level of the dependent variable. This is a very strong assumption about cognition, for what prevents poor thinking may not be what leads to unusually good thinking. But such problems are technical problems. They can be attacked by experimental design and perhaps by the development of new statistical methods.

Humans express their thinking in many ways. The student of thought must observe its multiple expressions.

NOTES

(1) This remark assumes that the expert and the novice begin in equivalent physical states. In practice, experts tend to be older than novices and thus may have less ideal brains _thought of as physical devices_. The fact that experts still outperform novices is an indication of the power of the content-specific mind over the content-free mind.

(2) Galen (quoted by Boorstin, 1983, p. 361) said, "So long as we are ignorant of the true essence of the cause which is operating, we call it a faculty." The concept of a faculty is no longer needed in medicine. Psychology is not so fortunate.

(3) A low estimate of split-half reliabilities of the MFF and the school skills test would be 0.8. Therefore, the expected value of the maximum correlation between these measures is 0.64, well above the 0.3 to 0.4 values cited by Pollit _et al_. (1983) in their literature review.

(4) An additional minor comment is worth making about the Pollit _et al_. study. They argue that their manipulation of the MFF is interesting because the MFF correlates with school skills. Their manipulation, however, was within-subject. The correlations between the MFF and the school skills variables were established by using between-subjects designs. To show that the within-subjects manipulation had predictive validity, Pollit _et al_. would have to show that the manipulated scores were more predictive than the unmanipulated scores whenever a subject was given a school skills test in the manipulated (here, no-breakfast) state.

(5) The partial correlation coefficients ranged from -0.13 for spelling to -0.29 for full-scale IQ.

(6) Throughout, Lester _et al_. treat a correlation as if it were zero when it has been found to "not be reliably different from zero."

(7) The study described was conducted in collaboration with P. Coggan, C. Berg, A. Davis, and P. Tianwattanatada.

REFERENCES

Boorstin, D.J. 1983. The Discoverers. Random House, New York.

Chi, M.T.H., P.J. Feltovich, and R. Glaser. 1981. Categorization and representation of physics problems by experts and novices. Cognit. Sci. 8:121-152.

Clark, H.H., and W.G. Chase. 1972. On the process of comparing sentences against pictures. Cognit. Psychol. 3:472-517.

Daneman, M., and P.A. Carpenter. 1980. Individual differences in working memory and reading. J. Verb. Learn. Verb. Behav. 19:450-466.

Gieselman, R.E., A.E. Woodward, and J. Beatty. 1982. Individual differences in verbal memory performance. A test of alternative information-processing models. J. Exp. Psychol. Gen. 111:109-134.

Heise, D.R. 1975. Causal Analysis. John Wiley & Sons, New York.

Jastak, J.F., and S. Jastak. 1978. Wide Range Achievement Test Manual. Jastak Association, Inc., Wilmington, Del.

Joreskog, K., and D. Sorboom. 1979. Advances in Factor Analysis and Structural Equation Models. ABT Books, Cambridge, Mass.

Kagan, J., B.L. Rosman, D. Day, J. Albert, and W. Phillips. 1964. Information processing in the child: significance of analytic and reflective attitudes. Psychol. Monogr. Vol. 78, No. 1 (Whole No. 578).

Larkin, J.H., J. McDermott, D.P. Simon, and H.A. Simon. 1980. Expert and novice performance in solving physics problems. Science 208:1335-1342.

Lester, M.I., R.W. Thatcher, and L. Monroe-Lord. 1982. Refined carbohydrate intake, hair cadmium, and cognitive functioning in children. Nutr. Behav. 1:3-13.

Pollit E., N.L. Lewis, C. Garza, and R.J. Shulman. 1983. Fasting and cognitive function. J. Psychiatr. Res. 17(2):169-174.

Raven, J.C. 1965. Advanced Progressive Matrices, Sets I and II. Lewis and Co., London.

Scribner, S. 1978. Models of thinking and ways of speaking: culture and logic reconsidered. Pp. 483-500 in P.N. Johnson-Laird and P.C. Wason, eds. Thinking: Readings in Cognitive Science. Cambridge University Press, New York.

Searle, S.R. 1971. Linear Models. John Wiley & Sons, New York.

Tarter, R.E., and O. Parson. 1971. Conceptual shifting in chronic alcoholics. J. Abnorm. Psychol. 77:71-75.

Wechsler, D. 1974. Wechsler Intelligence Scale for Children, Revised. The Psychological Corporation, New York.

Wechsler, D. 1975. Intelligence defined and redefined: a relativistic appraisal. Am. Psychol. 30:135-139.

TESTS OF ATTENTION

Christopher D. Wickens

The word attention has a wide number of meanings, a fact attested to by the diversity of chapters in Varieties of Attention (Parasuraman and Davies, 1984). For the purposes of the present paper, attention will be defined in three contexts (Wickens, 1984).

Focused attention refers to the ability to hold attention riveted on one task or source of information in the face of distraction from irrelevant stimuli. (This report will not consider the issue of sustained attention, sometimes used synonymously with vigilance, in which the only distraction from the relevant information source is sleep, boredom, or fatigue).

Selective attention refers to the ability to select relevant tasks and information sources from the environment when they are important. It includes two components: knowing when to sample what in an optimum fashion, and how rapidly and efficiently attention can be switched from one information source to another.

Divided attention refers to the ability to perform two tasks (or process two sources of information) concurrently as well as either task (or information source) can be performed (processed) alone. It is a concept used synonymously with time-sharing.

All three of these aspects of attention have in common the presence of two or more tasks or information sources. Some aspect of performance is measured in this dual task or dual source environment, and then variance in performance here is compared with that in the single environment. If there is variance in the dual environment that cannot be accounted for by variance in the single environment, then there is reason to assume the existence of an attentional phenomenon that could and should be assessed by a test. On the other hand, if all dual variance can be accounted for by single variance, then there is no need to seek particular attention measures, since measurements of the tasks in isolation provide all relevant information.

In the following pages, three sources of variance (single and dual) will be considered: variance between individuals--the issue of individual differences in attention ability; variance due to practice--the issue of whether special attention skills develop with training; and variance due to stressors such as noise or drugs. Is performance in the dual environment influenced by stressors in a qualitatively different way from performance in the single environment? This report will first review the literature in a three-by-three matrix form defined by the three kinds of attention and the three sources of variance. Studies that have looked for but not necessarily found dual variance will be discussed. The second part of the report will attempt to identify or propose those tests that will best tap the dual abilities whose existence was established in the first part.

FOCUSED ATTENTION

Individual Differences

Individual differences in the ability to focus attention on one source of information in the face of distraction from competing stimuli have been addressed in five investigations. Davidson et al. (1976) administered the Tellegen absorption test (Tellegen and Atkinson, 1974) to a sample of subjects. This test, used originally to assess openness to hypnotic influence, purports to measure a subject's ability to become completely absorbed in a source of environmental stimulus. Subjects with high and low scores on this scale were then asked to count either visual or tactile stimuli while ignoring concurrent stimuli in the opposite modality. The investigators found that subjects who had scored high on the absorption scale showed less electroencephalographic (EEG) processing activity over the cortex of the to-be-ignored modality than did subjects who scored low in absorption. This was particularly true when the visual channel was to be ignored.

The Stroop task, in which the subject is asked to report the color of ink in which a series of color name words are printed, is a classic demonstration of failure to focus attention. For example, when the word red is printed in blue ink, it is exceedingly difficult to ignore the meaning of the word and focus only on the color. In a recent investigation by Dupree (1984), 56 subjects were tested on the Stroop task, along with similar tests of focused attention and a battery of spatial tests. Performance on the Stroop task--a measure of the ability to resist distraction--showed a high degree of variation across subjects and was correlated highly with a spatial factor related to perceptual speed, as if the ability to focus attention is greatly facilitated by faster processing of relevant information.

Finally, four investigations have examined focused attention in a dichotic listening task, in which the subject must focus attention on one of a pair of word strings presented simultaneously to both ears. Lansman et al. (1983) had subjects press a button when a specified target letter appeared in the predesignated relevant ear. They found that differences in performance (latency and accuracy) in this focused-attention condition were perfectly predicted by performance in a single-task control condition. In contrast, a dichotic listening task employed by Gopher and Kahneman (Gopher, 1982; Gopher and Kahneman, 1971) showed a greater degree of predictive ability. Their dichotic listening task was separated into two phases. In the first phase, subjects heard a series of dichotically presented Hebrew verbs interspersed with occasional digits. Their task was to report the digits when they occurred in the relevant ear. Gopher and Kahneman found errors of both intrusion and omission. When the task was administered to flight cadets (100 cadets in Gopher and Kahneman, 1971; 2,000 cadets in Gopher, 1982), performance on both measures predicted the level of proficiency obtained by the cadets in flight training.

Using a smaller number of subjects (60), Braune and Wickens (1983, 1984) developed and administered an English version of the dichotic listening task. They also found robust individual differences in the focused attention measure. However, this study failed to obtain reliable correlations of this measure with performance proficiency in a flight simulator.

Focused Attention Under Stress

There is by now a large body of evidence suggesting that the focus of attention becomes more restrictive or narrow under increased levels of stress and arousal (see Broadbent, 1971; Hockey 1979, 1984; Kahneman, 1983, for a review). Typical studies in this area present subjects with a task in foveal vision and a concurrent task in peripheral vision. Arousal or stress is varied by imposing stressors such as noise, and performance on the peripheral tasks is found to deteriorate more than on the central tasks (e.g., Hockey, 1970; Weltman *et al.*, 1971; Williams, 1985); in fact, sometimes foveal performance does not decline at all.

In an interesting combination of this perceptual narrowing phenomenon with the Stroop task, investigators have observed that Stroop performance actually improves when an arousal increase restricts focal attention, eliminating the distraction of the irrelevant stimulus (the color name in the Stroop task). Houston (1969) has observed this by using noise as a stressor, and Calloway and Stone (1960) have done so by using drugs to influence arousal.

SELECTIVE ATTENTION

Individual Differences

There have been a number of studies indicating that the ability to switch attention from one information source to another is both a relatively enduring trait and an important component in complex multitask performance. During phase 2 of the dichotic listening task used by Gopher and Kahneman (1971), the relevant ear on which target digits are to be reported is suddenly changed (a change signaled by a tone), and subjects must now report digits along the new channel. Reporting errors are then recorded immediately after the switch as an index of how rapidly and effectively subjects can switch attention. Gopher (1982) reports that this switching measure does a better job than the focused attention measures of predicting success in flight training and accounts for variance unaccounted for by any of the traditional pilot selection tests. Braune and Wickens (in press) found reliable individual differences due to age in the switching measure of the English dichotic listening task, but failed to find that this measure correlated with success in time-sharing a reaction time and tracking task.

A slightly different approach to attention switching was taken by Keele and Hawkins (1982). In addition to assessing the dichotic listening switching measure, these investigators also assessed three other measures of internal switching. Internal switching refers to the ability of a subject who expects one kind of stimulus to process the unexpected occurrence of a different kind arriving from the same physical location. Keele and Hawkins found reliable correlations between all four of these switching measures, a finding they interpreted to support the existence of a single "attentional flexibility" trait. Unfortunately, however, their sample size (18) was lower than is optimal for correlational work, but the result is certainly consistent with the other data, suggesting that switching is an important source of individual differences.

An investigation conducted by Braune and Wickens (1985) examined the components of individual differences in dual-axis tracking, as control along the two axes was varied in different ways. Factor analysis on their data revealed one factor that was defined by task components in which subjects had to switch visual attention (i.e., scan between two displays separated by a wide visual angle).

Training. The second important aspect of selective attention is knowing when and where to allocate attention. Theoretical developments and research by investigators such as Senders (1964) and Moray (1981) have suggested that this ability depends critically on the operator's "internal model" of the statistical properties of the environment, a model that will guide attention to focus at the right place at the right time. Practically all work in this area has to be conducted in the area of training and skill development. These studies find that as expertise with the multichannel environment develops, subjects come closer and closer to allocating attention according to normative optimal policies (e.g., Moray et al., 1976; Senders, 1964). Typical is Senders's (1964) study, in which subjects monitored a set of dials moving in random paths at varying speeds; the speeds defined different information rates. Senders used a camera to discern the moment-by-moment orientation of foveal vision (and attention) and traced the improvement in allocation policy as subjects became more familiar with the statistical properties of the signals.

Stress. The findings related to the effects of stress on selective attention are basically the same as those reported for focused attention. Stressors that increase arousal typically cause attention to be directed to environmental information sources that are central, important, or salient at the expense of those that are peripheral or unimportant or have low physical and psychological salience.

DIVIDED ATTENTION

The ability of an operator to divide attention between two tasks or mental activities is clearly a critical facet of complex-task performance. We have probably all encountered individuals who are able to perform a multitude of activities at once and others who perform in a very "single-channel" fashion. The question is, what processing differences distinguish these two groups? The initial thought is a difference in a time-sharing ability. However, before postulating such an ability it is necessary to discount the possibility that good and bad time-sharers differ in terms of their single-task skill level (Damos and Wickens, 1980; Wickens and Benel, 1981). If two individuals differ in their level of skill on a given task, it is quite likely that the more practiced will require fewer attentional resources to perform the task. This in turn will leave more residual attention available to perform concurrent activities and so lead to more effective time-sharing (Norman and Bobrow, 1975). The source of this difference in time-sharing efficiency is inherent in the single-task skill level, not in a time-sharing skill or ability.

Individual Differences

To demonstrate that individuals differ in their basic time-sharing skills, three related approaches have been used. The first, called the correlational approach, is to assess performance across a number of subjects under both single and dual task conditions and then correlate single-task performance measures with some measure of time-sharing efficiency (i.e., the decrement in performance resulting from task concurrence). If these correlations are low, then there is a good deal of variance in time-sharing that cannot be accounted for by single-task skill levels, i.e., the variance then is attributed to time-sharing skill. Of course, as described above, there remains the possibility that this variance in time-sharing efficiency is simply the result of differences in single-task automation that are not reflected in single-task performance levels. To guard against this possibility, investigators should seek small differences in the way the tasks are time-shared to identify the factors underlying the ability.

A second approach is to seek variance in the performance of a complex criterion task (such as flying) that can be accounted for by variance in dual-task performance of predictor tasks but not in single-task performance. This may be labeled the complex criterion approach.

The third approach, which offers the strongest guarantee that individual differences in time-sharing efficiency do not reflect single-task automation, is to determine whether people who are efficient time-sharers on one dual-task combination (e.g., dual-axis tracking) are also efficient on a qualitatively different dual-task pair (e.g., reading while listening) so that there are no common elements between the two pairs of component tasks. Since it is less likely that people who are highly skilled in the single-task components of one task combination are also skilled in the components of the other, it is quite unlikely that automation is the common element shared by the good time-sharers. The usual approach taken here is factor analysis, in which a factor is sought that loads exclusively on dual-task measures of time-sharing efficiency. This then is the factor analysis approach.

Failure to find evidence for time-sharing ability from the factor analysis approach does not mean that such an ability does not exist; it only indicates that this factor is not a general one. It may well be that there are specific time-sharing abilities (specific to a restricted pair of tasks) that can be revealed by the correlational or complex criterion approaches. In this sense the situation is analogous to the issues in intelligence testing of whether there is a general (G) factor or a series of specific (S) factors of time-sharing. Only the factor-analytic approach can establish the presence of the G factor. The experiments described below will first identify positive evidence for some component related to time-sharing ability. Then, experimental evidence will be reviewed that considers whether this ability is general.

North and Gopher (1976) used the complex criterion approach and seemed to find evidence for at least an S factor. In their experiment, 40 flight-training students performed a tracking task concurrently with a reaction time task. In different conditions the emphasis of one task over the other was varied, and measures of attention manageability (how well performance changed with change in task emphasis) were derived in addition to measures of overall dual-task performance. North and Gopher found that both the manageability scores and dual-task performance measures accounted for variance in flight-training proficiency that was not accounted for by single-task performance measures.

Damos et al. (1983) followed up observations made in an earlier study by Damos and Wickens (1980) and used the correlational approach to identify an apparently stable difference in the ability to time-share two discrete information-processing manual tasks. The investigators identified two different strategies of processing. A massed strategy used by some subjects forced the processing of a series of stimuli from one task followed by the stimuli from the other, i.e., these subjects appeared to be unable to process in parallel. The other subjects, who used a simultaneous strategy, appeared to be quite capable of processing information from the two tasks in parallel, either executing simultaneous responses to both or rapidly alternating between responses to the two.

The study by Braune and Wickens (in press), discussed previously in the context of dichotic listening, also seemingly demonstrated the existence of time-sharing ability. Their subjects performed a series of different versions of the Sternberg memory search task, alone and concurrently with tracking. Correlational and factor analysis of single and dual task scores again revealed patterns of variance in the latter that could not be accounted for by the former. Hunt and Lansman (1981) conducted a study in which they asked subjects to perform three task configurations: a secondary task (probe reaction time) concurrently with an easy version of a primary task (a "keeping track" running memory task), the secondary task by itself, and a difficult version of the primary task by itself. They found that a secondary-task measure of reserve capacity, estimated while the subject was performing the easy primary task, could predict individual differences in performance of the difficult primary task. They used this observation to support the argument that subjects differed in the amount of general capacity they had available to deploy in the service of either task.

Finally, an investigation conducted by Stankov (1983) argued for the existence of a time-sharing ability when two auditory tasks are performed dichotically. The tasks represented different kinds of judgments made on two different tone sequences, one presented to each ear. When single- and dual-task correlations were computed between the component tasks, the dual-task correlations were found to be higher. This finding suggests that there is common variance across tasks that only emerges when tasks are time-shared or performed when attention is diverted. Stankov also argued that this attention ability may be related to general intelligence.

The five investigations reported above clearly point to the existence of time-sharing ability. Discussed in the following pages are investigations that have tried to determine whether this is a G factor of time-sharing by using variations on the factor-analytic approach. Studies that have obtained negative answers to this question will be discussed first, followed by those showing positive results.

Negative evidence for a G factor. Sverko (1977) assessed single- and dual-task performance on four tasks performed singly and in all possible dual-task combinations. The four tasks were pursuit rotor tracking, a visual-manual digit-processing task, a counting-backwards task, and an auditory discrimination task. Factor analysis of the results failed to suggest any general time-sharing factor, and the correlations between the time-sharing efficiency measures of all pairs of dual-task combinations that had no common elements were near zero. A subsequent investigation by Sverko *et al*. (1983) used the same general approach, looking at correlations between single and dual task measures of three different dual-task pairs. Sverko *et al*. did find common dual-task variance between two of the pairs (unaccounted for by single-task variance). However, the tasks making up these two pairs actually had quite a bit in common, both involving the pairing of two visual-motor reaction time tasks. Thus, it appears that the factor identified by Sverko *et al*. was probably similar to that found by Damos *et al*. (1983).

Jennings and Chiles (1977) used the same approach with six visual-manual information-processing tasks performed singly and in two triple-task combinations. The tasks ranged from simple monitoring to more complex problem solving. Factor analysis of their data provided no evidence for a general factor, but did point to the existence of a specific factor related to monitoring in multitask situations. This factor would seem to relate to visual scanning strategies, similar to those that reflect the internal model.

Hawkins *et al*. (1979) used a slightly different approach to reject the generality of time-sharing abilities. Their subjects performed a series of dual-task reaction time trials using the double-stimulation paradigm, in which the two stimuli begin a very short time apart. Stimulus and response modalities of one of the tasks, along with its difficulty, were varied over the trials, and a measure of time-sharing efficiency was derived. Hawkins *et al*. found that this measure correlated fairly highly across pairs of dual-task conditions that were similar in terms of input-output modality and difficulty, but the correlations of efficiency between dual-task pairs that were quite different from each other were near zero, indicating the absence of a G factor. Subsequent analysis of their correlations suggest that one specific ability relates to the switching of attention between auditory and visual input.

An investigation by Wickens *et al*. (1981) used a design identical to that used by Sverko (1977). The four tasks were a critical instability tracking task, an auditory running memory task, a visual digit classification task, and a visual-spatial line extrapolation task. Using factor-analytic and correlational procedures similar to those employed by Sverko, the investigators found no evidence for the G factor, but did observe one factor that seemed to be related to scanning strategies (time-sharing two visual tasks, one of which had high account demands) and a second factor related to the sharing of auditory and visual tasks. These findings converge with the S factors suggested by the data of Jennings and Chiles (1977) and Hawkins *et al*. (1979), respectively.

Finally, the investigation carried out by Lansman *et al*. (1983), as described in the context of focused attention measures, also provided evidence against a general ability related to attention switching. The authors presented single- and dual-task (divided attention) trials of auditory and visual detection. In the dual-task trials, the two visual or two auditory stimuli were presented simultaneously, so that attention had to be switched between sources. The authors' correlational and factor analyses revealed that practically all variance in each dual-task condition could be accounted for by single-task variance in the corresponding modality. There was no evidence for a general switching ability that transcended modalities.

Positive evidence for a G factor. Ackerman et al. (1984) examined the analytical procedures used in the studies discussed in the previous section and found a number of flaws that they believed might have prevented the investigators from finding a G factor for time-sharing. In particular they criticized the use of varimax rotational procedures in the factor analysis, which serve to minimize the likelihood of finding small sources of variance that might exist across all tasks (i.e., a general time-sharing factor). Reanalyzing the data from Wickens et al. (1981) with a Procrustean technique, they found that the data were consistent with the existence of a general time-sharing ability underlying the various dual-task combinations.

Fogarty and Stankov (1982) administered a series of 16 task configurations to 91 subjects. These included 10 single tasks and 6 dual-task combinations. All tasks were auditory, involving different operations performed on either letter series or tone series, and presented dichotically in the dual-task configurations. In each of the dual-task trials, one task was emphasized and the other deemphasized. Factor analysis revealed that primary-task performance variance was accounted for by single-task variance, but that secondary-task variance was not, thereby presumably reflecting a time-sharing factor. How general this factor is cannot be established from these data because all dual-task configurations had in common the sharing of auditorily presented letter material with auditorily presented tonal material.

The data collected by Braune and Wickens (1985) offer tentative evidence for a small general factor. Different experimental manipulations in a dual-axis tracking task were used to assess different components of attention when the two axes were to be tracked concurrently. Components related to attention switching, to the use of an internal model guiding fixation, to parallel processing, and to simultaneous performance of two dissimilar operations were each assessed. Across the 40 subjects, it was found that the first and third of these components were negatively correlated; subjects who were particularly proficient at carrying out dissimilar activities in parallel were not good at discretely switching between information sources, and vice versa.

Braune and Wickens also found that different components of dual-axis tracking could account for a small amount of variance in the performance of two dissimilar tasks, memory span time-shared with a Sternberg memory search task.

In summary, the data provide consistent evidence for the existence of at least some specific time-sharing abilities. These would seem to include (but not be limited to) the sharing of two discrete manual tasks, the time-sharing of continuous tracking and discrete reaction time, the allocation of visual attention between different stimulus sources, and the switching of attention between two auditory sources, and between an auditory and visual source. It appears likely that one element underlying the ability found by Fogarty and Stankov (1982) may be the same auditory switching ability revealed in the dichotic listening task of Gopher (1982). The existence of the G factor is less certain, with only the evidence provided by Ackerman et al. (1984) offering definitive support for the concept. Any general ability that does emerge will probably relate to a specific or "meta" skill in processing, such as visual attention switching, that will apply to some but not all dual-task combinations. The existence of fairly general time-sharing abilities is supported to some extent by studies cited in the following section that point to a general time-sharing skill.

Training and Divided Attention

Two investigations here seem to have gone farthest in establishing the existence of time-sharing skills acquired through practice. Damos and Wickens (1980) required subjects to perform a pair of discrete digit-processing tasks (one involving memory, the other classification) singly and in combination. With practice, dual-task performance continued to improve even as single-task performance remained constant. A detailed analysis of the dual-task performance data suggested that the former improvement was not simply the result of single-task automation. Further evidence that a time-sharing skill developed, and that it was a somewhat general skill, was provided by observing that subjects who had undergone this dual-task training with two discrete tasks showed positive transfer to a two-axis tracking task, as if the skill acquired in dual discrete processing assisted performance in the time-shared continuous tasks.

Reick et al. (1980) also found evidence for the acquisition of time-sharing skills. They administered subjects different amounts of single- and dual-task practice on a digit classification task paired with a discrete tracking task. Final dual-task performance was related to the amount of dual-task but not single-task training, suggesting that only the former was relevant for developing the time-sharing skill. Equally relevant to the generality issue was their finding that subsequent dual-task performance on a different dual-task combination (continuous tracking and choice reaction time) benefited from the amount of dual-task practice on the first pair.

Thus, it appears that practice in a dual-task setting develops at least some important time-sharing skills that cannot be acquired through single-task practice. Furthermore, the available evidence suggests that there is at least some generality to these skills.

Stress and Dual-Task Performance

Studies that have examined the effects of arousal and stressors on dual-task performance seem to posit two opposing influences, described by Kahneman (1973). On the one hand, stressors that increase arousal will make available more capacity, resources, or attention to deal with competing tasks. On the other hand, this arousal increase will narrow the focus of attention, presumably excluding tasks of lesser importance, as described earlier in this paper. Thus, in a primary-secondary task paradigm, we can expect stressors to increase performance on a primary task but decrease it on a secondary task.

RECOMMENDED TESTS OF ATTENTION

The preceding review of the literature leaves little doubt that there are important attention-related phenomena that cannot be predicted on the basis of single (task or stimulus) measures. Unfortunately the data probably remain too scattered to determine precisely what the optimal tests of these phenomena should be. Nevertheless the review indicates certain logical choices that have surfaced in one form or another. Certainly the dichotic listening test as a measure of both focused attention and switching is a candidate because it is well validated and well understood and because switching itself may underlie many different measures of time-sharing ability. However, certain precautions are important when developing versions of the test related to the timing of the signal to shift the relevant ear in phase 2. If the interval between this signal and the first phase 2 stimulus is too long (more than 0.5 s), the shift may become trivially easy to accomplish, and error measures will become too low to provide much evidence of variance.

Second, it seems that a test of visual environment sampling should be included in any attention testing battery. Such a test would not assess the speed of visual scanning but rather the ability of subjects to use statistical properties of the environment to their advantage--i.e., to develop and deploy the internal model (Moray, 1984; Moray and Fitter, 1973). It would also appear important to include at least one task of visual focused attention, such as the Stroop task.

Third, there is sufficient evidence for a divided attention time-sharing ability distinct from single-task abilities to warrant inclusion of at least three dual-task configurations. Candidates would include a dual discrete manual task configuration, such as that used by Damos et al. (1983), a discrete task time-shared with tracking, such as that employed by Braune and Wickens (1983) and by North and Gopher (1976), and a dual-axis tracking task as used by Damos and Wickens (1981).

At least one of these dual-task configurations should also be assessed under two conditions of priority allocation, one favoring one task and one favoring the other. On the one hand this will allow the investigator to assess attention manageability scores as done by North and Gopher (1976), clearly an important attentional concept. On the other hand it will allow examination of the predicted priority effects of stressors and arousal on primary and secondary task performance. Whenever a dual-task configuration is employed, it is essential to measure single-task performance as well.

There are probably other candidates that should be included, although at present not enough data exist to indicate precisely what these should be. One possibility that would seem to be important, however, relates to subjects' abilities to shift rapidly between two different mental-cognitive functions. As noted, Braune and Wickens (in press) have recently found that a dual-axis tracking task in which heterogeneous dynamics are used seems to tap this ability.

In conclusion, much more research is needed on attentional tests. This reflects the relatively limited number of studies in the literature, as well as the complexity of dual-task performance measures and the problems associated with measuring a mental phenomenon such as attention that cannot be observed but must be inferred from performance.

REFERENCES

Ackerman, P., W. Schneider, and C. Wickens. 1984. Deciding on the existence of a time-sharing ability. Hum. Factors 26:71-82.

Braune, R.J., and C.D. Wickens. 1984. Individual differences and age-related performance assessment in aviators II. University of Illinois Engineering Psychology Laboratory Technical Report No. EPL-83-7/NAMRL-83-2. University of Ill., Psychology Dept. Champaign, Ill.

Braune, R.J., and C.D. Wickens. 1985. Components of attention in dual axis tracking. In R. Jensen, ed. Third Symposium on Aviation Psychology. Ohio State University, Columbus.

Braune, R.J., and C.D. Wickens. In press. The functional age profile: an objective decision criterion for the assessment of pilot performance capacities and capabilities. Hum. Factors, Vol. 27.

Broadbent, D. 1971. Decision and Stress. Academic Press, London.

Calloway, E., and G. Stone. 1960. Re-evaluating the focus of attention. Pp. 393-398 in L. Uhr and J.G. Miller, eds. Drugs and Behavior. John Wiley & Sons, New York.

Damos, D., T. Smist, and A.C. Bittner. 1983. Individual differences in multiple task performance as a function of response strategies. Hum. Factors 25:215-226.

Davidson, R., G. Schwartz, and L. Rothman. 1976. Attentional style and the self regulation of mode-specific attention. J. Abnorm. Psychol. 85:611-621.

Damos, D., and C.D. Wickens. 1980. The acquisition and transfer of time-sharing skills. Acta Psychol. 6:569-577.

Dupree, D. 1984. Individual differences in visual information processing. Ph.D. dissertation. University of Illinois, Champaign.

Fogarty, G., and L. Stankov. 1982. Competing tasks as an index of intelligence. Personal. Indiv. Differences 3:407-422.

Gopher, D. 1982. A selective attention test as a prediction of success in flight training. Hum. Factors 24:173-184.

Gopher, D., and D. Kahneman. 1971. Individual differences in attention and the prediction of flight criteria. Percept. Motor Skills 33:1335-1342.

Hawkins, H.L., E. Rodriguez, and G.M. Reicher. 1979. Is time-sharing a general ability? University of Oregon Center for Cognitive and Perceptual Research, Office of Naval Research Technical Report No. 3, June 1979. University of Oregon, Eugene.

Hockey, G.R.J. 1970. Effect of loud noise on attentional selectivity. Q. J. Exp. Psychol. 22:28-36.

Hockey, G.R.J. 1979. Stress and the cognitive components of skilled performance. Pp. 141-178 in V. Hamilton and P.M. Warburton, eds. Human Stress and Cognition: An Information-Processing Approach. John Wiley & Sons, New York.

Hockey, G.R.J. 1984. Varieties of attentional state. Pp. 449-484 in R. Parasuraman and D.R. Davies, eds. Varieties of Attention. Academic Press, New York.

Houston, B. 1969. Noise, task difficulty, and Stroop color-word performance. J. Exp. Psychol. 82:403-404.

Hunt, E., and M. Lansman. 1981. Individual differences in attention. In R. Sternberg, ed. Advances in the Psychology of Intelligence. Vol. 1. Erlbaum Associates, Hillsdale, N.J.

Jennings, A.E., and W.D. Chiles. 1977. An investigation of time-sharing ability as a factor in complex performance. Hum. Factors 19:535-547.

Kahneman, D. 1973. Attention and Effort. Prentice-Hall, Englewood Cliffs, N.J.

Keele, S.W., and H.L. Hawkins. 1982. Explorations of individual differences relevant to high level skill. J. Motor Behav. 14:3-32.

Lansman, M., S. Poltrock, and E. Hunt. 1983. Individual differences in the ability to focus and divide attention. Intelligence 7:299-312.

Moray, N. 1981. The role of attention in the detection of errors and the diagnosis of errors in man-machine systems. Pp. 185-198 in J. Rasmussen and W. Rouse, eds. Human Detection and Diagnosis of System Failures. Plenum Publishing Corp., New York.

Moray, N. 1984. Attention to dynamic visual displays in man-machine systems. Pp. 485-513 in R. Parasuraman and D.R. Davies, eds. Varieties of Attention. Academic Press, New York.

Moray, N., and M. Fitter. 1973. A theory and measurement of attention. In S. Kornblum, ed. Attention and Performance. Vol. 4. Academic Press, New York.

Moray, N., M. Fitter, D. Ostry, D. Favreau, and V. Nagy. 1976. Attention to pure tones. Q. J. Exp. Psychol. 28:271-285.

Norman, D., and D.G. Bobrow. 1975. On data-limited and resource-limited processes. Cognit. Psychol. 7:44-64.

North, R., and D. Gopher. 1976. Measures of attention as predictors of flight performance. Hum. Factors 18:1-15.

Parasuraman, R., and D.R. Davies. 1984. Varieties of Attention. Academic Press, New York.

Reick, A., G. Ogden, and N. Anderson. 1980. An investigation of varying amounts of component task practice and dual task performance. Hum. Factors 22:373-384.

Senders, J. 1964. The human operator as monitor and controller of multi-degree of freedom systems. IEEE Trans. Hum. Factors Electronics 5:2-5.

Stankov, L. 1983. Attention and intelligence. J. Educ. Psychol. 75:471-490.

Sverko, B. 1977. Individual differences in time-sharing performance. Technical Report ARL-77-4/AFOSR-77-4. January 1977. Aviation Research Laboratory, Univ. of Ill.

Sverko, B., Z. Jercil, and Kulenovic. 1983. A contribution to the investigation of time-sharing ability. Ergonomics 12:151-160.

Tellegen, A., and G. Atkinson. 1974. Openness to absorbing and self altering experience ("absorption"): a trait related to hypnotic susceptibility. J. Abnorm. Psychol. 83:268-277.

Weltman, G., S.E. Smith, and G.H. Egstrom. 1971. Perceptual narrowing during simulated pressure chamber exposure. Hum. Factors 13:79-107.

Wickens, C.D. 1984. Engineering Psychology and Human Performance. Charles Merrill, Columbus, Ohio.

Wickens, C.D., and D. Benel. 1981. The development of time-sharing skills. Pp. 253-272 in J.A.S. Kelso and J. Clark, eds. Motor Development. John Wiley & Sons, New York.

Wickens, C.D., S.J. Mountford, and W. Schreider. 1981. Multiple resources, task hemispheric integrity, and individual differences in time-sharing. Hum. Factors 23:211-229.

Williams, L.J. 1985. Tunnel vision induced by foveal load manipulation. Hum. Factors 27:221-227.

THE NEUROPSYCHOLOGY OF MEMORY DYSFUNCTION AND ITS ASSESSMENT[1]

Larry R. Squire

Complaints about memory are very common, particularly among neurological and psychiatric patients. Among neurological patients with organic illness, memory problems are considered the single most frequent complaint (Strub and Black, 1977). Among psychiatric patients, memory problems are commonly reported in association with affective disorders (Sternberg and Jarvik, 1976: Stromgren, 1977) and schizophrenia (Chapman, 1966). In addition, memory functions are of special interest in cases of psychogenic amnesia (Nemiah, 1980) and when questions arise about the possible side effects of treatments such as psychotropic drugs and electroconvulsive therapy (ECT).

Evaluation of memory disorders with quantitative methods makes it possible to identify the various disorders that can occur, to understand the similarities and differences between them, and to follow their course reliably in individual patients. Several reviews have appeared in recent years that consider the subject of memory disorders and memory testing (Barbizet, 1970; Butters and Cermak, 1980; Erickson and Scott, 1977; Russell, 1981; Talland, 1965). This chapter will identify components of memory testing and develop a rationale for the neuropsychological evaluation of memory functions.

Of the higher cortical functions (e.g., perception, language, memory, and action), memory is perhaps the most studied and the best understood. Although our understanding is still very primitive, there is optimism among neuroscientists that memory will be among the first of the higher brain functions to be explained--both at the cellular level in terms of cellular events and synaptic change and at the systems (or neuropsychological) level in terms of brain regions and brain organization. Looking at the topic of memory disorders from the perspective of clinical neuropsychology, a good deal of relevant information can be gained from cognitive psychology and the neurosciences about normal memory and how it is organized in the brain. Readers may wish to consult any of several recent reviews and monographs that address these issues (Cermak, 1982; Kandel, 1977; Klatsky, 1975; Lynch et al., 1984; Mishkin, 1982; Norman, 1982; Squire, 1982a; Squire and Cohen, 1984; Squire and Davis, 1981; Thompson et al., 1983; Weiskrantz, 1982).

[1]From Neuropsychological Assessment of Neuropsychiatric Disorders, edited by Igor Grant and Kenneth M. Adams. Copyright 1985 by Oxford University Press, Inc. Reprinted by permission.

Memory dysfunction most often occurs in association with other disorders of intellectual function, as in depression and dementia. If attention is impaired, acquisition of memory will be deficient. If language is impaired, it may be difficult to remember words or to call them to mind with the usual facility. In this sense, it might seem unfruitful or even misleading to speak about memory and memory disorders in isolation from other cognitive functions. However, there is an important reason for considering memory as a distinct neurological function. Disorders of memory can sometimes occur as a relatively pure entity. Although such disorders are rather uncommon, study of them can aid in recognizing and understanding memory disorders when they occur in a web of other disorders.

FUNCTIONAL AMNESIA

Disorders of memory have many causes and can take many forms. A first step in classification is to distinguish disorders of functional or psychogenic origin from disorders that result from direct perturbation of brain function, i.e., from neurological injury or disease. Functional amnesias may be the best known, having been the subject of frequent treatment in films and literature. They are not nearly as common, however, as the so-called organic amnesias, and with proper testing these two general types of memory impairment are easy to distinguish.

A recently published case of functional amnesia is particularly instructive (Schacter et al., 1982). A 21-year-old man approached a policeman in downtown Toronto complaining of back pain. When brought to the hospital, it was found that he did not know his own name or anything about his past, except for an isolated period from about 1 to 2 years previously, when he had worked for a courier service. He also gave a nickname, Lumberjack (this was the pseudonym used in the published report to protect the patient's identity), but did not know why he had that name. During 4 days of testing in the hospital, he exhibited a good capacity for learning new material, and in informal conversation he also demonstrated a continuing awareness of what had occurred since his admission.

On a formal test of remote memory involving photographs of famous persons who had come into the news at different times in the past, he obtained a normal score. In contrast, his performance was quite abnormal when measured by a test of past autobiographical memory. This test asked for recall of past personal events in response to cue words (Crovitz and Schiffman, 1974). For example, he was given the word box and asked to recall a specific episode from his past that involved a box. Only 14% of his recalled episodes were dated by him as having occurred before his hospitalization. By contrast, 91% of the episodes recalled by a control subject came from the corresponding time period. In addition, when the patient's amnesia subsequently cleared, he was given a second form of the test and now dated 92% of his episodes from the period prior to hospitalization. Thus, during his period of amnesia, his past was nearly barren of personal memories; yet his store of general information about past public events was good.

On the fourth day in the hospital, while watching a climactic funeral scene in the televised rendition of "Shogun," his personal past began to return to him. The recent death of a favorite grandfather had provoked his memory loss. Lumberjack had been a nickname given him during his time at the courier service, which had been a particularly happy time of his life. In the end his memory of the past recovered fully, except for a 12-h period preceding his hospital admission. This case illustrates the essential features of functional disorders of memory. There is often loss of personal identity. Anterograde amnesia (loss of new learning capacity) does not usually occur. Retrograde amnesia (loss of memory for events that occurred before the onset of amnesia) is extensive but can be limited only to autobiographical memory.

AMNESIA DUE TO NEUROLOGICAL INJURY OR DISEASE

The organic amnesias have quite a different character. Amnesia can occur for a number of reasons: temporal lobe surgery, chronic alcohol abuse, head injury, hypotensive episode, encephalitis, epilepsy, tumor, and vascular accident. In addition, memory dysfunction is often a prominent and early sign of dementia, including Alzheimer's disease. Patients with memory dysfunction due to neurological injury or disease can appear normal to casual observation, even when the deficit is severe. These patients can have normal intelligence, as measured by conventional IQ tests, and normal ability to hold information in immediate memory, as measured by digit span tests. In conversation, they can show appropriate social skills, insight into their condition, and normal language ability.

The noted patient H.M., for example, was able to detect various kinds of linguistic ambiguity (e.g., "Racing cars can be dangerous"; "Charging tigers should be avoided" [Lackner, 1974]). In an often quoted passage (Milner, 1970, p. 37), H.M. expresses his own experience of his memory disorder.

> Right now, I'm wondering. Have I done or said anything amiss? You see, at this moment everything looks clear to me, but what happened just before? That's what worries me. It's like waking from a dream; I just don't remember.

Patient N.A. (Kaushall *et al.*, 1981; Teuber *et al.*, 1968) also has clear and continuing insight into his condition. Failure of insight, or outright denial of illness, seems to be related to two factors. First, when memory problems have a gradual onset, as with dementia or a slowly growing tumor, memory loss is often underappreciated. Second, when memory problems occur as part of a confused and disoriented state, as in acute conditions like Wernicke-Korsakoff syndrome, patients can be unaware of their memory impairment.

Brain Regions Affected in Amnesia

It has been known for almost 100 years that amnesia depends on disruption of normal function in one of two brain regions, the medial surface of the temporal lobes and the diencephalic midline of the brain. For the medial temporal region, it has usually been supposed that amnesia depends on bilateral damage to the hippocampal formation, but the evidence on this point is by no means decisive. The idea that amnesia depends on hippocampal damage is based largely on a famous series of epileptic and psychotic patients, some of whom had sustained bilateral resection of the amygdala and did not exhibit amnesia, and two others (including the well-known patient H.M.) who sustained more radical excisions of both amygdala and hippocampus and developed profound amnesia (Scoville and Milner, 1957). Yet it is also possible, on the same evidence, to suppose that amnesia occurs only when there is damage to both amygdala and hippocampus, and the findings of one study in monkeys are consistent with this idea (Mishkin, 1978). Further studies of memory and amnesia in the monkey are needed to settle this issue (cf. Mahut, in press; Mahut et al., 1981; Squire and Zola-Morgan, 1983; Zola-Morgan, 1984; Zola-Morgan et al., 1982).

For diencephalic damage, a large body of neuropathological data has associated amnesia with damage to the mammillary bodies and to the dorsomedial thalamic nucleus (Brierley, 1977; Mair et al., 1979; Victor et al., 1971). Damage to these structures is prominent in the neuropathology of Korsakoff's alcoholic syndrome, perhaps the best-studied example of amnesia, but there is disagreement concerning whether separate damage to one structure or the other can cause amnesia and which structure deserves the greater emphasis. The dorsomedial nucleus of the thalamus appeared to be a site of damage in patient N.A. (Squire and Moore, 1979), a well-studied single case. N.A. became amnesic in 1960, primarily for verbal material, as the result of an accidental stab wound to the brain with a miniature fencing foil (Kaushall et al., 1981; Teuber et al., 1968). In addition, individual cases of amnesia with vascular damage to the region of the dorsomedial nucleus due to occlusion of the paramedian artery have been reported (Michel et al., 1982; Mills and Swanson, 1978; Speedie and Heilman, 1982). Nevertheless, the relative importance of the dorsomedial nucleus and the mammillary bodies in amnesia is not yet clear. Studies of monkeys with circumscribed surgical lesions of these structures, alone and in combination, will be needed to address these issues decisively.

Similarities and Differences Among the Amnesias: A First Approximation

Korsakoff patients are sometimes taken as typical of amnesic patients in general, because this syndrome is the most thoroughly studied example of amnesic disorder. But the Korsakoff syndrome differs in several ways from other examples of amnesia. For example, Talland (1965, p. 108) wrote that the memory disorder in patients with Korsakoff's syndrome "does not present simply a derangement in memory," and Zangwill (1977, p. 113) contended that "other and more extensive psychological dysfunction must co-exist with amnesia for the classic picture of Korsakoff's syndrome to emerge."

Some of these other deficits may derive from frontal lobe dysfunction. Neuropsychological signs of frontal lobe impairment are commonly present in these patients and can influence neuropsychological test performance (Moscovitch, 1982; Squire, 1982b).

The confabulation sometimes exhibited by Korsakoff patients is by no means a common feature of either amnesia in general or the Korsakoff syndrome itself. Confabulation refers to misstatements of fact, often amusing, bizarre, and self-contradictory, which occur as short-latency responses to questions (Mercer *et al.*, 1977). Confabulation is seen most often in the acute phase of the disease, particularly in association with denial of illness, and is seen much less often in the chronic phase.

Neuropsychological examination of patients with memory dysfunction must take into account this idea that etiologically distinct forms of amnesia may present particular deficits superimposed on the amnesia. In addition, the nature of the disorder itself may depend to some extent on the locus of the effective lesions. Recent evidence suggests that diencephalic and bitemporal amnesia may differ from each other in the rate of forgetting (Huppert and Piercy, 1979; Squire, 1981), though the evidence on this point is not yet decisive (Zola-Morgan and Squire, 1985).

NEUROPSYCHOLOGICAL ASSESSMENT OF MEMORY DYSFUNCTION

Time-Line Measurements: Anterograde and Retrograde Amnesia

Because many patients have insight into their conditions, one useful way to obtain information about memory loss is to construct a time-line of the deficit. This technique was used by Barbizet (1970) to identify portions of past time periods that were affected. The method can reveal in an approximate way the duration of both anterograde and retrograde amnesia and can show how the deficit changes with the passage of time. For Barbizet's patient, who had suffered a severe closed head injury, anterograde amnesia remained fixed at about 3.5 months even after memory capacities had largely recovered. This occurred because the anterograde amnesia reflected a time when memories could not be formed in the normal way. Accordingly, even though the capacity to form new memories eventually recovered, memories did not return for events that had occurred during the period of anterograde amnesia. Retrograde amnesia was initially severe and extensive, but gradually shrank to 2 weeks; the oldest memories recovered first.

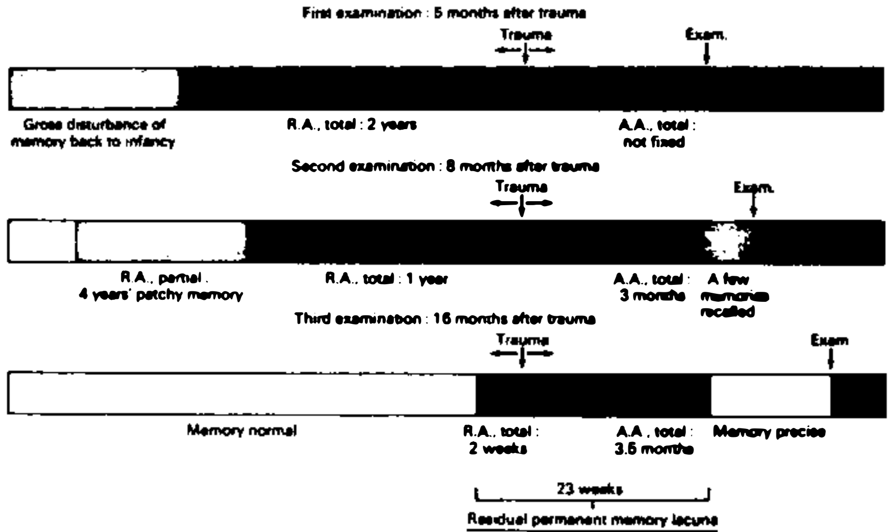


FIGURE 1. Illustration of the time periods that were difficult to remember at three different intervals after a severe head injury. (From Barbizet, 1970.) R.A., retrograde amnesia, A.A., anterograde amnesia.

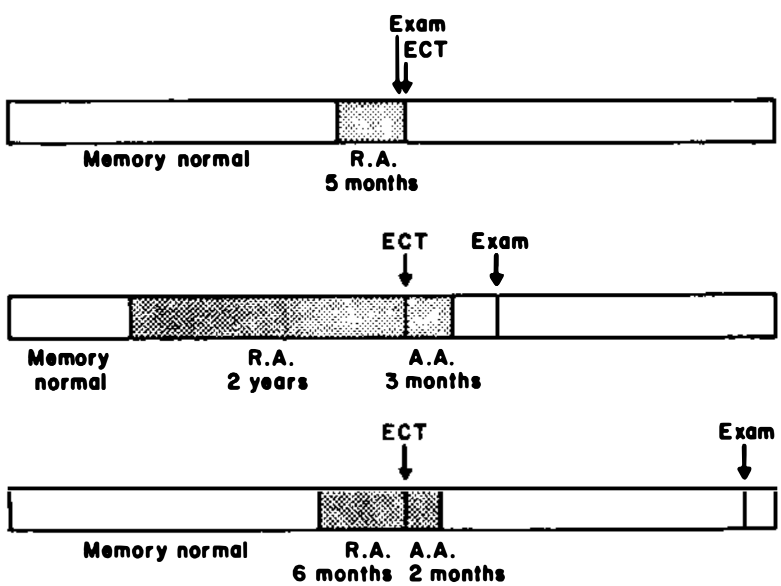


FIGURE 2. Estimates of time periods that were difficult to remember obtained before (top), 7 months after (middle), and 3 years after (bottom) bilateral ECT ($n = 31$). Shaded areas represent the median time period perceived as affected both from the period before ECT (retrograde amnesia [R.A.]) and after ECT (anterograde amnesia [A.A.]). Since the first time estimate was obtained just prior to ECT (top), the 5 months perceived as affected at that time presumably reflected memory problems associated with depressive illness. (From Squire and Slater, 1983.)

A similar relationship between anterograde and retrograde amnesia holds for psychiatric patients undergoing a prescribed course of bilateral ECT. Thirty-one patients were interviewed before and 7 months and 3 years after a course of ECT. Before treatment, patients on average reported having difficulty remembering the 5 months prior to ECT, presumably because of the depressive illness that had brought them to this treatment. Seven months later, patients reported difficulty remembering events that occurred during the 3 months after treatment and for 2 years before it. Three years later, retrograde amnesia had shrunk to about the pre-ECT level, and anterograde amnesia remained fixed at 2 months (Squire and Slater, 1983).

The largest study of disorders of this type, which occur without penetration of brain tissue, comes from the classic work of Russell and Nathan (1946) on traumatic amnesia. Their study of 1,031 consecutive cases makes several points about the relationship between anterograde and retrograde amnesia. First, based on 972 of the cases for which information was available about memory loss, retrograde amnesia was typically brief, less than 30 min in 90% of the cases. Second, the longer the anterograde amnesia, the longer the retrograde amnesia. Third, retrograde amnesia was more severe in closed (concussive) head injury than in gunshot wounds or other cases of penetrating brain injury. During the period of amnesia, neither the anterograde nor

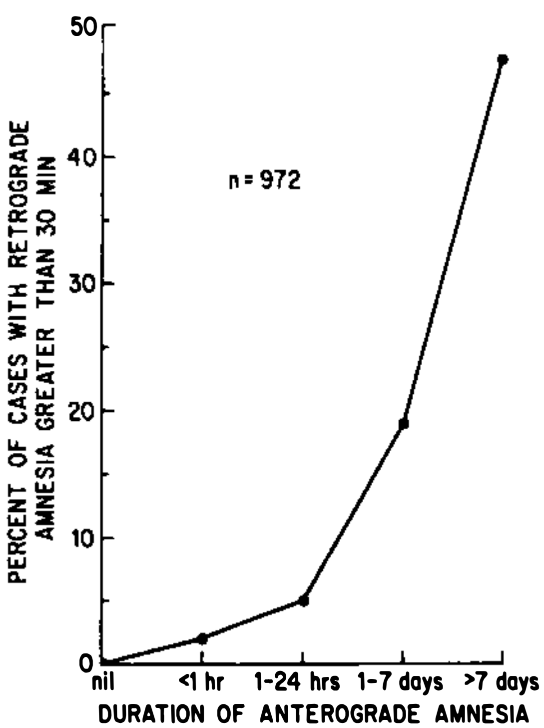


FIGURE 3. Relationship between anterograde and retrograde amnesia in 972 head injury patients. (Reconstructed from Russell and Nathan, 1946.)

retrograde component of memory loss could be influenced significantly by hypnosis or barbiturates.

Quantitative Measurement: Standardized Tests and Batteries

Although the time-line technique can be of considerable help in diagnosis and in obtaining a rough idea of the temporal dimensions of the memory deficit, it is of less help in judging the severity of the deficit because it provides no quantitative information. Formal tests are needed to obtain reliable quantitative data. Because of the rich tradition of basic research that has been directed at problems of memory during the past few decades, many good tests are available.

One of the advantages of formal neuropsychological testing is that it permits objective comparison of the scores of a given patient with a known group average. This protects us from the natural tendency to rationalize, minimize, or forget clinical observations when they do not conform to expectation. These tendencies are, of course, not confined to clinical observers. In his autobiography, Charles Darwin (Darwin, 1892, p. 42) wrote:

> I had also, during many years, followed a golden rule, namely, that whenever a published fact, a new observation or thought came across me, which was opposed to my general results, to make a memorandum of it without fail and at once; for I had found by experience that such facts and thoughts were far more likely to escape from the memory than favorable ones.

It is also true that we have been so influenced by literature and film to look for psychological explanations of behavior that in the absence of formal neuropsychological testing, we often tend to develop psychological explanations for organic memory disorders. A few years ago, I received a letter from a woman telling me of her son's memory problems, which had resulted from a traumatic head injury sustained in an auto accident. Feeling that his memory problems seemed greater at some times than at others, she had said to him, "You only remember what you want to," whereupon he had replied, "Don't we all?" Her son was probably right. Persons with memory dysfunction, so long as it is not so severe as to be absolute, have the same tendencies that we all have to remember more reliably things that seem important than things that seem trivial and in general to exercise the same denial, repression, and selection when processing new information. In amnesia these normal selective factors seem to be superimposed on an overall reduced retentive capacity.

The best-known and most often used neuropsychological batteries provide for only limited testing of memory functions (Golden et al., 1980; Reitan and Davison, 1974). The well-known Wechsler Memory Scale (WMS) (Wechsler, 1945) is inadequate to the task of detecting circumscribed memory problems, especially when they are mild. At 24 h after the fifth treatment, patients receiving bilateral ECT were reliably impaired only on the paired-associate subtest of the WMS (Small, 1974), even though tests of memory functions showed clear defects at the same time. In addition, because the WMS contains questions about contemporary facts (e.g., Who is president? governor of your state? mayor of your city?), the WMS will yield lower scores for patients who have been amnesic for a long time than it will for patients who have become amnesic only recently. Thus, even if two patients have equivalent difficulty in acquiring new information, the patient who has been amnesic for several years may not know the answers to these questions, but the patient who has become amnesic recently may know them. Russell (1975) has greatly improved the usefulness of the WMS by revising it to include delayed-recall measures. Delayed-recall tasks are critical to bringing out an amnesic deficit, and these tasks will be discussed more fully below.

Another standardized battery, the NYU Memory Test (Randt et al., 1980), was constructed specifically to detect memory dysfunction of the organic type. It contains delayed-recall measures, a test of paired-associate learning (see next section), and a sensitive test of incidental learning. This group of tests has been used extensively in the evaluation of memory in aging subjects (Osbourne et al., 1982). Because five alternative forms are available, it can be used in longitudinal studies.

Those who prefer to use one of the popular, general-purpose neuropsychological batteries could include an explicit test of memory dysfunction by adding a simple procedure at the conclusion of testing. Give the patient a test form that includes simple phrases describing several of the tests that he has taken together with phrases describing bogus tests that were not taken. Ask the patient to indicate for each description whether he recognizes having taken such a test. The sensitivity of this recognition measure could be increased by preceding it with a recall test, i.e., asking patients to describe by unaided recall the tests that they have taken.

Paired-Associate Learning

Among tests specialized for the detection and quantification of memory dysfunction, perhaps the most sensitive is paired-associate learning. Figure 4 shows the performance of three kinds of patients with memory problems who attempted to learn 10 noun-noun word pairs (e.g., army-table, door-sky). In this case, three consecutive presentations of the word pairs were given, and after each presentation subjects were asked to try to produce the second word of the pair upon hearing the first. Patient N.A., patients with alcoholic Korsakoff syndrome, and psychiatric patients tested 1 to 2 h after their fifth bilateral ECT all performed poorly on this test, obtaining an average score of less than 3 correct responses out of 10 on the third learning trial. Control subjects for each kind of patient performed considerably better. On this same test, patient H.M. was unable to produce any correct responses after three trials, even when he was instructed in the use of imagery techniques for associating the words in each pair (Jones, 1974).

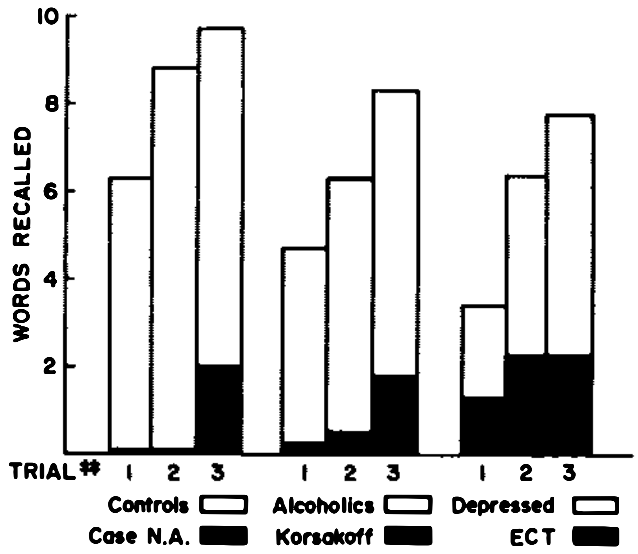


FIGURE 4. Impairment in paired-associate learning by three kinds of amnesic patients and separate control groups: patient N.A., patients with alcoholic Korsakoff syndrome, and patients receiving bilateral ECT. Subjects were presented with 10 noun-noun pairs on each of three trials. After each presentation they tried to recall the second word of each pair after seeing the first word.

Free Recall and Reminding Techniques

Another useful single test, which can generate a good deal of information for individual subjects, involves assessing the free recall of words by a method of selective or restricted reminding (Buschke and Fuld, 1974). In both cases, 10 words from a single category, e.g., animal names, are presented, and subjects are asked to recall as many of them as possible after each of several presentations. In selective reminding, each successive presentation of the list includes only those words not recalled on the preceding trial. Trials continue until the subject succeeds in recalling all the words in a single trial. In restricted reminding, each successive presentation of the list includes only those words that have not yet been recalled on any trial. Trials continue until the subject has succeeded in recalling each word once. These techniques can yield useful data on single subjects about the ability to acquire new information and the consistency of performance.

Exceeding Immediate Memory Capacity

The critical feature of both paired-associate learning tasks and selective or restricted reminding tasks that accounts for their sensitivity to amnesia is that the information presented to the patient exceeds the immediate memory capacity. Even severely amnesic patients can have normal digit spans and a normal ability to report back the relatively small amount of information that can be maintained in conscious awareness. William James (1890, pp. 646-648) termed this capacity primary memory.

> An object in primary memory . . . was never lost; its data was never cut off in consciousness from that of the immediately present moment. In fact it comes to us as belonging to the rearward portion of the present pace of time, and not to the genuine past. . . . Secondary memory, as it might be styled, is the knowledge of a former state of mind after it had already once dropped from consciousness. . . . It is brought back, recalled, fished up, so to speak, from a reservoir in which, with countless other objects, it lay buried and lost from view.

This concept of primary memory remains quite useful in understanding the nature of the memory impairment in amnesia. This point has been illustrated in a formal way by an interesting study of five amnesic patients, including the noted surgical patient H.M. (Drachman and Arbit, 1966). Patients and control subjects were given digit strings of increasing length to repeat back until an error occurred. An error was defined as failing three times in succession on three different digit strings of the same length. At that point a different string of digits of the same length was given repeatedly until it was reproduced correctly or until 25 repetitions of the same digit string had been given. Each time a correct response was given, a new string of digits was presented that was one digit longer than the preceding string. With this procedure, normal subjects were able to increase their digit span to at least 20 digits. Amnesic patients, however, had great difficulty once their digit span capacity had been reached, i.e., at that digit string length when their first error occurred. H.M. was unable, even after 25 repetitions of the same digit string, to increase his digit span by one digit beyond his premorbid level of six digits.

Thus, for amnesic patients, performance on tests involving immediate recall depends on whether the amount of information to be remembered exceeds a finite processing capacity, termed primary memory or immediate memory capacity. It is also true that memory performance will be poor even when the amount of information to be remembered is within the limits of that capacity, so long as a delay filled with distraction is interposed between learning and retention testing in order to prevent active rehearsal. If the delay is very long, e.g., an hour or more, the natural distraction of ongoing activity is sufficient to prevent rehearsal and to reveal a deficit if one is present. If the delay is short, seconds or minutes, a formal distraction procedure is needed to prevent rehearsal. The following observation of H.M. (Milner, 1970, p. 37) makes this point.

> Forgetting occurred the instant his focus of attention shifted, but in the absence of distraction his capacity for sustained attention was remarkable. Thus he was able to retain the number 584 for at least 15 minutes, by continuously working out mnemonic schemes. When asked how he had been able to retain the number for so long, he replied: "It's easy. You just remember 8. You see, 5, 8, and 4, add to 17. You remember 8, subtract it from 17 and it leaves 9. Divide 9 in half and you get 5 and 4, and there you are: 584. Easy."

The Importance of Delayed-Recall Measures

With these considerations in mind, it is easy to understand that the hallmark of the organic amnesias is considered to be impaired performance on tests of delayed recall (with interpolated distraction). These tests form the cornerstone of any thorough neuropsychological assessment of memory functions. Tests of delayed recall are simply formal versions of the well-known bedside examination used by neurologists. Typically the names of three objects are presented to the patient, with the instruction to repeat them in order to demonstrate comprehension and attention. After a delay of several minutes, which is filled with the continuing mental status examination, the patient is asked to recall the words. This informal method can be expanded to yield even more information by first asking patients at the time of recall how many words they had been asked to remember, by cueing them with synonyms or rhymes for any words that could not be produced in unaided recall, and finally by offering several words and asking the patients to pick out the ones that had been presented earlier. Failure to recognize words as having been previously presented is a more reliable sign of memory disorder than failure to recall, and recognition failure denotes a more severe disorder as well. The use of recall and recognition tests together makes it possible to detect more subtle and earlier signs of impairment.

Since in many medical settings patients are tested repeatedly by different physicians and students, it would be useful if the number of words given for memorization were sometimes four, not always three, and of course the words themselves must be different in each test session. Patients with memory problems, like the rest of us, may forget unique events but still retain information about events that are repeated. This sometimes enables them to defeat the purpose of mental status examinations by rehearsing the answers. A colleague tells the story of encountering a neurological patient one day in the hospital outside a room where the patient was about to be presented at rounds.

As my colleague passed by, the patient approached him, asking anxiously, "Say, doc, who's the president of the United States?"

Delayed Recall of Prose Material

There are several formal neuropsychological tests of delayed recall that can be used to good advantage. One useful test of this kind uses connected prose, such as the logical memory subtest of the WMS, and asks patients to repeat the story immediately after hearing it and then again after some delay (Milner, 1958). Figure 5 shows the performance of 15 patients who had been prescribed bilateral ECT on a test of this type (not the WMS) with a delay of 24 h. Both immediate and delayed recall were assessed before ECT and then again, with an alternative form, 6 to 10 h after the fifth treatment of the series. ECT had no effect on immediate recall of the prose passage, and by this measure memory functions might have seemed very good. Indeed, by 6 to 10 h after the fifth treatment, patients can score normally on tests of verbal IQ, show no signs of confusion or disorientation, and carry on a conversation in a normal and appropriate way. It is not uncommon at this time to hear comments from hospital staff not familiar with ECT that the patient's memory is all right or that the patient can remember what he wants to. Yet delayed-recall tests show that memory functions are not normal at this time. Whereas patients tested before their prescribed series of ECT showed on average considerable retention 24 h after hearing the prose passage, only 2 of 15 could recall any part of the prose passage after ECT, and some could not remember having heard any passage or having seen the experimenter previously. Figure 5 also shows the performance of another group of subjects tested 6 to 9 months

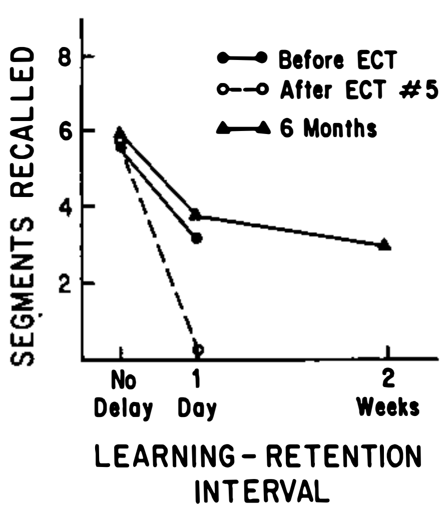


FIGURE 5. Delayed recall of a short prose passage by 15 patients prescribed bilateral ECT. Testing occurred before treatment, 6 to 10 h after the fifth treatment, and for a different group of 16 patients 6 to 9 months after bilateral treatment.

after the completion of treatment. The ability to learn and retain prose material has considerably recovered by this time after treatment.

This pattern of deficit, impaired recall at a delayed test but not at an immediate test, helps to distinguish the so-called organic memory disorders from memory disorders caused by depression. Indeed, delayed-recall tests provide sensitive measures of neurological dysfunction and can detect early signs of progressive impairment. By contrast, depression affects immediate memory, presumably because patients are preoccupied or inattentive, but does not affect delayed recall beyond what would be expected from the level of recall achieved at immediate testing (Cronholm and Ottosson, 1961; Sternberg and Jarvik, 1976).

Material-Specific Memory Dysfunction and Its Assessment

In addition to addressing the issue of severity, neuropsychological testing of memory must also address the fact that memory dysfunction can be different depending on whether the neurological injury or disease is bilateral and whether it affects structures in the left or right cerebral hemisphere. The effect on memory follows from the asymmetry of hemispheric function with respect to language: verbal impairment from left-side damage, nonverbal impairment from right-side damage, global impairment from bilateral damage. This point has been best demonstrated in the thorough work of Brenda Milner on temporal lobe function (Milner, 1958, 1971). Patients with resections of the left medial temporal lobe complain, for example, that they cannot remember what they have read, and they do poorly on verbal memory tests. Delayed-recall tests of short prose passages, as just described, are useful in bringing out this deficit. Patients who have sustained right medial temporal lobe resection complain, for example, that they do not remember where they have put things, and they do poorly on tests of memory for faces, spatial relationships, and other things that are not ordinarily encoded in words.

These disorders of memory, which arise from unilateral brain injury or disease, were termed material-specific disorders (Milner, 1968a) to capture the fact that the side of the brain affected determines the kind of material that is difficult to learn and remember. The sensory modality through which material is learned (e.g., auditory, visual, or tactile) is ordinarily not important. In a left medial temporal lobe injury, a short prose passage will be difficult to remember regardless of whether the patient reads the story or hears it read. These material-specific effects have been demonstrated for left and right temporal lobe surgical lesions (Milner, 1971), epileptic foci of the left or right temporal lobe (Delaney, 1980), left or right unilateral ECT (where the two electrodes are applied to the same side of the head, in contrast to bilateral ECT where one electrode is applied to each temple) (Squire, 1982c), unilateral diencephalic lesions (Michel *et al*., 1982; Speedie and Heilman, 1982; Squire and Slater, 1978; Teuber *et al*., 1968), and unilateral diencephalic brain stimulation (Ojemann, 1971).

Ross (1980a,b) has described five patients with medial temporal lobe dysfunction who provided evidence for the existence of fractional, modality-specific memory disorders. The deficit affected the tactile or visual modality and in three of the five cases did not affect memory functions in other modalities. In addition, three patients with acute unilateral lesions of the medial temporal region were reported to have "tactile recent memory loss" for stimuli applied to the hand opposite the lesion, but not for stimuli applied to the ipsilateral hand. These observations raise the possibility that left and right medial temporal lobe lesions need not always cause equally severe deficits in all modalities, as has been traditionally believed. Moreover, even deficits affecting all modalities might sometimes be more severe when information is presented to the visual field or body half contralateral to the lesion. Additional systematic comparisons of modality (e.g., visual versus auditory versus tactile), side of presentation (left versus right), and kind of material (verbal versus nonverbal) are needed before the relative importance of all three factors can be appreciated.

Delayed Recall of Nonverbal Material

Of the tests that have been used to assess delayed recall of nonverbal material, perhaps the best known is the Rey-Osterrieth figure (Figure 6). The subject is asked to copy the figure and then after a delay is asked to reconstruct it from memory without forewarning. An immediate test of reconstruction can be used as well. Brenda Milner and her colleagues have developed an alternative form for this figure and a standardized 36-point scoring system for both figures (Milner and Teuber, 1968). An alternative method for testing delayed recall of nonverbal memory, suggested by Russell (1975), is to use the figures from the WMS. The ability to reconstruct these figures is tested in the normal way, immediately after presentation, and again 30 min later. When patients have constructive deficits or other problems that make the use of memory tests involving drawing unsatisfactory, other tests

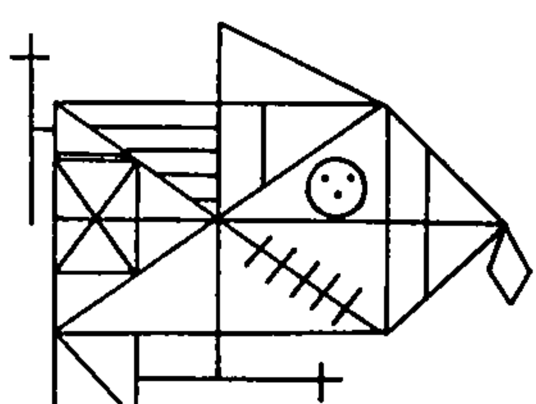


FIGURE 6. The Rey-Osterrieth figure, used to assess delayed recall of nonverbal material. (From Osterrieth, 1944; also see Milner and Teuber, 1968.)

can be used, e.g., recognition of previously presented faces after a delay (Milner, 1968b), recollection of the position of a dot on an 8-in. horizontal line (Milner, 1974), or recognition by touch of wire shapes or other objects that are not easily remembered by verbal labels (Milner and Teuber, 1968).

Skill Learning

Although the deficit exhibited by amnesic patients with bilateral brain damage has often been described as global, as if it affects the learning and memory of any or all material, it is now clear that the memory impairment in fact is considerably selective. Perceptuo-motor skills, such as the eye-hand coordination skills involved in certain tracking tasks and in mirror drawing, can be acquired by amnesic patients (Corkin, 1968; Milner, 1962). Acquisition of such tasks can occur at a normal rate over a period of days, although patients may deny having worked at the task before. More recently this observation has been extended to cognitive skills, such as mirror reading and the solution to certain puzzles (Cohen and Corkin, 1981; Cohen and Squire, 1980). These observations have suggested a distinction between information based on skills or procedures (procedural knowledge), which is intact in amnesia, and information based on facts or data (declarative knowledge), which is impaired.

The capacity for acquiring skills apparently does not require integrity of the medial temporal or diencephalic brain regions damaged in amnesia, whereas the capacity to acquire facts about the world, e.g., the ability to recognize the words that have been read in a mirror-reading task, depends on the integrity of these brain regions. These ideas have been developed in more detail elsewhere, with the suggestion that the nervous system honors the distinction between two different kinds of memory systems (Cohen, 1981, 1984; Squire, 1982a; Squire and Cohen, 1984).

This idea is reminiscent of other classic distinctions about the representation of knowledge (e.g., knowing how versus knowing that (Ryle, 1949) and procedural versus declarative knowledge (Winograd, 1975). These distinctions seem to capture the difference between what amnesic patients can and cannot do, as we presently understand them. Other terminologies have also been suggested in order to embody the finding that amnesic patients (or experimental animals with presumably similar lesions) can successfully accomplish some kinds of learning and memory, but not other kinds (conscious recollection versus skills [Moscovitch, 1982], memory versus habits [Mishkin, 1982], mediational versus semantic memory [Warrington and Weiskrentz, 1982], episodic versus semantic memory [Kinsbourne and Wood, 1975], taxon versus locale [O'Keefe and Nadel, 1978], reference versus working memory [Olton *et al.*, 1979], vertical versus horizontal associative memory [Wickelgren, 1979]). Whatever language is used, there now seems good reason to suppose that there is more than one kind of memory. The notion of two (or more) kinds of memory or memory systems has interesting implications for a variety of issues, such as infantile amnesia, phylogeny of memory, and psychoanalytic theory (Nadel and Zola-Morgan, 1984; Schacter and Moscovitch, 1984).

Recent work with amnesic patients suggests that the domain of learning and memory that is spared may include not only skills but also classical conditioning (Weiskrantz and Warrington, 1979) and what has been termed priming (Graf et al., 1984) or perceptual fluency (Jacoby, 1982). It has been suggested (Cohen, 1984; Squire and Cohen, 1984) that what is spared in amnesia depends on changes in already existing knowledge structures. By contrast, the domain of learning and memory that is affected in amnesia appears to provide additional information, i.e., directly accessible information, about the time and place of occurrence of individual events, facts about the world that derive from these events, and the awareness and sense of familiarity that a particular event has previously occurred.

The sparing of skill learning should be demonstrable in amnesic patients to the extent that amnesia is relatively free from other cognitive deficits. The presence of frontal lobe signs, a common feature of Korsakoff syndrome, or the presence of dementing illness may interfere with the capacity for skill learning. It remains to be seen whether the preserved capacity for learning and memory in amnesic patients has any implications for rehabilitation.

Assessment of Remote Memory

One of the striking features of amnesia is that it commonly involves some loss of information for events that occurred before onset. Before considering the retrograde aspect of memory dysfunction, it is important to note that in memory disorders having a gradual onset the distinction between anterograde and retrograde amnesia is necessarily blurred. It is not always possible to determine whether information has been lost because it was not acquired in the first place or because it was first acquired and then later lost as the result of the onset of amnesia.

Several techniques have been developed to assess memory for remote events in a formal quantitative way. When amnesia has a very recent or known time of onset, test questions that access retrograde amnesia can be used. Of the tests available for clinical neuropsychological testing, most ask for information about public events that occurred at specific times in the past. This ensures that the information is verifiable and accessible to all subjects. The first remote memory tests to explore the nature of memory disorders in amnesic patients consisted of multiple-choice questions about persons or events that had been in the news in Great Britain or faces of famous persons who had come into prominence in Great Britain at different times in the past (Sanders and Warrington, 1971). By presenting a number of questions for each of several past decades, it was possible to obtain a sampling of an individual's knowledge of past events. Subsequently, similar tests based on public events or famous faces (Fig. 7) were developed for use in the United States (Albert et al., 1979; Seltzer and Benson, 1975; Squire, 1974). In addition to multiple-choice techniques, these tests can all be given in a free-recall format prior to recognition testing, and many of them can be given in a detailed-recall format, in which subjects are asked to tell all they can about a previous event

FIGURE 7. Items from the Famous Faces Test of the Boston Retrograde Amnesia Battery (Albert et al., 1979). From left to right, top row: John L. Lewis (1930s), Fulton Sheen (1940s), Joe McCarthy (1950s); bottom row: Jimmy Hoffa (1960s), H.R. Haldeman (1970s).

(Cohen and Squire, 1980; Squire and Cohen, 1982). In this case, the responses are recorded, transcribed, and then scored for number of details produced.

Another kind of remote memory test that has found useful application in the study of memory disorders is a test of former one-season television programs (Squire and Fox, 1980; Squire and Slater, 1975). This test was designed to overcome one important limitation of all other available remote memory tests, i.e., the difficulty in comparing scores from different past time periods.

To make valid comparisons across time periods, the items selected must satisfy the criterion of equivalence, i.e., they must sample past time periods in an equivalent way so that the events from different time periods are likely to have been learned to about the same extent and then forgotten at similar rates. The television test appears to satisfy this criterion, as demonstrated by studies with updated versions of the test over a 7-year period (Squire and Fox, 1980).

Remote memory tests cannot be assumed to satisfy the criterion of equivalence just because normal subjects obtain the same score (e.g., 75% correct) across all time periods sampled by the test. As discussed in detail elsewhere (Squire and Cohen, 1982), it is possible that the events selected from more remote time periods were initially more salient and more widely known than the events selected from more recent time periods. In addition, the events from more remote periods could have been forgotten more slowly. Accordingly, findings from patients taking these tests will necessarily be ambiguous, to the extent that the point of testing is to compare performance across time periods.

These issues notwithstanding, all the remote memory tests have some useful application to the quantitative examination of memory disorders. As with any scientific tool, its appropriateness depends on the particular question being asked. Thus, the television test, although advantageous in some respects, is limited by the relatively short time span that it can reliably cover (about 20 years) and by the fact that it yields variable results when used clinically to explore the memory capacity of a single patient. Indeed, many of the tests available for the assessment of memory, including the television test, were originally designed for the investigation of specific research questions in groups of patients, and they are not always as useful as one might wish when trying to understand the nature of memory impairment in an individual patient. For this reason it is advisable to use several different tests when studying a single patient.

An important general point about remote memory tests for public events is the issue of test sensitivity. Unfortunately, multiple-choice tests and recall tests that ask for a single word or phrase do not appear to be sufficiently sensitive to rule out definitively the presence of remote-memory impairment. Patient N.A., who had a circumscribed diencephalic lesion involving the region of the dorsomedial thalamic nucleus, performed six such tests normally but was impaired on two tests that assessed his detailed recall for past events (Cohen and Squire, 1980; Zola-Morgan et al., 1983). Until detailed recall tests are given to other patients with dorsomedial thalamic damage, the status of retrograde amnesia in this group remains uncertain. Despite these uncertainties about patients with focal diencephalic lesions, remote memory tests have been useful in characterizing and differentiating clinical groups (Butters and Albert, 1982; Cohen and Squire, 1980).

All the remote memory tests just described share the advantage of being based on verifiable and publicly accessible information. A final remote-memory test should be mentioned here that does not have this feature, but can nevertheless be of considerable value in the neuropsychological assessment of memory disorders. This test derives from early quantitative studies by Galton (1879) and was recently modified and applied to the study of memory by Crovitz and Schiffman (1974). The test is designed to obtain autobiographical remote memories about specific past episodes of a patient's life. Patients are given 10 standard cue words (e.g., window, tree, ticket, bird) and are asked in each case to recall a specific memory from the past that involves this word. Various scoring procedures can be used, a 0 or 1 method that can be done during the testing session, or a 0 to 3 method, where the responses are recorded, transcribed, and then scored by some predetermined system of partial credits. After recalling a memory for a given cue word, subjects are asked to date the memory as best they can. This test can provide useful information about the quality and quantity of recalled information, as well as vital information about the time periods from which recall is possible. Even amnesic patients who obtain normal scores for recall may draw their memories from different time periods than normal subjects. This was true in the case of the patient with functional amnesia described above, who atypically drew most of his memories from the 4 days immediately preceding the onset of his amnesic disturbance (Schacter et al., 1982). Patients with alcoholic Korsakoff syndrome, though they could obtain normal or near normal scores on the recall test, drew their memories from 10 years earlier than their alcoholic control subjects (Zola-Morgan et al., 1983).

The material obtained in an autobiographical test is not easily corroborated. One way to check for outright fabrication is to ask subjects again, some time after the initial test, to date the memories they had recalled (Schacter et al., 1982). This procedure typically shows close agreement between the dates produced in the two sessions.

Self-Ratings of Memory

This section began by pointing out how much can be learned about memory disorders by constructing a time-line of the disorder, based in part on what patients themselves report about their impairment. Quantitative and qualitative information can be obtained with formal tests that take advantage of patients' own sense of their memory problems. Of course, self-reports of memory function can be misleading. For example, in depressed elderly patients memory complaints appeared to be related more to depression than to performance on memory tests (Kahn et al., 1975). Conversely, patients receiving ECT who were clinically improved often denied memory impairment even though impairment could be documented by formal tests (Cronholm and Ottosson, 1963). Nevertheless, relations between memory self-ratings and objective measures of performance can be demonstrated (Baddeley et al., 1982; Zelinski et al., 1980).

Following is an 18-item self-rating scale that has been used with patients receiving ECT (Squire et al., 1979):

My ability to search through my mind and recall names or memories I know are there is:

I think my relatives and acquaintances now judge my memory to be:

My ability to recall things when I really try is:

My ability to hold in my memory things that I have learned is:

If I were asked about it a month from now, my ability to remember facts about this form I am filling out would be:

The tendency for a past memory to be "on the tip of my tongue" but not available to me is:

My ability to recall things that happened a long time ago is:

My ability to remember the names and faces of people I meet is:

My ability to remember what I was doing after I have taken my mind off it for a few minutes is:

My ability now to remember things that have happened more than a year ago is:

My ability to remember what I read and what I watch on television is:

My ability to recall things that happened during my childhood is:

My ability to know when the things I am paying attention to are going to stick in my memory is:

My ability to make sense out of what people explain to me is:

My ability to reach back in my memory and recall what happened a few minutes ago is:

My ability to pay attention to what goes on around me is:

My general alertness to things happening around me is:

My ability to follow what people are saying is:

This scale takes advantage of the fact that memory problems due to depression are different from memory problems due to amnesia (Cronholm and Ottosson, 1961). The former impairs immediate recall but has no special effect on delayed recall. The latter impairs delayed recall, but has little or no effect on immediate recall, especially if the amount of material to be remembered is small and can be held in immediate memory (see above, Delayed Recall of Prose Material).

The scale was designed with the thought that patients might reflect this difference in their own assessments of their memory abilities. Patients given the scale were asked to rate each item from -4 (worse than ever before) through 0 (same as before) to +4 (better than ever before). The items listed in the scale were given to patients in a random order. One week after ECT, self-ratings overall were worse than before ECT, but some items were rated worse than others (Fig. 8). That is, the profiles of self-ratings obtained before and after ECT were different, indicating that ECT changed the patients' own experience of

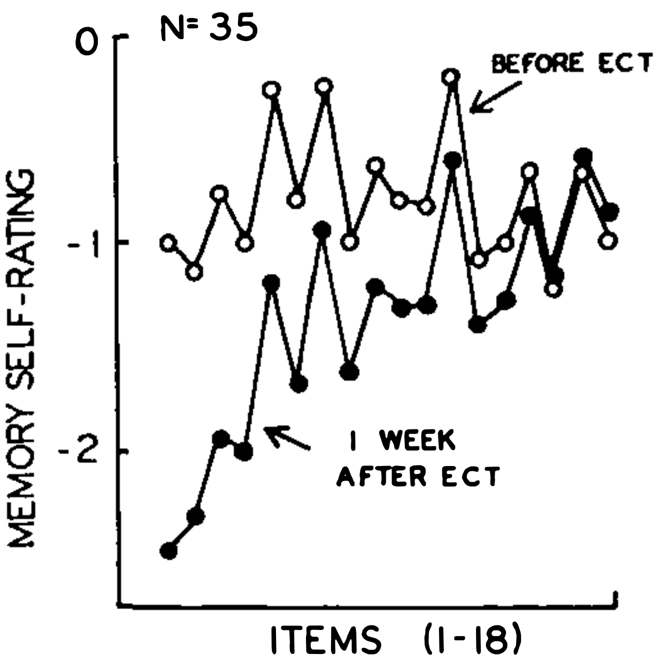


FIGURE 8. Memory self-ratings before and 1 week after a course of bilateral ECT for 35 subjects. Items have been ordered from the one yielding the largest before-after difference (item 1, to left) to the one yielding the smallest difference (item 18, to right). The results show that ECT both increased memory complaints and changed their character. The test items are shown in above list for the self-rating scale. (From Squire _et al._, 1979.)

memory functions. Whereas before ECT, reports of poor memory can probably be attributed to depression, we suppose that after ECT reports of poor memory were influenced largely by the amnesic effect of the treatment.

This finding makes it possible to study patients long after ECT and to ask whether persisting complaints about memory functions are influenced by depression or by amnesia. The results shown in Fig. 9 for the same 35 patients answer this question. First, self-ratings of memory functions improved between 1 week and 7 months after the completion of treatment, corresponding to the findings from objective tests given at these times (Price, 1982; Squire, 1982c). Second, self-ratings retained the form they had 1 week after ECT, when patients were amnesic, and differed from the form before ECT, when patients were depressed. These findings suggest that persisting memory complaints long after ECT reflect an experience of amnesia rather than an experience of depression or some other condition that was present both before and after ECT. An additional follow-up of these patients 3 years later (Squire and Slater, 1983) suggested that the experience of memory impairment is in part veridical, insofar as it relates to the gap in memory that patients experience for events that occurred close to the time of treatment. In addition, it may be in part related to a natural tendency to attribute normal memory problems to the prior ECT experience. So far, formal memory testing 6 months or more after ECT has not provided any evidence for permanent impairment in new learning capacity after a typical course of ECT (i.e., 6 to 12 treatments) (Price, 1982; Squire, 1982a; Weeks *et al*., 1980).

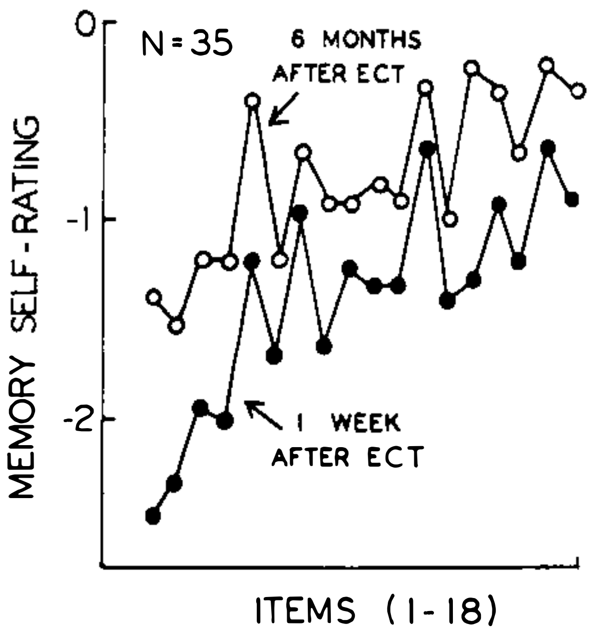


FIGURE 9. Memory self-ratings 1 week and 7 months after bilateral ECT for 35 subjects. Items are arranged as in Fig. 8. The results show that time diminished the severity of memory complaints, but did not alter the pattern observed 1 week after ECT. (From Squire *et al*., 1979.)

Self-rating scales could find useful application in a variety of settings where there is interest in understanding memory complaints and relating them to objective test performance. For example, they might be useful in distinguishing depression in the elderly from the effects of incipient brain disease. In addition, they could be of use in assessing patients with memory disorders who are given pharmacological treatments in an effort to improve their memory functions. Some of these treatments might work in part by improving mood rather than by affecting memory mechanisms directly, and such effects might be teased apart by self-rating scales.

HEAD INJURY AND DEMENTIA

Two instances in which memory dysfunction is commonly encountered by the clinical neuropsychologist are traumatic head injury and dementia, especially Alzheimer's disease. In both circumstances, memory loss occurs in a constellation of other disorders. The specific testing of memory functions, however, can follow the same guidelines that have been developed here for more circumscribed disorders.

The neuropsychology of head injury has been reviewed thoroughly (Newcombe, 1983; Levin et al., 1982). Although memory dysfunction may be the most ubiquitous sign of closed head injury, it commonly occurs together with a reduction in general intellectual capacity, disorders of language and visual perception, apraxia, impairment of attention, and personality change. These multiple disorders are consistent with the variable and widespread pattern of cerebral damage that can result from severe head injury.

Whereas the great majority of patients admitted to hospitals with a head injury eventually make a good recovery, Newcombe (1983) notes that perhaps 1% have persisting signs of impairment. This impairment is now recognized in most cases to have a neurological basis, not a functional basis (e.g., compensation neurosis) as was once believed. The neuropsychologist must address questions about the severity and nature of the deficit and the predicted rate of recovery, as well as questions about possible rehabilitation.

The memory deficits themselves have been carefully reviewed by Schacter and Crovitz (1977). Interestingly, tests of immediate and delayed prose recall, considered above, correlate better with judgments by relatives of a patient's memory abilities than do other formal laboratory tests of memory (Baddeley et al., 1982). In addition to specialized memory tests, assessment might be facilitated by the use of one or more scales that have been developed especially for head-injured patients (Jennett et al., 1981; Teasdale and Jennett, 1974). The Galveston Orientation and Amnesia Test (Levin et al., 1979) can be used repeatedly and yields estimates of retrograde amnesia and of the duration of posttraumatic amnesia. The duration of posttraumatic amnesia is the best available index of severity of injury and is a good predictor of recovery (Jennett, 1976).

In general, techniques that improve memory functions in normal subjects can be expected to improve memory functions in head injury patients and other amnesic patients as well. Accordingly, the techniques most often mentioned in the context of rehabilitation or retraining are elaborate note keeping, imagery, and rehearsal. Imagery may be of special value in working with patients with memory impairment due to left-hemisphere dysfunction (Pattern, 1972). However, for a balanced perspective on the difficulties of making simple generalizations about rehabilitation, see Newcombe (1983).

Memory problems are also commonly encountered by the clinical neuropsychologist in dementia. Dementia is estimated to affect 4 to 5% of the population over 65 years of age. Alzheimer's disease, the most common form of dementia, accounts for at least 50% of the cases (Katzman, 1976; Terry and Davies, 1980). Memory problems are perhaps the most common and often the earliest sign of the disease, but disorders of language (especially anomia) and visuospatial disorders can also occur early. The disease is progressive, eventually involving most or all intellectual functions. It typically develops over a period of 5 to 10 years. At present, the diagnosis of Alzheimer's disease cannot be made definitively without neuropathological data.

Accordingly, a primary question for the neuropsychologist concerns the extent to which the disease can be detected early and differentiated from the benign forgetfulness of normal aging (Kral, 1978), from depression or pseudodementia (McAllister, 1983; Wells, 1979), and from other sometimes treatable forms of dementia.

There is not yet sufficient information to permit Alzheimer's disease to be identified consistently on the basis of neuropsychological reviews, see Fuld, 1978; Kaplan, 1979; Miller, 1981; Corkin et al., 1982). In many respects, the memory loss exhibited by patients diagnosed as having Alzheimer's disease seems to resemble the memory loss exhibited by patients with more circumscribed amnesias. This conclusion is consistent with the finding that neuropathological changes in Alzheimer's disease are prominent in the medial temporal region (Corsellis, 1970). However, two features of memory dysfunction in Alzheimer's disease have been identified that are not observed in the circumscribed amnesias. First, patients with even mild Alzheimer's disease have poor digit span scores, but amnesic patients usually have normal scores on digit span tasks (Corkin, 1982). Second, patients with Alzheimer's disease have been reported to exhibit an extensive deficit in remote memory without any sign of a temporal gradient (Wilson et al., 1981). With additional work, it should be possible to determine, for different stages of the disease, whether the memory deficit more closely resembles diencephalic or bitemporal amnesia and the extent to which the memory loss, perhaps because of superimposed cognitive deficits, has a characteristic and identifiable form.

SELECTION OF RECOMMENDED MEMORY TESTS

The preceding sections considered some of the issues involved in undertaking a clinical neuropsychological evaluation of memory functions. Special techniques are available for assessing memory dysfunction, and these can be used to distinguish functional from organic disorders, to identify material-specific disorders, to assess both anterograde and retrograde amnesia, to assess skill learning, and to explore patients' self-reports of memory functions. Table 1 lists a minimal selection of tests that should serve most clinical needs. More than one test of each type can be used if the findings in any one area are ambiguous. Additional tests can be used as needed, once the clinician becomes familiar with the status of the patient's memory functions.

Neuropsychological testing of memory is most informative when memory tests are supplemented with some additional tests. Additional tests can establish valuable reference points that help in interpreting memory test scores. They can also identify or rule out other kinds of impairment which if present would influence memory test scores. The Wechsler Adult Intelligence Scale is probably the most helpful, as it can define the general test-taking ability of the patient and, in conjunction with the WMS, can give a rough indication of whether memory problems are present. A naming test (Goodglass and Kaplan, 1972) can also be valuable, because patients with anomia will perform poorly on some memory tests (e.g., recall tests of remote memory). Anomia is not always detectable by casual bedside examination. A formal quantitative method that includes many low-frequency object names (protractor, trellis, paddle) will often reveal impairment that is missed by a briefer, more casual exam. Finally, tests of frontal lobe function are often helpful, since frontal lobe dysfunction can influence scores on many memory tests (Moscovitch, 1982; Squire, 1982b).

TABLE 1. Recommended tests for neuropsychological assessment of memory

Test	References
Construction of time-line	Barbizet, 1970; Squire and Slater, 1983
Immediate recall of prose passage	Milner, 1958
Copy of Rey-Osterrieth figure	Milner and Teuber, 1968
Paired-associate learning	Jones, 1974
Delayed recall of prose passage	
Delayed reconstruction of Rey-Osterrieth figure	
Remote memory for famous faces	Albert _et al_., 1979
Cued recall of autobiographical memory	Crovitz and Schiffman, 1974
Self-rating scale	Squire _et al_., 1979

CONCLUSION

A review of the available methods for testing memory and various clinical settings in which these methods can be usefully applied, indicates that much progress has been possible because of basic research. An enormous amount has been learned in recent years about how the brain accomplishes memory storage, and an inventory of memory tests has been one of the fruits of this enterprise.

The neuropsychological study of memory is part of a broader program of research aimed at understanding the biology of memory at all levels of analysis--from cellular and synaptic events to the whole behavior of complex animals like humans. It seems certain that a broad, basic research approach to problems of memory and the brain will continue to inform us about mechanism and organization and at the same time will produce improved methods for neuropsychological assessment of patients with memory dysfunction. Ultimately, perhaps, these research efforts may lead to methods for the treatment of memory dysfunction.

REFERENCES

Albert, M.S., N. Butters, and J. Levin. 1979. Temporal gradients in the retrograde amnesia of patients with alcoholic Korsakoff's disease. Arch. Neurol. 36:211-216.

Baddeley, A., A. Sunderland, and J. Harris. 1982. How well do laboratory-based psychological tests predict patients' performance outside the laboratory? Pp. 141-148 in S. Corkin, K.L. Davis, J.H. Growden, W. Usdin, and R.J. Wurtman, eds. Alzheimer's Disease: A Report of Progress in Research. Raven Press, New York.

Barbizet, J. 1970. Human Memory and Its Pathology. W.H. Freeman & Co., San Francisco. 198 pp.

Brierley, J.B. 1977. Neuropathology of amnesic states. Pp. 199-223 in C.W.M. Whitty and O.L. Zangwell, eds. Amnesia, 2nd Ed. Butterworths, London.

Buschke, H., and P.A. Fuld. 1974. Evaluating storage, retention, and retrieval in disordered memory and learning. Neurology 24:1019-1025.

Butters, N., and M. Albert. 1982. Processes underlying failures to recall remote events. Pp. 257-273 in L. Cermak, ed. Human Memory and Amnesia. Lawrence Erlbaum, Hillsdale, N.J.

Butters, N., and L.S. Cermak. 1980. Alcoholic Korsakoff's Syndrome: An Information Processing Approach to Amnesia. Academic Press, New York. 188 pp.

Cermak, L.S., ed. 1982. Human Memory and Amnesia. Lawrence Erlbaum, Hillsdale, N.J. 388 pp.

Chapman, J. 1966. The early symptoms of schizophrenia. Br. J. Psychiatry 112:225-251.

Cohen, N.J. 1981. Neuropsychological evidence for a distinction between procedural and declarative knowledge in human memory and amnesia. Ph.D. dissertation, University of California, San Diego. Unpublished. [175 pp.]

Cohen, N.J. 1984. Preserved learning capacity in amnesia: evidence for multiple memory systems. Pp. 83-103 in L.R. Squire and N. Butters, eds. The Neuropsychology of Memory. Guilford Press, New York.

Cohen, N.J., and S. Corkin. 1981. The amnesic patient H.M.: learning and retention of a cognitive skill. Soc. Neurosci. Abstr. 7:235.

Cohen, N.J., and L.R. Squire. 1980. Preserved learning and retention of pattern-analyzing skill in amnesia: dissociation of knowing how and knowing that. Science 210:207-209.

Cohen, N.J., and L.R. Squire. 1981. Retrograde amnesia and remote memory impairment. Neuropsychologia 19:337-356.

Corkin, S. 1968. Acquisition of motor skill after bilateral medial temporal lobe excision. Neuropsychologia 6:255-265.

Corkin, S. 1982. Some relationships between global amnesias and the memory impairments in Alzheimer's disease. Pp. 149-164 in S. Corkin, K.L. Davis, J.H. Growdon, E. Usdin, and R.J. Wurtman, eds. Alzheimer's Disease: A Report of Progress in Research. Raven Press, New York.

Corkin, S., K.L. Davis, J.H. Growdon, E. Usdin, and R.J. Wurtman, eds. 1982. Alzheimer's Disease: A Report of Progress in Research. Raven Press, New York. 525 pp.

Corsellis, J.A.N. 1970. The limbic areas in Alzheimer's disease and in other conditions associated with dementia. Pp. 37-50 in G.E.W. Wolsterholme and M. O'Connor, eds. Alzheimer's Disease and Related Conditions. Churchill, London.

Cronholm, B., and J.O. Ottosson. 1961. Memory functions in endogenous depression. Arch. Gen. Psychiatry 5:101-107.

Cronholm, B., and J.O. Ottosson. 1963. The experience of memory function after electroconvulsive therapy. Br. J. Psychiatry 109:251-258.

Crovitz, H.F., and H. Schiffman. 1974. Frequency of episodic memories as a function of their age. Bull. Psychonom. Soc. 4:517-518.

Darwin, F., ed. 1892. Life of Charles Darwin. John Murray, London. 348 pp.

Delaney, R.C., A.J. Rosen, R.H. Mattson, and R.A. Novelly. 1980. Memory function in focal epilepsy: a comparison of non-surgical, unilateral temporal lobe and frontal lobe samples. Cortex 16:103-117.

Drachman, D.A., and J. Arbit. 1966. Memory and the hippocampal complex. Arch. Neurol. 15:52-61.

Erickson, R.C., and M.L. Scott. 1977. Clinical memory testing: a review. Psychol. Bull. 84:1130-1149.

Fuld, P.A. 1978. Psychological testing in the differential diagnosis of the dementias. Pp. 185-193 in R. Katzman, R.D. Terry, and K.L. Bick, eds. Alzheimer's Disease: Senila Dementia and Related Disorders. Raven Press, New York.

Galton, F. 1879. Psychometric experiments. Brain 2:148-162.

Golden, C.J., T.A. Hammeke, and A.D. Purisch. 1980. The Luria-Nebraska Neuropsychological Battery: Manual. Western Psychological Services, Los Angeles.

Goodglass, H., and E. Kaplan. 1972. The Assessment of Aphasia and Related Disorders. Henry Kimpton, London. 80 pp.

Graf, P., L.R. Squire, and G. Mandler. 1984. The information that amnesic patients do not forget. J. Exp. Psychol. Learn. Mem. Cognition 10:164-178.

Huppert, F.A., and M. Piercy. 1979. Normal and abnormal forgetting in organic amnesia: effect of locus of lesion. Cortex 15:385-390.

Jacoby, L.L. 1982. Knowing and remembering: some parallels in the behavior of Korsakoff patients and normals. Pp. 97-122 in L.S. Cermak, ed. Human Memory and Amnesia. Lawrence Erlbaum, Hillsdale, N.J.

James, W. 1890. Principles of Psychology. Vol. 1. Dover, New York. 689 pp.

Jennett, B. 1976. Assessment of severe head injury. J. Neurol. Neurosurg. Psychiatry 39:647-655.

Jennett, B., J. Snoek, M.R. Bond, and N. Brooks. 1981. Disability after severe head injury: observations on the use of the Glasgow Outcome Scale. J. Neurol. Neurosurg. Psychiatry 44:285-293.

Jones, M.K. 1974. Imagery as a mnemonic aid after left temporal lobectomy: contrast between material-specific and generalized memory disorders. Neuropsychologia 12:21-30.

Kahn, R.L., S.H. Zarit, N.M. Hilbert, and G. Niedereke. 1975. Memory complaints and impairment in the aged. Arch. Gen. Psychiatry 32:1560-1573.

Kandel, E. 1977. Neuronal plasticity and the modification of behavior. Pp. 1137-1182 in J.M. Brookhart, V.B. Mountcastle, E.R. Kandel, and S.R. Geiger, eds. Handbook of Physiology. American Physiological Society, Bethesda, Md.

Kaplan, O.J. 1979. Psychological testing of seniles. Pp. 45-78 in O.J. Kaplan, ed. Psychopathology of Aging. Academic Press, New York.

Katzman, R. 1976. The prevalence and malignancy of Alzheimer's disease. Arch. Neurol. 33:217-218.

Kaushall, P.I., M. Zetin, and L.R. Squire. 1981. A psychosocial study of chronic, circumscribed amnesia. J. Nerv. Ment. Dis. 169:383-389.

Kinsbourne, M., and F. Wood. 1975. Short-term memory processes and the amnesic syndrome. Pp. 258-291 in D. Deutsch and J.A. Deutsch, eds. Short-Term Memory. Academic Press, New York.

Klatsky, R.L. 1975. Human Memory: Structures and Processes. W.H. Freeman, San Francisco. 276 pp.

Kral, V.A. 1978. Benign senescent forgetfulness. Pp. 47-52 in R. Katzman, R.D. Terry, and K.L. Bick, eds. Alzheimer's Disease: Senile Dementia and Related Disorders. Raven Press, New York.

Lackner, J.R. 1974. Observations on the speech processing capabilities of an amnesic patient: several aspects of H.M.'s language function. Neuropsychologia 12:199-207.

Levin, H.S., V.M. O'Donnell, and R.G. Grossman. 1979. The Galveston Orientation and Amnesia Test: a practical scale to assess cognition after head injury. J. Nerv. Ment. Dis. 167:575-684.

Levin, H.S., A.L. Benton, and R.G. Grossman. 1982. Neurobehavioral Consequences of Closed Head Injury. Oxford University Press, New York. 279 pp.

Lynch, G., J. McGaugh, and N. Weinberger. 1984. Neurobiology of Learning and Memory. Guilford Press, New York. 528 pp.

Mahut, H. In press. Consolidation of memory--the hippocampus revisited. In N. Butters and L.R. Squire, eds. Neuropsychology of Memory. Guilford Press, New York.

Mahut, H., M. Moss, and S. Zola-Morgan. 1981. Retention deficits after combined amygdala-hippocampal and selective hippocampal resections in the monkey. Neuropsychologia 19:201-225.

Mair, W.G.P., E.K. Warrington, and L. Weiskrantz. 1979. Memory disorder in Koraskoff's psychosis: a neuropathological and neuropsychological investigation of two cases. Brain 102:749-783.

McAllister, T.W. 1983. Overview: pseudodementia. Am. J. Psychiatry 140:528-533.

McGaugh, J.L. 1982. Hormonal influences on memory. Annu. Rev. Psychol. 24:297-323.

Mercer, B., W. Wagner, H. Gardner, and F. Benson. 1977. A study of confabulation. Arch. Neurol. 34:429-433.

Michel, D., B. Laurent, N. Foyatier, A. Blanc, and M. Portafaix. 1982. Infarctus thalamique paramedian gauche. Rev. Neurol. 138:533-550. [In French.]

Miller, E. 1981. The nature of the cognitive deficit in senile dementia. Pp. 103-120 in N.E. Miller and G.D. Cohen, eds. Clinical Aspects of Alzheimer's Disease and Senile Dementia. Raven Press, New York.

Mills, R.P., and P.D. Swanson. 1978. Vertical oculomotor apraxia and memory loss. Ann. Neurol. 4:149-153.

Milner, B. 1958. Psychological defects produced by temporal lobe excision. Res. Proc. Assoc. Res. Nerv. Ment. Dis. 36:244-257.

Milner, B. 1962. Les troubles de la memoire accompagnant des lesions hippocampiques bilaterales. Pp. 257-272 in P. Passonant, ed. Physiologies de l'Hippocampe. Centre National de la Recherche Scientifique, Paris.

Milner, B. 1968a. Disorders of memory after brain lesions in man. Preface: material-specific and generalized memory loss. Neuropsychologia 6:175-179.

Milner, B. 1968b. Visual recognition and recall after right temporal lobe excision in man. Neuropsychologia 6:191-209.

Milner, B. 1970. Memory and the medial temporal regions of the brain. Pp. 211-226 in K.H. Pribram and D.E. Broadbent, eds. Biology of Memory. Academic Press, New York.

Milner, B. 1971. Interhemispheric differences in the localization of psychological processes in man. Br. Med. Bull. 27:272-277.

Milner, B. 1974. Hemispheric specialization: scope and limits. Pp. 75-89 in F.O. Schmitt and F.G. Worden, eds. The Neurosciences: Third Study Program. MIT Press, Cambridge, Mass.

Milner, B., and H.-L. Teuber. 1968. Alteration of perception and memory in man: reflection on methods. Pp. 268-375 in L. Weiskrantz, ed. Analysis of Behavioral Change. Harper & Row, New York.

Mishkin, M. 1978. Memory in monkeys severely impaired by combined but not by separate removal of amygdala and hippocampus. Nature 273:297-298.

Mishkin, M. 1982. A memory system in the monkey. Phil. Trans. R. Soc. London 298:85-95.

Moscovitch, M. 1982. Multiple dissociations of function in amnesia. Pp. 337-370 in L. Cermak, ed. Human Memory and Amnesia. Lawrence Erlbaum, Hillsdale, N.J.

Nadel, L., and S. Zola-Morgan. 1984. Infantile amnesia: a neurobiological perspective. Pp. 145-172 in M. Moscovitch and E. Saroka, eds. Infant Memory. Plenum Publishing, New York.

Nemiah, J.C. 1980. Dissociative disorders. Pp. 1544-1561 in H. Kaplan, A. Freedman, and B. Sadock, eds. Comprehensive Textbook of Psychiatry. Williams & Wilkins, Baltimore.

Newcombe, F. 1983. The psychological consequences of closed head injury: assessment and rehabilitation. Injury 14:11-136.

Norman, D.A. 1982. Learning and Memory. W.H. Freeman, San Francisco. 129 pp.

Ojemann, G.A. 1971. Alteration in nonverbal short term memory with stimulation in the region of the mammillothalamic tract in man. Neuropsychologia 9:195-201.

O'Keefe, J., and L. Nadel. 1978. The Hippocampus as a Cognitive Map. Oxford University Press, London. 570 pp.

Olton, D.S., J.T. Becker, and G.E. Handlemann. 1979. Hippocampus, space, and memory. Behav. Brain Sci. 2:313-365.

Osbourne, D.P., E.R. Brown, and C.T. Randt. 1982. Qualitative changes in memory function: aging and dementia. Pp. 165-170 in S. Corkin, K.L. Davis, J.H. Growdon, E. Usdin, and R.J. Wurtman, eds. Alzheimer's Disease: A Report of Progress in Research. Raven Press, New York.

Osterrieth, P. 1944. Le test de copie d'une figure complexe. Arch. Psychol. 30:206-356.

Pattern, B.M. 1972. The ancient art of memory. Arch. Neurol. 26:25-31.

Price, T.R.P. 1982. Short- and long-term cognitive effects of ECT. I. Effects on memory. Psychopharmacol. Bull. 18:81-91.

Randt, C.T., E.R. Brown, and D.P. Osbourne. 1980. A memory test for longitudinal measurement of mild to moderate deficits. Clin. Neuropsychol. 2:164-194.

Reitan, R.M., and L.A. Davison. 1974. Clinical Neuropsychology: Current Status and Applications. V.A. Winston, Washington, D.C.

Ross, E.D. 1980a. Sensory-specific and fractional disorders of recent memory in man. I. Isolated loss of visual recent memory. Arch. Neurol. 37:193-200.

Ross, E.D. 1980b. Sensory-specific and fractional disorders of recent memory in man. II. Unilateral loss of tactile recent memory. Arch. Neurol. 37:267-272.

Russell, E.W. 1975. A multiple scoring method for assessment of complex memory functions. J. Consult. Clin. Psychol. 43:800-809.

Russell, E.W. 1981. The pathology and clinical examination of memory. Pp. 287-319 in S.B. Filskov and T.J. Boll, eds. Handbook of Clinical Neuropsychology. John Wiley & Sons, New York.

Russell, W.R., and P.W. Nathan. 1946. Traumatic amnesia. Brain 69:280-300.

Ryle, G. 1949. The Concept of Mind. Hutchinson, London. 334 pp.

Sanders, H.I., and D.K. Warrington. 1971. Memory for remote events in amnesic patients. Brain 94:661-668.

Schacter, D., and H.F. Crovitz. 1977. Memory function after closed head injury: a review of the quantitative research. Cortex 13:150-176.

Schacter, D.L., and M. Moscovitch. 1984. Infants, amnesics, and dissociable memory systems. Pp. 173-216 in M. Moscovitch and E. Saroka, eds. Infant Memory. Plenum Publishing Corp., New York.

Schacter, D., P.L. Wang, E. Tulving, and P.C. Freedman. 1982. Functional retrograde amnesia: a quantitative case study. Neuropsychologia 20:523-532.

Scoville, W.B., and B. Milner. 1957. Loss of recent memory after bilateral hippocampal lesions. J. Neurol. Neurosurg. Psychiatry 20:11-21.

Seltzer, B., and D.F. Benson. 1974. The temporal pattern of retrograde amnesia in Korsakoff's disease. Neurology 24:527-530.

Small, I.F. 1974. Inhalent convulsive therapy. Pp. 65-77 in M. Fink, S. Kety, J. McGaugh, and T.A. Williams, eds. Psychobiology of Convulsive Therapy. John Wiley & Sons, New York.

Speedie, L.J., and K.M. Heilman. 1982. Amnesic disturbance following infarction of the left dorsomedial nucleus of the thalamus. Neuropsychologia 20:597-604.

Squire, L.R. 1974. Remote memory as affected by aging. Neuropsychologia 12:429-435.

Squire, L.R. 1981. Two forms of human amnesia: an analysis of forgetting. J. Neurosci. 1:635-640.

Squire, L.R. 1982a. The neuropsychology of human memory. Annu. Rev. Neurosci. 5:241-273.

Squire, L.R. 1982b. Comparisons between forms of amnesia: some deficits are unique to Korsakoff syndrome. J. Exp. Psychol. Learn. Mem. Cognition 8:560-571.

Squire, L.R. 1982c. Neuropsychological effects of ECT. Pp. 169-186 in W.B. Essman and R. Abrams, eds. Electroconvulsive Therapy. Spectrum, New York.

Squire, L.R., and N.J. Cohen. 1982. Remote memory, retrograde amnesia, and the neuropsychology of memory. Pp. 275-303 in L. Cermak, ed. Human Memory and Amnesia. Lawrence Erlbaum, Hillsdale, N.J.

Squire, L.R., and N.J. Cohen. 1984. Human memory and amnesia. Pp. 3-64 in G. Lynch, J.L. McGaugh, and N.M. Weinberger, eds. Neurobiology of Learning and Memory. Guilford Press, New York.

Squire, L.R., and H.P. Davis. 1981. The pharmacology of memory: a neurobiological perspective. Annu. Rev. Pharmacol. Toxicol. 21:323-356.

Squire, L.R., and M.M. Fox. 1980. Assessment of remote memory: validation of the television test by repeated testing during a seven-year period. Behav. Res. Methods Instrument. 12:483-586.

Squire, L.R., and R.Y. Moore. 1979. Dorsal thalamic lesion in a noted case of chronic memory dysfunction. Ann. Neurol. 6:503-506.

Squire, L.R., and P.C. Slater. 1975. Forgetting in very long-term memory assessed by an improved questionnaire technique. J. Exp. Psychol. Hum. Learn. Mem. 104:50-54.

Squire, L.R., and P.C. Slater. 1978. Anterograde and retrograde memory impairment in chronic amnesia. Neuropsychologia 16:313-322.

Squire, L.R., and P.C. Slater. 1983. Electroconvulsive therapy and complaints of memory dysfunction: a prospective three-year follow-up study. Br. J. Psychiatry 142:1-8.

Squire, L.R., and S. Zola-Morgan. 1983. The neurology of memory: the case for correspondence between the findings for man and non-human primate. Pp. 200-268 in J.A. Deutsch, ed. The Physiological Basis of Memory. 2nd Ed. Academic Press, New York.

Squire, L.R., C.D. Wetzel, and P.C. Slater. 1979. Memory complaint after electroconvulsive therapy: assessment with a new self-rating instrument. Biol. Psychiatry 14:791-801.

Sternberg, D.E., and M.E. Jarvik. 1976. Memory functions in depression. Arch. Gen. Psychiatry 33:219-224.

Stromgren, L.S. 1977. The influence of depression on memory. Acta Psychiatr. Scand. 56:109-128.

Strub, R.L., and F.W. Black. 1977. The Mental Status Examination in Neurology. F.A. Davis, Philadelphia. 182 pp.

Talland, G.A. 1965. Deranged Memory. Academic Press, New York. 356 pp.

Teasdale, G., and B. Jennett. 1974. Assessment of coma and impaired consciousness: a practical scale. Lancet ii:81-84.

Terry, R.D., and P. Davies. 1980. Dementia of the Alzheimer type. Annu. Rev. Neurosci. 3:77-95.

Teuber, H.-L., B. Milner, and H.G. Vaughan. 1968. Persistent anterograde amnesia after stab wound of the basal brain. Neuropsychologia 6:267-282.

Thompson, R., T. Berger, and J. Madden. 1983. Cellular processes of learning and memory in the mammalian CNS. Annu. Rev. Neurosci. 6:447-492.

Victor, M., R.D. Adams, and G.H. Collins. 1971. The Wernicke-Korsakoff Syndrome. F.A. Davis, Philadelphia. 206 pp.

Warrington, E.K., and L. Weiskrantz. 1982. Amnesia: a disconnection syndrome? Neuropsychologia 20:233-248.

Wechsler, D. 1945. A standardized memory scale for clinical use. J. Psychol. 9:87-93.

Weeks, D., C.P.L. Freeman, and R.E. Kendell. 1980. ECT. III. Enduring cognitive deficits? Br. J. Psychiatry 137:26-37.

Weiskrantz, L. 1982. Comparative aspects of studies of amnesia. Phil. Trans. R. Soc. London 298:97-109.

Weiskrantz, L., and E.K. Warrington. 1979. Conditioning in amnesic patients. Neuropsychologia 17:187-194.

Wells, C.E. 1979. Pseudodementia. Am. J. Psychiatry 136:895-900.

Wickelgren, W.A. 1979. Chunking and consolidation: a theoretical synthesis of semantic networks, configuring in conditioning, S-R v. cognitive learning, normal forgetting, the amnesic syndrome and the hippocampal arousal system. Psychol. Rev. 86:44-60.

Wilson, R.S., A.W. Kaszniak, and J.H. Fox. 1981. Remote memory in senile dementia. Cortex 17:41-48.

Winograd, T. 1975. Understanding natural language. Pp. 185-210 in D. Bobrow and A. Collins, eds. Representation and Understanding. Academic Press, New York.

Zangwill, O.L. 1977. The amnesic syndrome. Pp. 104-117 in C.W.M. Whitty and O.L. Zangwill, eds. Amnesia. Butterworths, London.

Zelinski, E.M., M.K. Gilewski, and L.W. Thompson. 1980. Do laboratory tests relate to self-assessment of memory ability in the young and old? Pp. 519-544 in L.W Poon, J.L. Foxard, L.S. Cermak, D. Arenberg, and L.W. Thompson, eds. New Directions in Memory and Aging. Lawrence Erlbaum, Hillsdale, N.J.

Zola-Morgan, S. 1984. Toward an animal model of human amnesia: some critical issues. Pp. 316-329 in L.R. Squire and N. Butters, eds. The Neuropsychology of Memory. Guilford Press, New York.

Zola-Morgan, S., and L.R. Squire. 1985. Complementary approaches to the study of memory: human amnesia and animal models. Pp. 463-477 in N. Weinberger, J. McGaugh, and G. Lynch, eds. Memory Systems of the Brain: Animal and Human Cognitive Processes. Guilford Press, New York.

Zola-Morgan, S., L.R. Squire, and M. Mishkin. 1982. The neuroanatomy of amnesia: the amygdala-hippocampus vs. temporal stem. Science 218:1337-1339.

Zola-Morgan, S., N.J. Cohen, and L.R. Squire 1983. Recall of remote episodic memory in amnesia. Neuropsychologia 21:487-500.

NEUROPSYCHOLOGICAL EVALUATION

Youngjai Kim

This presentation has a twofold purpose: to briefly outline and discuss what a comprehensive neuropsychological assessment is and to discuss ways in which subtle neuropsychological deficits are to be assessed.

Neuropsychological evaluation can be defined as a procedure in which the nature and extent of behavioral deficits that result from central nervous system dysfunctions are objectively measured. This procedure is predicated upon what we know of brain-behavior relationships. As new information about this relationship continues to emerge, it is reasonable to assume that the assessment procedures will undergo considerable change and refinement in years to come.

Neuropsychological assessment has been used in diagnosis, in research, and in making realistic patient management and rehabilitation plans. While the specific nature and scope of a neuropsychological test will be determined by the purpose for which it is used, any initial assessment of an individual or group of subjects must be comprehensive. Although the impairment of cognitive abilities will provide the most critical information in neuropsychological assessment, this index of impairment must be interpreted not in isolation, but within the context of an overall picture of cognitive ability. This comprehensive information is particularly important when the assessment is used in diagnosis, in which profile analysis plays an important role, and in making patient management plans, for which delineating the patient's strengths and weaknesses is essential to making realistic and effective plans. Comprehensive evaluation, in this context, does not refer to the depth of evaluation, but to its width, in the sense that a number of essential cognitive abilities that constitute an overall neuropsychological picture must be covered by the assessment procedure.

In any neuropsychological assessment, two frames of reference are kept in mind simultaneously. One pertains to behavioral representation of higher cortical functions from a neurological perspective, and the other pertains to our knowledge of human intelligence from a psychological perspective. Studies have shown that there is a functional asymmetry between the two hemispheres of the brain: verbal materials tend to be processed in the left hemisphere, nonverbal materials in the right hemisphere. Within each hemisphere, different lobes are, in turn, known to process different types of cognitive tasks. Crudely, the frontal lobes process conceptual and abstraction tasks, the temporal lobes process memory and auditory materials, the parietal lobes process visuoconstructional tasks, and the occipital lobes process visuoperceptual tasks.

Parallel to this corticobehavioral frame of reference is the psychological frame of reference. In any comprehensive assessment, a number of major cognitive functions must be evaluated. These functions include verbal ability, visuospatial ability, memory, attention and concentration abilities, and higher-level conceptual abilities.

VERBAL ABILITY

Assessment of verbal ability entails evaluation of two somewhat discrete aspects of verbal skills: basic language skills and verbal intelligence. Tests of aphasia reflect the basic language skills that should be evaluated, and tests such as the verbal subtests of the Wechsler Adult Intelligence Scale (WAIS) are used to assess verbal intelligence. Materials obtained during the interview and the administration of the verbal subtests of the WAIS provide the examiner with information on the subject's auditory comprehension, verbal expression, and some naming skills; however, they do not provide information on the subject's reading, writing, and repetition skills, which play important roles in making a diagnosis of aphasia. Repetition, in particular, plays a significant role in the differential diagnosis of subtypes of aphasia. In conduction aphasia, for instance, the subject manifests severe repetition deficits in the presence of relatively intact speech output (Goodglass and Kaplan, 1972), and without a test of repetition, this aphasic problem is not likely to be diagnosed.

The cortical areas associated with language disturbance have been extensively examined and will not be elaborated on here. For a right-handed person, various language problems are associated with a left-hemisphere lesion ranging from Broca's area anteriorly to Wernicke's area posteriorly.

VISUOSPATIAL ABILITY

At least three aspects of visuospatial ability should be examined separately. If a subject is unable to perform a visuoconstructional task such as the block design subtest of the WAIS, it is not clear whether the deficit lies in his deficient visuoperceptual ability, in his inability to mentally manipulate the required task, or in his inability to handle the motor aspect of the task.

Visuospatial deficits are associated with fronto-temporo-parieto-occipital lobe involvement. Whereas a right-hemisphere lesion, especially in the post-Rolandic fissure, produces the most pronounced visuospatial deficits, left parietal lobe lesions also produce visuoconstructional deficits. Some investigators (Hecaen and Assal, 1970) claim that there are qualitative differences in the types of errors committed on constructional tasks between patients with left or right parietal lobe lesions.

ATTENTION AND CONCENTRATION

Attention is one of the cognitive functions sensitively affected by any central nervous system (CNS) dysfunction. Also, in the context of a neuropsychological assessment, selective impairment of verbal and nonverbal attention has been seen in association with left- and right-hemisphere lesions, respectively (Diller and Weinberg, 1972; Kim *et al*., 1980).

The anatomical structure responsible for maintaining attention is the diencephalic extension of the reticular formation (Magoun, 1966). As attention represents a complex interaction of the limbic, neocortical, and ascending activating systems, attentional deficit could occur with damage in many areas of the brain.

Memory

Disturbances in memory often are the most common initial symptoms and complaints seen in early phases of CNS dysfunctions. In a neuropsychological assessment, verbal memory and nonverbal memory are assessed separately, as they frequently reflect lateralized CNS involvement. It is not uncommon to see a patient with significantly impaired nonverbal memory yet well-preserved verbal memory.

Temporal lobe lesions, left for verbal memory and right for nonverbal memory, are generally associated with memory deficits. However, subcortical structures such as the hippocampus, dorsomedial nucleus of the thalamus, and mammillary bodies are also implicated in memory functions.

Conceptual and Abstraction Abilities

As the ability to perform these higher-order cognitive tasks entails an intact ability to process basic attention, memory, and language skills, it is highly susceptible to the effects of neurological dysfunctions. At times, deficits in these higher-level functions are detected before deficits in other basic skills, and these tasks are sensitive in identifying subtle neurological involvement or an early phase of a neurological disease process.

The higher cognitive functions, although localized in the cortex, are not as focused as language functions, but are considered widely represented throughout the cortex. This diffuse representation makes this function particularly vulnerable to dysfunctions in various parts of the cortex.

Given the results of this initial comprehensive neuropsychological tests, a neuropsychologist asks the following set of questions.

What are the nature and the extent of cognitive deficits?
Are there indications of either focal or diffuse CNS involvement?
If the lesion is focal, what is the locus of the lesion?
Is the neuropsychological profile consistent with any of the known profiles associated with a specific neurological syndrome?
Are there any indications for further testing of certain specific cognitive functions?

A number of considerations should be kept in mind in interpreting the answers to these questions. Few of the focal deficits occur in pure form; rather, they tend to occur in a background of some degree of general psychological impairment. On the other hand, an absence of psychological deficit does not necessarily imply an absence of neurological involvement. Very mild neurological problems or lesions located in relatively silent areas of the brain might not produce measurable behavioral changes. Certain configurations or clusters of psychological deficits might signal specific neurological dysfunctions. When this is suspected and the available test data are not complete, additional tests should be used to check the hypothesis. This aspect of the procedure highlights the competence of a neuropsychologist. The quickness with which the examiner can formulate a number of possible hypotheses on the basis of minimum test data and the ingenuity with which he or she proceeds to attain missing information would vary from one examiner to another. Let us examine a hypothetical case of an associative visual agnosia. Patients with this syndrome have adequate visual perception and an adequate ability to recognize objects and demonstrate their use, but they are unable to name these objects. Left occipital lobe and posterior corpus callosum infarctions are associated with this syndrome. When available test data present a pattern in which severe naming deficit is seen in the presence of otherwise adequate language skills and visuoperceptual ability, a neuropsychologist should be alerted and carry out further evaluation to test the hypothesis and associative visual agnosia. The patient should be asked to demonstrate actual use of objects, and the discrepancy between his ability to name visually presented and verbally described objects should be tested. As these patients also frequently manifest alexia without agraphia, this pattern of skills should also be examined.

In this sense, a comprehensive neuropsychological assessment battery such as I have described, cumbersome as it already is, is just a basis from which further evaluation may proceed.

To a certain extent, there is a built-in system of validation for neuropsychological assessment when used as an individual diagnostic procedure. The results of the test interpretation are validated against clinical neurological examination findings or neuroradiological findings, such as computerized tomography (CT) scans. The majority of systematic, well-controlled studies in which the validity of neuropsychological tests is examined tended to involve a single test or cluster of tests designed to evaluate a particular cognitive function. However, a few studies evaluated the validity of a comprehensive neuropsychological assessment battery. Reitan (1962) attempted to infer the locus, extent, and etiology of lesions on the basis of the Halstead-Reitan Battery. He found generally high agreement between the inferences made on the basis of test results and independently derived neurological ratings. In another study (Kim *et al*., 1979), the efficacy of another comprehensive neuropsychological test battery was examined.

In that study, my colleagues and I examined the relationship between neuropsychological test results and CT scan findings for 20 stroke patients. The paradigm used was as follows. A comprehensive test battery and CT scans, both with and without the contrast material, were administered to each subject. Two neuropsychologists were independently assigned the following tasks: to identify the locus and extent of cerebral lesions based on the patient's test results, age, and level of education, and to specify the extent of psychological deficits on the basis of the CT scan readings provided by a neuroradiologist. The age and level of education of each patient were also provided. The tests of verbal ability used were the WAIS (information, comprehension, simlarities, vocabulary subtests), the Boston diagnostic aphasia examination, and the Wingfield object-naming test. Tests of nonverbal ability included general (WAIS performance scale), visual (Rey-Osterrieth figure test, Hooper visual organization test, Royer-Holland block rotation test), auditory (seashore rhythm test), and tactile (stereognosis test, Seguin formboard test) components. For attention and concentration, the digit sequencing test, word sequencing test, Knox cube test, 4 X 4 block sequencing test, WAIS (digit span and mental arithmetic subtests), and WMS (mental control subtest) were used. Tests of memory included verbal (WMS), visual (Benton visual retention test, Rey-Osterrieth figure test [recall]), auditory (seashore tonal memory), and tactile (Seguin formboard test) components. Abstraction and conceptualization were assessed with the Porteus maze test, Wisconsin card-sorting test, proverbs, and trail-making test. Other tests included color naming and matching, drawing geometric figures, clock drawing and setting, finger agnosia test, left-right orientation test, written arithmetic test, word fluency test, and the Minnesota Multiphasic Personality Inventory.

The psychologists' test scores were transformed into deficit scores (Table 1). The CT scan readings were also scored (Table 2). The locus of lesion to be specified was the frontal, temporal, parietal, or occipital lobe for both the left and right hemispheres. In addition, ventricular and sulcal enlargement and other categories were available for use. Briefly, the findings indicated that identification of the laterality of the lesion based on the test results was accurate on 80%

TABLE 1. Deficits as scored by psychologists

Test	Score[a]			
WAIS: Verbal IQ	1	2	3	4
Information	X			
Comprehension		X		
Similarities	X			
Vocabulary	X			
Rey-Osterrieth copy			X	
Hooper visual organization			X	

[a] Scores ranged from 1 (no deficit) to 4 (severe deficit).

TABLE 2. CT lesions as scored by radiologists

	Site and size of lesion	
Site of lesion	Left hemisphere	Right hemisphere
Frontal lobe		4
Temporal lobe		3
Parietal lobe		3
Occipital lobe		
Ventricular enlargement		
Sulcal enlargment		
Other		

[a] Scores ranged from 1 (no lesion) to 4 (severe lesion).

of the subjects; specification of psychological deficits based on CT scan readings was more accurate than identification of lesion sites based on test results was; agreement on the lesion site between the two neuropsychologists was greater than that between the neuroradiologist and the psychologist; and generally, the larger the size of the lesion, the more severe the psychological deficits.

On the whole, neuropsychological assessment of focal lesions tends to be relatively straightforward and its reliability and validity are generally adequate.

Meier (1970), in his discussion of assessing subtle psychological deficits associated with cardiovascular disease, stated that such deficits require elaborate, time-consuming, and complicated behavioral testing devices.

When a patient with unequivocal neurological symptoms fails to show behavioral deficits on psychological tests, it could be due to either a true absence of behavioral changes accompanying the neurological symptoms or the insensitivity of the tests used.

The areas of cognitive function shown to be sensitively affected by known or suspected diffuse CNS dysfunction provide clues about where to begin. Attention and concentration, new learning and short-term memory, speed of reaction, and ability to handle higher-level cognitive tasks appear to be more sensitive in reflecting diffuse and subtle CNS involvement than other tasks. Goodglass and Kaplan (1972) and Lezak (1977) describe impairment in conceptual thinking, slowing of ideational processes, reduced attention, stimulus boundness, and memory impairment as deficits associated with nonspecific neurological deficits among brain-damaged patients. Spieth (1962, 1967) and Szafron (1963, 1965), on the other hand, demonstrated that among those with cardiovascular insufficiency, impairment in speed of reaction, decision time in complex judgments, and short-term memory under conditions of interference are the most frequently seen cognitive deficits.

When very mildly impaired, equivocal performance is noted on many of these sensitive tasks against a background of intact performance on others, subtle CNS dysfunction should be suspected.

Certain complex tasks have been shown to be sensitive in reflecting subtle deficits. Witt et al. (1984) administered psychological tests to 16 anorectics, 16 age-matched depressed patients, 16 juvenile-onset diabetics, and 16 normal controls. The tests administered were the information and digit span subtests of the WAIS, the trailmaking and visual reproduction tests of the Wechsler Memory Scale (WMS), and the symbol-digit learning test (SDLT). Their findings indicated that SDLT was the only test that discriminated the anorectics from the other groups and was also found to correlate with duration of illness. In this test, seven symbols were paired with seven digits. After being shown a pair, the subject was shown one of the symbols alone and asked to supply the digit that was paired with it.

The digit-symbol subtest of the WAIS, which is similar to the above test, was also shown to be sensitive to generalized CNS involvement among brain-damaged patients. In the WAIS test, nine digits are paired with nine symbols. A series of digits is presented, and the subject is asked to write in the matching symbols as fast as possible. When the nature of these tasks is analyzed, they are found to entail sustained attention, memory, scanning, transformation of symbols into digits, and speed. It is possible that some of these components play greater roles in enhancing the sensitivity of the task than others. For instance, the transformation of the materials from one stimulus modality into another, such as digits to symbols, might involve rapid shifting, activation, and integration of different parts of the brain.

Examining the qualitative aspect of subjects' performance strategy could also provide important clues to the level of difficulty the subjects experience. The two performances of the copying task shown in Fig. 1 scored for accuracy alone would yield the same score. However, if the different strategies used in arriving at the same score are examined, the two performances appear quite different and may represent different levels of CNS integrity.

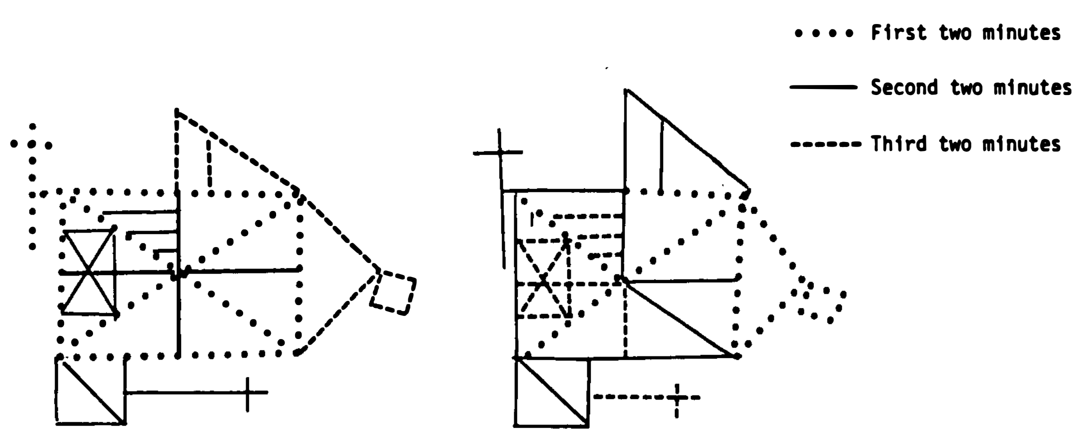


FIGURE 1. Rey-Osterrieth figure copying test. Two different sequences in copying strategy. Symbols: ...., first 2 min; ----, second 2 min; - - - -, third 2 min.

The tests that have been shown to be sensitive to subtle deficits are complex tasks that call for manipulating multiple factors with speed or that require efficient strategy to carry out.

The problem of validating the assessment of subtle deficits is more complex than it is for focal deficits. This aspect of assessment is just beginning to receive more attention. Systematic and rigorous examination of some of the issues described earlier, such as comparison of conventionally and qualitatively scored test results, may provide a starting point for further exploration.

Finally, a brief comment is in order with reference to two prevailing approaches to neuropsychological evaluation. One is a fixed test battery approach and the other is a flexible test battery approach. The Halstead-Reitan battery and the Luria-Nebraska neuropsychological battery are examples of the former. In this approach, the same test battery is administered to all candidates for neuropsychological assessment in as uniform a manner as possible. Each candidate's test score pattern is interpreted against the prototypic patterns for various diagnostic categories. The advantage of this approach is that the batteries can be administered by trained, competent technicians. This is an economical way of attaining neuropsychological information. The standardized test score interpretation also enables someone with less than expert knowledge of neuropsychological assessment to interpret a given patient's neuropsychological test data. The disadvantages of this approach are that it is difficult to assess and specify the precise nature and extent of neuropsychological deficits of a given patient in satisfactory detail and that deficits in certain idiosyncratic cognitive parameters that are not included in the battery might remain unassessed. In the flexible test approach, the major emphasis is placed on the uniqueness of the individual behavioral manifestations that result from neurological dysfunctions. Thus, the selection and administration of the tests depend on the given patient's specific pattern of deficits. Concentrated assessment of areas of deficit enables the neuropsychologist to explicate the specific nature of the deficit as well as the degree of deficit in detail. The drawback of this approach, however, is that the quality of the neuropsychological assessment greatly depends on the competence of a given neuropsychologist. There are, thus, advantages and disadvantages in both approaches. The choice to subscribe to either approach is a personal matter. Some of the important characteristics of a neuropsychologist, regardless of the specific approach one adheres to, are open-minded curiosity and an ability to observe and assess behavior in a systematic and creative manner.

REFERENCES

Diller, L.S., and J. Weinberg. 1972. Differential aspects of attention in brain damaged patients. Percept. Motor Skills 35:71-81.

Goodglass, H., and E. Kaplan. 1972. Assessment of Aphasia and Related Disorders. Lea & Febiger, Philadelphia.

Hecaen, H., and G. Assal. 1970. A comparison of constructive deficits following right and left hemispheric lesions. Neuropsychologia 8:289-303.

Kim, Y., F. Boller, F. Roger, and C. Bonstelle. 1979. Efficacy of neuropsychological examination. Paper presented at the International Neuropsychological Society Meeting, New York, February 1979.

Kim, Y., F. Roger, C. Bonstelle, and F. Boller. 1980. Temporal sequencing of verbal and nonverbal materials: the effect of laterality of lesion. Cortex 16:135-143.

Lezak, M. 1977. Neuropsychological Assessment. Oxford University Press, New York.

Magoun, H.G. 1966. Neural plasticity and the memory process. In M. Rinkel, ed. Biological Treatment of Mental Illness. L.C. Page, New York.

Meier, M. 1970. Objective behavioral assessment in diagnosis and predictor. In A. Benton, ed. Behavioral Changes in Cerebrovascular Disease. Harper & Row, New York.

Reitan, R. 1962. Psychological deficit. Annu. Rev. Psychol. 13:415-444.

Spieth, W. 1962. Abnormally slow perceptual motor task performances in individuals with stable, mild to moderate heart disease. Aerospace Med. 33:370.

Spieth, W. 1967. Reaction to cardiovascular change. Mayo Clinic Proc. 42:632-636.

Szafron, J. 1963. Age differences in choice reaction time and cardiovascular status among pilots. Nature 200:904-906.

Szafron, J. 1965. Age differences in the rate of gain of information, signal detection strategy, and cardiovascular status among pilots. Am. Psychol. 20:581.

Witt, E., C. Ryan, and G. Hsu. 1984. Learning impairments associated with anorexia nervosa. Paper presented at the International Neurological Society Meeting, February 1984, Houston, Tex.

SESSION II

INTRODUCTION--EFFECTS OF SELECTED PHYSIOLOGICAL AND ENVIRONMENTAL VARIABLES ON COGNITION

The number of studies done specifically to evaluate the effects of nutritional variables on cognitive performance is rather small, especially for studies done recently with state-of-the-art cognitive assessment techniques. Nevertheless, cognitive testing methods have been shown to be useful in assessing the effects of a variety of factors on mental performance. The purpose of this session was to review four areas in which cognitive performance tests have been used. The first paper reviews what is known about the effects of nutrition on cognition; these results come largely from studies done prior to 1960. Later chapters review the cognitive effects of cholinergic drugs, noise, heat, and other stressors, and neurotoxic agents. In each area the authors emphasize the testing methods that have been most useful.

NUTRITIONAL DEFICIENCIES AND COGNITION

David A. Levitsky and Barbara J. Strupp

The question of whether nutrition, or, perhaps more importantly, poor nutrition, affects human cognition is as old as civilization. Certainly every culture had ideas about the relationship between foods or the lack of certain foods and the mind. Even today, such diverse behavioral pathologies as schizophrenia, depression, anxiety, criminal behavior, hyperkinetic syndrome, memory disorders, premenstrual tension, and many others have been linked to food intake.

Despite the keen interest in the effects of nutrition on the mind and behavior, relatively little research has been done. The reason for this dearth of research is not clear. This relationship bridges two disciplines, nutrition and psychology, which have a number of overlapping research areas, such as obesity and feeding behavior, social and cultural determinants of food intake, neurotransmitter metabolism, etc. In addition, the particular problems involved in determining the effect of nutrition on behavior do not require unusual or costly techniques for solution. Why then has research not proceeded in this area of science at the same rate that it has in other areas that interface nutrition and psychology? The answer, we believe, is that scientists in general don't really believe that food can affect their minds and their behavior.

The reason most scientists don't believe nutrition can affect the mind is twofold. First, most of the research published so far has shown that even severe nutritional deficiencies produce only a minimal alteration in commonly measured aspects of cognitive function. Second, and perhaps more important, people (including scientists) are not generally concerned about subtle influences on cognition. They know from personal experience that going without breakfast doesn't have much of an effect on their own thinking, except perhaps that they may be preoccupied with food. It is difficult, therefore, to convince people, particularly scientists, to study a phenomenon that they intuitively don't believe and moreover appears to be weak at best. Tenure is not obtained by trying to publish studies claiming "no effect."

One major exception to this lack of interest in the possible subtle effects of nutrition on behavior is the military. They are very concerned with the need to understand the conditions that will optimize the performance of their personnel. It is therefore not surprising that the most enlightening research that has been or possibly ever will be done in this area was sponsored by the U.S. Army. Most of what we know today about the effects of calorie and vitamin restriction on adult behavior and physical performance was obtained through research sponsored by the military.

However, before discussing these and other data about the effects of malnutrition on adult behavior, we would like to review briefly the effects of dietary restriction on the cognitive development of children. More is known about the effects of malnutrition on the cognition of children than of adults, and in fact many strong similarities exist between the findings for these two groups.

NUTRITION AND THE COGNITION OF CHILDREN

The reason that the effects of nutrition on cognition have been studied much more intensely in children than in adults is that adult cognition is greatly influenced by the conditions that existed during childhood. Events during childhood that interfere with the development of intelligence may permanently limit the potential productivity of the adult--a link that therefore has important ramifications not only for the individual, but for society as well.

Perhaps the most serious question relating the effects of nutrition to cognition in children is whether malnutrition suffered early in life permanently limits adult intelligence (see Levitsky and Strupp, 1985, for a comprehensive review). The question was raised because clinical observations reported in the early 1960s suggested that malnourished children in certain parts of the world were showing signs of delayed intellectual development. These observations were coupled with the results of laboratory research demonstrating that much of the decrease in brain size caused by malnutrition was not reversed with nutritional therapy. It was not difficult for many to jump to the conclusion, therefore, that the developmental lags observed in malnourished children would be permanent.

Does malnutrition cause permanent mental retardation? The answer seems to depend on the kind of environment in which the children are raised. Malnourished children raised in impoverished environments in which there is a paucity of intellectual stimulation and educational encouragement show substantial deficits in IQ. In contrast, previously malnourished children raised in environments in which educational achievement is highly valued show almost no permanent effects of the poor nutrition on intelligence.

Understanding how the child's environment can modulate the effects of malnutrition on intelligence can teach us a great deal about how poor nutrition may be expected to affect adult cognition. There is very little evidence from studies on either animals or humans that even very severe malnutrition impairs the ability to learn or remember important environmental information. Apparently, the brain mechanisms responsible for these cognitive processes are well protected against the destructive effects of malnutrition.

Rather than impairing the ability to learn and retain information, malnutrition seems to cause a deficit in the propensity of a child or animal to actively acquire and process new information. Children, and indeed all young mammals, accrue most information about their environment on their own by being curious and exploring. In educationally enriched environments, the child is constantly surrounded by information that is important for intellectual development. Unfortunately, in educationally impoverished environments, relevant information and encouragement are not readily available and must be actively sought by the child.

The major effect of malnutrition is to suppress the young mammal's normal tendency to explore and acquire information about its environment. This deficit in endogenously motivated learning would be expected to have a greater effect on the development of intelligence in children raised in educationally impoverished environments than in educationally enriched environments simply because the acquisition of educationally relevant information is left to the child to a much greater degree in less-advantaged environments. Not only is the environment of socially advantaged children more saturated with educationally relevant material, but the acquisition of that information is more apt to be required by the parents. Thus, in an educationally enriched environment, although the effects of malnutrition may impair the child's endogenous mechanisms for acquiring information to the same degree as those of a malnourished child living in an impoverished environment, the deficit in the curiosity to acquire information is compensated for by the increased amount of available information and by the parents' encouragement and demands.

Besides inhibiting the child's curiosity to learn, malnutrition also seems to impair the degree to which information is processed and encoded. Children as well as adults sort, classify, and categorize information before encoding it into memory. This degree of information processing, which occurs spontaneously in well-fed children, appears to be suppressed by malnutrition. It must be emphasized, however, that malnourished children are fully capable of this degree of processing if it is required by the testing situation; they simply don't engage in the same depth of information processing when not forced to do so.

The conclusion that has emerged from the study of malnutrition and cognition in children is that although malnourished children are fully capable of learning and processing environmental information, the normal endogenous motivation for acquiring and processing that information is inhibited by poor nutrition, but their basic ability to acquire and process information seems to be unaffected.

OBSERVATIONS OF COGNITION DURING FAMINE

In an excellent review of the literature on the behavior of humans during famines, Keys et al. (1950) point out a common observation: no matter what the cause of malnutrition, the response is a reduction of spontaneous activity, apathy, irritability, and depression. Blanton (1919) made astute observations of German schoolchildren whose food ration was limited during World War I. He observed no deterioration in intelligence or school performance in children "of good breeding stock," or upper class, but the children of lower class showed rather marked declines in school performance. The teachers complained that all the children were frequently falling asleep in class, had difficulty sustaining attention and working for long periods, and were more restless and harder to discipline.

Fliederbaum et al. (1946) observed the effect of food restriction on the Jewish population of the Warsaw Ghetto during World War II. They emphasized the apathy and lack of interest these people had in their surroundings. Chernorutskii (1943) also noted apathy, dulling of emotions, and narrowing of intellectual interests among the inhabitants of Leningrad during a famine caused by the German blockade in World War II. Dols and van Arcken (1946) commented on the behavior of the Dutch people during food shortages of World War II: "The most striking symptom of starving persons is their total apathy, in spite of the fact that they often had lost their self control and were most irritable." Sorokin (1942), in reviewing the data from a number of famines, spoke of food restriction as causing an impairment in creative thinking.

In almost every one of the 29 observations of a natural famine, this same pattern of response had been observed: domination of all thoughts by food, severely decreased activity, lack of interest in anything (including sex) except food, and increased irritability. In no situation was impairment in cognitive ability mentioned.

EXPERIMENTAL CALORIC RESTRICTION AND COGNITION

The first systematic investigation of the effects of dietary restriction on mental functioning was done by Weygandt (1904). Despite the universally reported changes in the behavior of people during a natural famine, Weygandt found that fasting for 24 to 72 h had no effect on perceptual speed (the length of time necessary to recognize words), word association reaction time, or learning ability. Benedict (1915), in one of his classic studies of the effect of a 31-day fast on a "professional" faster, noted that although the subject started the fast "cheerful," he became progressively more "depressed." Benedict noted, however, that his intellectual capacity and expressive powers, as judged in conversations, did not seem to be affected.

Marsh (1916) measured the effects of a 1-week fast on himself and his wife. He noted that his memory deteriorated with prolonged fasting while his wife's improved. Unfortunately, the lack of a control practice period and the absence of statistics prevent any definitive conclusion from being made. Glaze (1928) subjected himself, his wife, and a professional faster to a 10-day fast and like the others before him was unable to detect any deterioration in cognition as measured by the ability to do mental arithmetic, name colors, and read letters backwards.

The first as well as the most careful and comprehensive quantitative study of the cognitive and behavioral effects of caloric restriction ever done was performed by Ancel Keys and Josef Brozek and their associates at the University of Minnesota (Keys et al., 1950). Brozek, a well-trained and methodical experimental psychologist, knew all the problems of testing humans for long periods of time: practice effects, declining ambition, monotony, and many others that plague and could possibly confound long-term studies of behavior.

Following a 12-week control period during which 32 subjects were trained on various tests of physical and mental performance, their diet was reduced from 3,492 to 1,658 kcal/day, a reduction of 52.5%. This period of caloric restriction continued for 24 weeks and resulted in a 24.2% reduction in body weight. The subjects were then studied for another 12 weeks during the recovery period. All the performance tests used were highly standardized and well accepted as sound tests of cognitive and physical performance. Table 1 shows the tests that were used.

All the subjects were well trained on each of the tests before they started the semistarvation part of the study. The tests were also administered during the rehabilitation period. In effect, the experiment included baseline, treatment, and recovery periods.

Considering first the tests of sensory function, it was clear that no change in visual acuity was observed during any part of the experiment. The visual threshold at 100 and 1 fc (1,076 and 10.76 lx) was not altered by semistarvation. Fluctuations in perception of the Necker cube also did not prove to be sensitive to the effects of semistarvation.

Flicker fusion frequency thresholds were statistically affected. The threshold dropped from 36.3 $\pm$ 3.2 (SD) during the control period to 35.3 $\pm$ 3.6 by the end of the starvation period. A within-subject comparison yielded a highly statistically significant effect ($F = 19.41$; $P > 0.01$). Unfortunately, the biological meaning of this decrease is not known. The threshold does decrease with age (Brozek and Keys, 1945; Simonson et al., 1941) and stress (Brozek and Keys, 1944). The authors, unconvinced of the biological importance of this difference, calculated a displacement ratio by dividing the mean within-subject difference in threshold by the between-subject SD obtained during the control period. The value they calculated, -0.31, they considered "not impressive." They concluded that "changes in the visual functions during the period of semi-starvation were either absent or negligible."

TABLE 1. Performance tests used by the Minnesota study

Function	Tests used
Sensory function	Auditory acuity (manual audiometer); visual (determined by the diameter of the smallest circle that can be seen at 15 ft.); flicker fusion frequency
Perceptual fluctuations	Necker cube fluctuation
Motor performance	Psychomotor test battery (Brozek *et al.*, 1946b)
Strength	Standard handgrip and back-lift dynometer
Speed	Pencil tapping
Coordination	Pattern-tracing task
Intelligence	Army General Classification Test; CAVD scale; Repeatable Test Battery (Guetzkow and Brozek, 1947)
Learning	Crossing-out 4's (crossing out the number 4 from a long series of numbers)
Personality	Interviews by three psychologists; diaries; MMPI; inventory of temperamental traits (Guilford and Martin, 1943); self-ratings; complaint questionnaire; man-to-man rating (subjects rated each other); drive ratings (food, sex, activity)

The effect of semistarvation on hearing acuity was quite different. Not only was the effect statistically significant, but contrary to expectation, semistarvation actually caused an improvement in hearing. This increase in hearing acuity occurred at almost every sound frequency tested, including 128, 512, 2,048, and 4,096 cycles/s, although the effect was not statistically significant at the highest frequency, 8,192 cycles/s. Hearing acuity returned to baseline after nutritional rehabilitation. Table 2 shows the results for one testing frequency, 2,048 cycles/s. The authors reported that the subjects frequently complained that their tolerance for loud noise, music, or speech was lowered during semistarvation.

One of the most striking effects of semistarvation is the decrease in muscular strength and endurance. Statistically significant decreases

TABLE 2. Hearing loss[a] during Minnesota study

Period	Hearing loss (dB), Mean ± SD (n = 32)	t	P
Control	0.58 ± 5.89		
Semistarvation	-4.36 ± 4.92		
Change[b]	-4.94 ± 4.05	7.36	0.01
Change during recovery[c]	5.75 ± 4.42	6.90	0.01

[a]Hearing loss as measured by a Maico D-5 audiometer at a frequency of 2,048 cycles/s.
[b]Individual-change scores from end of control period to end of semistarvation period.
[c]Individual-change scores from end of semistarvation period to end of recovery period.

due to semistarvation were found in hand dynometer, back dynometer, leg reaction time, gross body reaction time, manual speed, tapping speed, and pattern tracing measurements.

Although all the behavioral measures described so far involve neural systems, the major question we want to address is whether semistarvation affects cognition. As indicated in Table 1, the Minnesota studies used three different assessments of cognitive performance; the Army General Classification Test, the Completions, Arithmetic, Vocabulary, and Directions (CAVD) scales, and the Repeatable Test Pattern, although only the latter two were used during the period of semistarvation and recovery. In addition, a crossing-out test was used to analyze the rate of learning. All subjects were well practiced during the control period.

Before discussing the scales, it may be instructive to obtain a glimpse of the researchers' impression of whether semistarvation affects intellect and cognition. The authors (Keys et al., 1950, pp. 859-860) state:

> According to our clinical impressions, the intellective capacity was essentially unchanged. Throughout the stress the subjects talked intelligently, though with decreasing speed, and appeared to think clearly. Memory disturbances and a decrease in expressive power were only rarely encountered and did not seem to go beyond the range of normality. Quantitative tests of intellective performance were used to supplement the clinical judgment.

Indeed, almost none of the tests of mental function taken during the period of semistarvation showed a statistically significant change from the control period. The Repeatable Test Battery of intellective functions is made up of six subscales, none of which show an effect of starvation. These scales include flags (a test of the perception of spatial relations), first letter (examines word fluency), word-number recall (measures memory), multiplication (measures number facility), letter series (examines inductive reasoning), and number checking (examines perceptual speed).

The results of the CAVD test of intellective functions, however, were a bit different (Table 3). The CAVD scale was developed by E.L. Thorndike, the famous psychometrician at the Institute of Educational Research of Columbia University, and contains several different forms. To control for any practice effect, different forms were used at each of the three test periods: control, after 12 weeks of semistarvation, and after 24 weeks of semistarvation. The only scale that was not affected by starvation was the completion scale. In the completion test, the subject was required to supply words to a sentence to make it true or false.

The arithmetic scale required logical reasoning and computation. The vocabulary scale tested the ability of the subjects to choose among various synonyms of a given word. The directions scale examined the "ability to understand corrected discourse."

TABLE 3. CAVD test of intellective function[a]

	Scores (mean ± SD)				
Test	C	S12	*F*	S24	*F*
Completions	32.6 ± 7.7	32.9 ± 7.3	0.17	32.6 ± 7.6	0.00
Arithmetic	32.9 ± 10.4	36.7 ± 11.3	15.04[b]	35.9 ± 11.0	21.68[b]
Vocabulary	29.8 ± 8.3	32.1 ± 7.1	8.59[b]	32.1 ± 6.8	8.02[b]
Directions	35.6 ± 5.0	43.1 ± 7.4	1.93	32.9 ± 6.2	15.63[b]

[a]Scores for subtests and total raw scores were calculated during the control (C) and after 12 and 24 weeks of semistarvation (S12 and S24, respectively). *F* was calculated on the basis of within-subject effect.
[b]$p < 0.01$.

Thus, contrary to the rather definitive conclusion that the Minnesota researchers reached, their data contain suggestions that at least some aspects of cognition may be affected by semistarvation. However, the cognitive effects they did observe, although statistically significant, were very small--approximately a 7% decrease after 24 weeks of semistarvation. Moreover, performance on one of three scales, the vocabulary scale, actually improved rather than deteriorated with starvation.

One final cognitive test was performed by the Minnesota researchers--the test of crossing-out 4's. This test examined the learning of a number cancellation problem and was given at the end of the 24 weeks of starvation and again at the end of 18 weeks of rehabilitation. No differences were observed in the rate of learning this task. The researchers concluded that "this fits in with the other test results and with the clinical observations and increases our confidence that 'intelligence' was not adversely affected by the reduction of food intake and the consecutive decrease in body weight."

Although neither the clinical psychologists nor the cognitive tests could detect any evidence of deterioration in mental ability, the subjects all believed that they were worse off. "As semi-starvation progressed in the Minnesota Experiment, complaints of inability to concentrate for any period of time and of difficulty in developing thoughts became numerous. By the end of the semi-starvation period a large proportion of the men felt that their judgment had been impaired. They reported further that their general alertness and comprehension had declined."

Moreover, contrary to the tests of "intellective functions," tests of personality, particularly the Minnesota Multiphasic Personality Inventory (MMPI), consistently showed statistically significant changes in the personality of subjects during the period of caloric restriction. The major changes in MMPI measures can be seen in Table 4.

Of the scores, the greatest changes were found with the scales of hyperchondriasis, depression, and hysteria. These scales place a great deal of weight on physical complaints, bodily preoccupation, hypersensitivity, emotional shallowness, and lethargy (Dahlstrom and Welsh, 1960), all of which are undoubtedly increased during the extended period of semistarvation.

Two other personality tests were administered, the Rosenzweig picture-frustration test and the Rorschach test. Neither test detected any significant effect of caloric restriction (Franklin and Brozek, 1949).

TABLE 4. Scores on MMPI[a]

	Scores (mean ± SD)				
Scale	Control	S12	S24	F	R12
Hyperchondriasis	45.7 ± 3.4	58.2 ± 7.1	63.0 ± 6.7	110.1[b]	54.1
Depression	54.2 ± 6.5	64.8 ± 7.2	73.9 ± 8.2	109.6[b]	65.7
Hysteria	59.0 ± 6.1	65.8 ± 7.1	70.0 ± 8.2	52.2[b]	64.9
Psychopathic deprivation	52.2 ± 7.5	53.0	52.9 ± 6.5	0.3	53.0
Masculinity-femininity	69.8 ± 10.3	68.0	66.9 ± 8.1	6.4[c]	66.9
Paranoia	53.5 ± 5.0	53.4	54.1 ± 5.9	0.3	53.2
Psychastenia	45.7 ± 7.4	46.8	51.9 ± 7.7	21.1[b]	49.1
Schizophrenia	47.5 ± 7.5	49.2	55.1 ± 8.9	43.2[b]	48.5
Hypomania	51.0 ± 7.1	50.2	51.4 ± 8.7	0.1	50.6

[a]Scores were obtained during the control period, after 12 and 24 weeks of semistarvation (S12 and S24, respectively), and after 1, 6, and 12 weeks (R12) of recovery.
[b]$P < 0.01$
[c]$P < 0.05$

Although it was not measured directly, the Minnesota researchers remarked throughout their monograph that during the period of semistarvation the subjects became extremely depressed, apathetic, and irritable. As mentioned earlier, others have noted this kind of behavioral change during natural famines: Markowski (1945) in semistarved prisoners of war and Dols and van Arcken (1946) in the semistarved victims of the Dutch famine during World War II. The Minnesota researchers were so impressed by the changes induced by the semistarvation that they stated, "Perhaps the outstanding feature in both famine and the Minnesota experiment is depression and apathy" (Keys *et al*., 1950, p. 907).

VITAMIN DEFICIENCY AND COGNITION

A profile remarkably similar to that observed with caloric restriction has been seen with vitamin deficiencies. Spies *et al*.

(1943) noted that patients suffering from pellegra complained about difficulty in concentration, memory, and comprehension. A similar observation was made by Wilder (1944).

In one of the first experimental studies of the effects of vitamin B complex restriction on cognition in humans, O'Shea et al. (1942) observed a deterioration in the rate of learning various maze tasks after vitamin deficiency and a recovery in the maze-learning rate after vitamin restoration. Sebrell (1943) also noted disturbances in cognition with experimentally induced vitamin B deficiencies. He reported that advanced deficiency led to disturbances in memory, clouding of consciousness, and disorientation. Unfortunately, these findings were based on clinical impressions rather than quantitative measurements.

The first study of vitamin deficiency and behavior by the Minnesota group examined the effect of riboflavin. The period of deficiency lasted only 12 days, and no effect was observed in perceptual-motor tasks, visual perception, or strength measures (Keys et al., 1944). More recently, Sterner and Price (1973) studied the effects of a diet low in riboflavin by using a battery of psychomotor and cognitive tests. The study used a 14-day control period, a 56-day depletion period (0.07 mg of riboflavin per day), and a 14-day recovery period. The behavioral tasks included a brightness discrimination task and an Ortho-Rater task to measure visual perception and acuity, the Crawford Small Parts Dexterity Test (CSPDT), a hand dynometer, a pursuit motor task, a test of reaction time, the Wechsler Adult Intelligence Scale, the MMPI, and the Rorschach inkblot test. These researchers were unable to observe any effect of riboflavin deficiency on any of the measures of psychomotor performance or in any of the intelligence scales. However, they did find large changes in the MMPI, but not the Rorschach test: increases in hyperchondriasis, depression, psychopathological deviancy, and hypomania. Notice the increase in the scale of depression (Table 4).

In a study of the effects of a diet low but not deficient in total B vitamins on psychomotor and cognitive performance, Keys et al. (1945) also could not find any detectable effect on the speed of repetitive motor responses, whole-body reaction time, pattern tracing, speed and accuracy of a motor coordination task, critical flicker fusion frequency, or perceptual thresholds for pressure or vibration. Moreover, no effect of the restriction in B vitamins could be found in tests of cognition with the Porteus maze or in personality as measured by the MMPI and Cattell Cursive Miniature Situation Personality Test.

The effects of an acute deficiency were a bit different. Although no effect of an acute deficiency in B vitamins was observed on any of the tests of strength, a clear deterioration in psychomotor coordination was found, which recovered after 10 days of rehabilitation (Brozek et al., 1946b). Besides a sharp increase in complaints by the subjects due primarily to the appearance of anorexia and nausea after 10 days on the deficient diet, significant increases in MMPI scores were observed on the scales for hyperchondriasis, depression, and hysteria, the same three scales that were affected by semistarvation (Keys et al., 1950).

Berryman et al. (1947) performed a study similar to that of Brozek et al. (1946b) in which they examined the effect of a diet restricted in B vitamins on performance. Unlike the Minnesota group, Berryman et al. could not find any evidence of deterioration in psychomotor performance even after 18 weeks on the restricted ration. However, they did note a decrease in physical performance as well as an increase in hysteria, depression, and hyperchondriasis on the MMPI (Henderson et al., 1947).

Brozek et al. (1946a) observed that the increase in hysteria, depression, and hyperchondriasis scores brought about by an acute deficiency of the B vitamins could be almost totally reversed by administering thiamine. This finding suggests that the "neurotic triad" is caused by thiamine depletion and not by lack of the other B vitamins. This suggestion is supported by the observation of Williams et al. (1940) that depression and emotional instability were quite evident in subjects made thiamine deficient.

B vitamins are not the only nutrients that seem to affect behavior. Wilter et al. (1946) observed patients suffering from clinical scurvy. They reported that many of the patients were disoriented and depressed. Cutforth (1958) observed a similar picture in patients suffering from scurvy, but was most struck by their profound lethargy, apathy, and depression.

The first and only attempt to quantify the behavioral changes that accompany experimentally produced vitamin C deficiency was reported by Kinsman and Hood (1971). They observed that after a prolonged deficiency of ascorbic acid, physical performance decreased as measured by the Fleshman Physical Fitness test battery, but there was no deterioration in cognition as measured by several tests. However, like semistarvation and B vitamin deficiency, a vitamin C deficiency produces increases in the neurotic triad scores, as indicated by an increase in hyperchondriasis, depression, and hysteria in the MMPI.

DEPRESSION AND COGNITION

The results of the studies described so far seem to suggest that prolonged nutritional deficiencies have very little effect on intellectual abilities. Therefore, can we conclude that cognitive function is not impaired even if an individual is poorly nourished? Based on several lines of indirect evidence, it seems that the answer is no.

At the outset it is important to distinguish between cognitive ability, as assessed under laboratory conditions, and the spontaneous mental activity that characterizes an individual in daily life. The distinction was emphasized by Keys et al. (1950) to explain the marked contrast between the unimpaired test performance of the men during starvation and their striking apathy and the decrement in their "self-initiated mental activities" scores. Although these starved men were generally capable of normal intellectual functioning when they were in the testing situation and encouraged to perform well, they seemed totally uninterested in the intellectual activities that had previously been important to them.

In animal studies, too, it seems clear that malnutrition may differentially affect an animal's ability to learn and its propensity to acquire information (Strupp and Levitsky, 1983). Specifically, malnourished animals show unimpaired learning ability when learning is required to meet basic needs while showing profound deficits in learning when information is available but unnecessary for immediate survival or comfort. They seem capable of learning when they are forced to do so, but do not spontaneously acquire information to the same extent as do well-nourished animals when not forced.

It seems fair to conclude that even if nutritionally deficient subjects are able to perform adequately on structured tests in a formal testing situation, they may not be functioning optimally in their day-to-day life. This possibility is also suggested by the profound mood changes that have universally been found to accompany prolonged nutritional deficiencies--in particular, depression. In fact, many of the subtle but consistent cognitive changes that have been reported in experimental deficiency studies are qualitatively very similar to those that are characteristic of clinically depressed patients, although perhaps less severe.

Overall there does seem to be a correlation between severity of depression and degree of cognitive impairment in hospitalized depressed patients (e.g., Stromgren, 1977). However, similar to poorly nourished individuals, depressed patients are able to perform normally on the majority of cognitive tests in formal testing situations. In fact, many investigators have commented on the striking disparity between the profound affective changes seen in these patients and their relatively minor cognitive deficits (Friedman, 1964; Kendrick and Post, 1967).

However, it has also been noted that the performance of depressed patients in these testing situations may not be a good predictor of their cognitive functioning in life outside the testing situation. For example, Colbert and Harrow (1967) observed that although a group of depressed patients demonstrated no psychomotor retardation when tested after their admission to the hospital, there had been profound cognitive retardation (e.g., impairment of initiative, reduction in concentration, retardation in thinking) for 2 two weeks preceding hospitalization. The authors stress that except for severely depressed patients, most can "mobilize themselves" in these structured testing situations. Other studies of depressed patients have also emphasized that under the conditions of active support and encouragement usually present in these situations, these patients can achieve normal cognitive performance, often in contrast to the level of cognitive functioning in their daily life (e.g., Stromgren, 1977). In light of the depression that seems to accompany nutritional deficiencies (albeit minor in comparison to the depression suffered by many of these patients), it seems likely that here too there may be a disparity between the test performance of these individuals and their cognitive functioning under less structured conditions.

As noted above, although the cognitive deficits observed in clinically depressed patients are minor in comparison to the severity of their affective disorder, these patients do seem to be impaired cognitively compared with both normal controls and with themselves in a nondepressed state (e.g., Henry *et al*., 1973; Stromgren, 1977; Weingartner *et al*., 1981). However, studies that have tried to specify the type of cognitive impairment associated with depression have revealed that the deficit is rather specific and does not show up on all cognitive tasks.

The parameter that seems to be important in predicting the cognitive performance of depressed patients is the amount of effort required to perform optimally on the tasks. Scores on cognitive tasks that can be accomplished relatively automatically do not distinguish between control subjects and depressed patients. In contrast, tasks that require sustained effort and elaborate encoding for optimal performance do reveal cognitive deficits in these patients.

For example, in one recent study, depressed patients were asked to listen to and remember different word lists that differed both in the relatedness of the words in the list and in the clustering of the related words. Specifically, there were seven different 32-word lists: one list of random (unrelated) words; two lists that each contained two sets of 16 related words (e.g., 16 fruits and 16 vegetables); two lists that each contained words belonging to four categories; and two eight-category word lists. One of the two-category lists clustered each category together (AAA....BBB), while the other was unclustered (ABBABA...). Similarly, there was both a clustered and an unclustered version of the four- and eight-category word lists.

The depressed patients did not differ from the control subjects in recalling the lists that were highly organized and in which the organization was shown by obvious clustering. However, these depressed patients recalled significantly fewer words when the organization of the words was not obvious (unclustered lists). The results indicated that even if the related words were not clustered together, the controls were able to take advantage of the relatedness of the words and impose organization on the lists. In contrast, the depressed patients seemed to be forming nontransformed copies of the lists presented and therefore were only able to take advantage of the relatedness of the words in the clustered lists.

This task and others like it (Cohen et al., 1982; Weingartner et al., 1981) suggest that when subjects must impose structure or use elaborate encoding strategies to perform well, depressed patients perform more poorly than controls.

In addition, depressed patients have been found to be impaired in cognitive tasks that require sustained concentration (e.g., Friedman, 1964). For example, in two different tasks (reaction time to a light signal and a task requiring a response shift), Friedman found that the depressed subjects were only impaired on the last trial of each test, strongly suggesting a fatigue or sustained-attention effect. Tests that primarily assess psychomotor speed, such as reaction time tests and the digit-symbol test, have also revealed deficits in depressed subjects.

That fatigue and psychomotor retardation are associated with depression does not seem surprising. These findings are consistent with those of Cohen and his colleagues (1982) that a significant correlation existed between the performance of depressed patients on certain cognitive tests and their muscular strength as assessed with a dynometer. These authors interpreted the results as indicating a deficit in effortful processing. The reduced dynometer performance generally found in subjects with nutritional deficiencies might also be associated with this type of cognitive deficit.

In summary, the cognitive functions that seem to be specifically altered in the depressed state include the initiation of intellectual activities in daily life, effortful and elaborate encoding strategies (information seems to be only superficially processed), and psychomotor speed and sustained concentration. Several points, however, should be stressed about the cognitive alterations associated with depression. First, although significant, the reported deficits are generally small. For example, even in one study that found highly significant correlations between the scores on several psychological tests and the severity of depression, the correlation coefficients ranged from 0.15 to 0.46. Only 2 to 21% of the variation in test performance could be attributed to the severity of the depression (Stromgren, 1977). Second, the cognitive deficits seem to be rather specific and only show up on a minority of cognitive tasks. For example, in one investigation in which 82 test scores were reported (33 cognitive, perceptual, and psychomotor tasks), only 8 revealed significant differences between the psychotically depressed patients and normal controls (Friedman, 1964). As noted above, these latter scores primarily reflected deficits in sustained concentration, effortful processing, and psychomotor speed. Many investigators have found that depressed patients are not impaired on typical intelligence tests. Since most of the experimental deficiency studies have used such tests rather than those that have been shown to be sensitive to depression, our contention that nutritionally deficient subjects may not be cognitively normal despite the apparent lack of deficits has some support.

If depression does alter specific cognitive processes, it might be asked why the subjects who experienced nutritional deficiencies (and who were depressed) were not generally found to be cognitively impaired. There are several possible contributing factors. First, as noted above, perhaps the most appropriate tests were not used. This comment pertains to both the type of formal tests chosen and the possible fallacy in extrapolating from cognitive performance in structured test situations to cognitive functioning in daily life. This last aspect may be particularly pertinent in light of the active encouragement during the testing in the experimental deficiency studies. This factor may partly explain why these subjects were not as impaired as some clinically depressed patients (for whom the degree of tester involvement seemed to vary greatly from study to study). Certainly this factor may also explain Brozek's distinction between mental ability and mental activity.

Second, in the experimental deficiency studies the subjects were highly practiced on the cognitive test battery specifically to abolish practice effects. This procedure may well have overcome the cognitive effects of the depression and certainly is in marked contrast to the methodology of the depression studies, in which subjects were not generally practiced at all.

Finally, these experimental subjects were surely not as severely depressed as the patients who participated in the depression studies. However, even if the cognitive changes associated with depression would therefore be expected to be less severe under conditions of nutritional deficiency, these types of cognitive changes would be expected to result.

REFERENCES

Benedict, F.G. 1915. A study of prolonged fasting. Carnegie Institute Publication No. 203. Carnegie Institute, Washington, D.C. 415 pp.

Berryman, G.H., C.R. Henderson, N.C. Wheeler, R.C. Cogwell, Jr., F.R. Spinella, W.E. Grundy, H.C. Johnson, M.E. Wood, C.W. Denko, T.E. Friedemann, C.S. Harris, A.C. Ivy, and J.B. Youmans. 1947. Effect in young men consuming restricted quantities of B-complex vitamins and protein, and changes associated with supplementation. Am. J. Physiol. 148:618-647.

Blanton, S. 1919. Mental and nervous changes in the children of the Volkschulen of Trier, Germany, caused by malnutrition. Mental Hyg. 3:343-386.

Brozek, J., and A. Keys. 1944. General aspects of interdisciplinary research in experimental human biology. Science 100:507-512.

Brozek, J., and A. Keys. 1945. Changes in flicker fusion frequency with age. J. Consult. Psychol. 9:87-90.

Brozek, J., H. Guetzkow, and A. Keys. 1946a. A study of personality of normal young men maintained on restricted intakes of vitamins of the B complex. Psychosomat. Med. 8:98-109.

Brozek, J., H. Guetzkow, O. Michelsen, and A. Keys. 1946b. Motor performance of normal young men maintained on restricted intakes of vitamins of the B complex. J. Appl. Psychol. 30:359-379.

Chernorutskii, M.V. 1943. Problem of alimentary dystrophy. Stud. Leningrad Physicians 3:3-13. Cited in A. Keys, J. Brozek, A. Henschel, O. Mickelson, and H.L. Taylor. 1950. The Biology of Human Starvation, p. 793. University of Minnesota Press, Minneapolis.

Cohen, R.M., H. Weingartner, S. Smallberg, and D.C. Murphy. 1982. Effort in cognitive processes in depression. Arch. Gen. Psychiatry 39:593-597.

Colbert, J., and M. Harrow. 1967. Psychomotor retardation in depressive syndromes. J. Nerv. Mental Dis. 145:405-419.

Cutforth, R.H. 1958. Adult scurvy. Lancet i:454.

Dahlstrom, W.G., and G.S. Welsh. 1960. An MMPI Handbook: A Guide to Use in Clinical Practice and Research. University of Minnesota Press, Minneapolis.

Dols, M.J.L., and D.J.A.M. van Arcken. 1946. Food supply and nutrition in the Netherlands during and immediately after World War II. Milbank Mem. Fund Q. 24:319-355.

Fliederbaum, J., A. Heller, K. Zweibaum, J. Zarchi, S. Szejnfinkel, T. Goliborska, R. Elbinger, and F. Ferszt. 1946. Recherches cliniques et biochimiques sur les malades en famine. Pp. 79-172 in E. Apfelbaum, ed. Maladie de Famine. Recherches Cliniques sur la Famine Executées dans le Ghetto de Varsovie en 1942. American Joint Distribution Committee, Warsaw.

Franklin, J.C., and J. Brozek. 1949. The Rosenzweig P-F test as a measure of frustration response in semi-starvation. J. Consult. Psychol. 13:293-301.

Friedman, A.S. 1964. Minimal effects of severe depression on cognitive functioning. J. Abnorm. Soc. Psychol. 69:237-243.

Glaze, J.A. 1928. Psychological effects of fasting. Am. J. Psychol. 40:236-253.

Gruetzkow, H., and J. Brozek. 1947. Intellective tests for longitudinal experiments on adults. Am. J. Psychol. 60:350-366.

Guilford, J.P., and H.G. Martin. 1943. The Guilford-Martin inventory of factors GAMIN (blanks, and manual of directions and norms). Sheridan Supply Co., Beverly Hills, Calif.

Henderson, C.R., N.C. Wheeler, H.C. Johnson, R.C. Cogswell, G.H. Berryman, A.C. Ivy, T.E. Friedemann, and J.B. Youmans. 1947. Changes in personality appraisal associated with a restricted intake of B vitamins and proteins. Am. J. Med. Sci. 213:488-493.

Henry, G.M., H. Weingartner, and D. Murphy. 1973. Influence of affective states and psychoactive drugs on verbal learning and memory. Am. J. Psychiatry 130:966-971.

Kendrick, D.C., and F. Post. 1967. Influences in cognitive status between healthy, psychiatrically ill, and diffusely brain damaged elderly subjects. Br. J. Psychiatry 113:75-81.

Keys, A., A. Henschel, O. Mickelsen, J. Brozek, and J.H. Crawford. 1944. Physiological and biochemical functions in normal young men on a diet restricted in riboflavin. J. Nutr. 27:165-178.

Keys, A., A. Henschel, H.L. Taylor, O. Mickelsen, and J. Brozek. 1945. Experimental studies on men with a restricted intake of the B vitamins. Am. J. Physiol. 144:5-41.

Keys, A., J. Brozek, A. Henschel, O. Mickelsen, and H.L. Taylor. 1950. The Biology of Human Starvation, University of Minnesota Press, Minneapolis. 1,385 pp.

Kinsman, R.A., and J. Hood. 1971. Some behavioral effects of ascorbic acid deficiency. Am. J. Clin. Nutr. 24:455-464.

Levitsky, D.A., and B.J. Strupp. 1985. Nutrition and the behavior of children. In W. Walter and A. Watkins, eds. Nutrition in Pediatrics: Basic Science and Clinical Application. Little, Brown & Co., Boston.

Markowski, B. 1945. Some experiences of a medical prisoner of war. Br. Med. J. 2:361-366.

Marsh, H.D. 1916. Individual and sex differences brought out by fasting. Psychol. Rev. 23:437-445.

O'Shea, H.E., K. Elsom, and R.O. Higbe. 1942. Studies of B vitamins in human subjects: mental changes in experimental deficiency. Am. J. Med. Sci. 203:388-397.

Sebrell, W.H. 1943. The mental and neurological aspects of vitamin B-complex deficiency. Proc. Assoc. Res. Nerv. Mental Dis. 22:113-121.

Simonson, E., N. Enzer, and S.S. Blankenstein. 1941. The influence of age on fusion frequency of flicker. J. Exp. Psychol. 29:252-255.

Sorokin, P.A. 1942. Man and Society in Calamity. Dutton, New York. 352 pp.

Spies, T.D., J. Bradley, M. Rosenbaum, and J.R. Knott. 1943. Emotional disturbances in persons with pellagra, beri-beri and associated deficiency states. Res. Publ. Assoc. Nerv. Mental Dis. 22:122-140.

Sterner, R.T., and W.R. Price. 1973. Restricted riboflavin: within subject behavioral effects in humans. Am. J. Clin. Nutr. 26:150-160.

Stromgren, L.S. 1977. The influence of depression on memory. Acta Psychiat. Scand. 56:109-128.

Strupp, B.J., and D.A. Levitsky. 1983. Early brain insult and cognition: a comparison of malnutrition and hypothyroidism. Dev. Psychobiol. 16:535-549.

Weingartner, H., R.M. Cohen, D.L. Murphy, J. Martello, and C. Gerdt. 1981. Cognitive processes in depression. Arch. Gen. Psychiatry 38:42-47.

Weygandt, W. 1904. Uber die Beeinflussung geistiger Funktionen durch Hungern. Psychol. Arbeit. 4:45-143. Cited in A. Keys, J. Brozek, A. Henschel, O. Mickelson, and H.L. Taylor. 1950. The Biology of Human Starvation, p. 858. University of Minnesota Press, Minneapolis.

Wilder, J. 1944. Malnutrition and mental deficiency. Nerv. Dis. Child. 3:178-186.

Williams, R.D., H.L. Morgan, R.M. Wilder, and B.F. Smith. 1940. Observations on induced thiamine (vitamin Bl) deficiency in man. Arch. Intern. Med. 66:785.

Wilter, R.W., R.M. Woolford, and T.D. Spies. 1946. Severe scurvy. J. Lab. Clin. Med. 31:609-630.

COGNITIVE EFFECTS OF CHOLINERGIC DRUGS

Myron F. Weiner and Kenneth L. Davis

Recent investigations of cholinergic pathways and function in the central nervous system have shed much light on the physiological basis of learning and memory and have also illuminated certain aspects of Alzheimer's disease.

Acetylcholine (ACh), the neurotransmitter of the cholinergic system, is found in both central and peripheral nervous system tissue. Its muscarinic effects are due to its action at postganglionic parasympathetic nerve endings on smooth muscle, cardiac muscle, and endocrine glands. It is the chemical mediator at all autonomic ganglion synapses, at the motor fiber terminals at skeletal muscle neuromuscular junctions (its nicotinic effect), and at various sites within the central nervous system (Goth, 1984).

Cholinergic pathways have been well delineated in rat brain (Fig. 1), but are less well known in humans because ACh is rapidly hydrolyzed and is therefore not readily detectable in human postmortem specimens. In humans, ACh activity is usually inferred by choline acetyltransferase (CAT) activity (Davies, 1978).

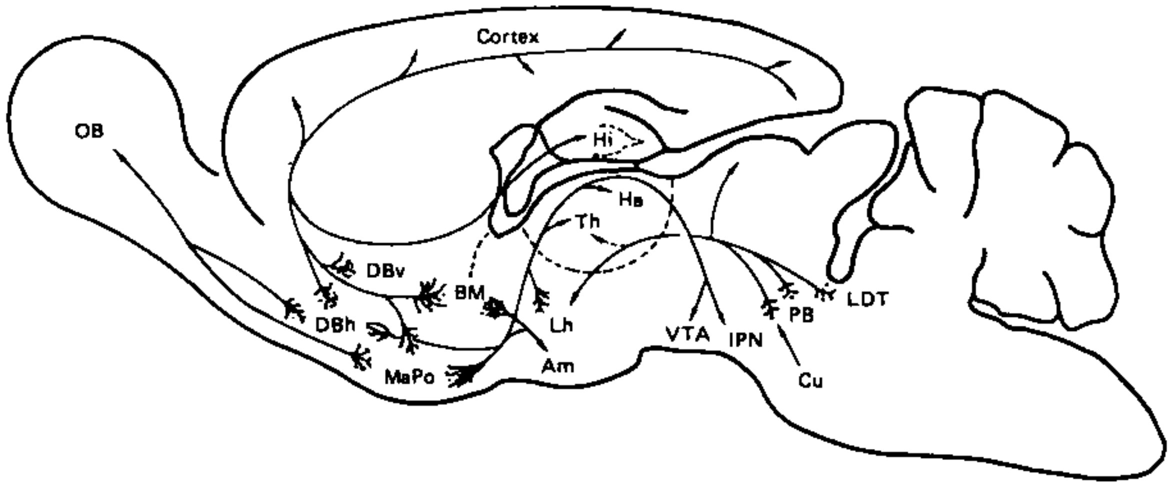
SYNTHESIS AND DEGRADATION

ACh is synthesized by nerve cell bodies and terminals in a reaction catalyzed by CAT (Cooper *et al*., 1982), as shown:

$$\text{acetyl coenzyme A} + \text{choline} \rightleftarrows \text{ACh} + \text{coenzyme A}$$

At nerve terminals, about half the ACh is found in synaptic vesicles and half is found in the synaptosomal cytoplasm. It has been suggested that as nerve cells are depolarized, vesicles near their terminals fuse with the presynaptic cell membrane and ACh is released into the synaptic cleft, where it changes the ion permeability of the postsynaptic receptor. In the synaptic cleft, ACh is rapidly hydrolyzed by the membrane-bound enzyme acetylcholinesterase, as indicated below.

$$\text{ACh} \rightarrow \text{acetate} + \text{choline}$$



Am	amygdala	IPN	interpeduncular nucleus
BM	nucleus basalis magnocellularis	Lh	lateral hypothalamus
Cu	nucleus cuneiformis	LDT	laterodorsal tegmental nucleus
DbH	horizontal limb of the nucleus of the diagonal band	MaPO	magnocellular preoptic nucleus
		OB	olfactory bulb
DBv	vertical limb of the nucleus of the diagonal band	PB	parabrachial nucleus
		Th	thalamus
Ha	habenular nuclei	VTA	ventral tegmental area
Hi	hippocampus		

FIGURE 1. Cholinergic central nervous system tracts. (From Cooper et al., 1982; adapted from Fibiger, 1982.)

The choline produced by this process is available for reuptake into the presynaptic neuron (Seiden and Dykstra, 1977), as indicted in Fig. 2, which illustrates a hypothetical cholinergic terminal. The rate-limiting step in ACh synthesis is the transport of choline into cholinergic terminals by a high-affinity uptake system (Simon and Kuhar, 1975).

Release of ACh from the presynaptic neuron is blocked by botulinum toxin and Mg^{2+}. Central muscarinic receptors are blocked by atropine, scopolamine, and quinuclidinyl benzylate, whose tertiary amine structure allows them to penetrate the blood-brain barrier. Other similarly acting anticholinergics do not pass the blood-brain barrier. Acetylcholinesterase is inhibited centrally by physostigmine (reversibly) or irreversibly by diisopropyl fluorophosphate (Cooper et al., 1982).

AUGMENTATION OF CHOLINERGIC FUNCTION

Cholinergic transmission can be affected presynaptically, postsynaptically, or in the synaptic cleft itself. The synthesis and

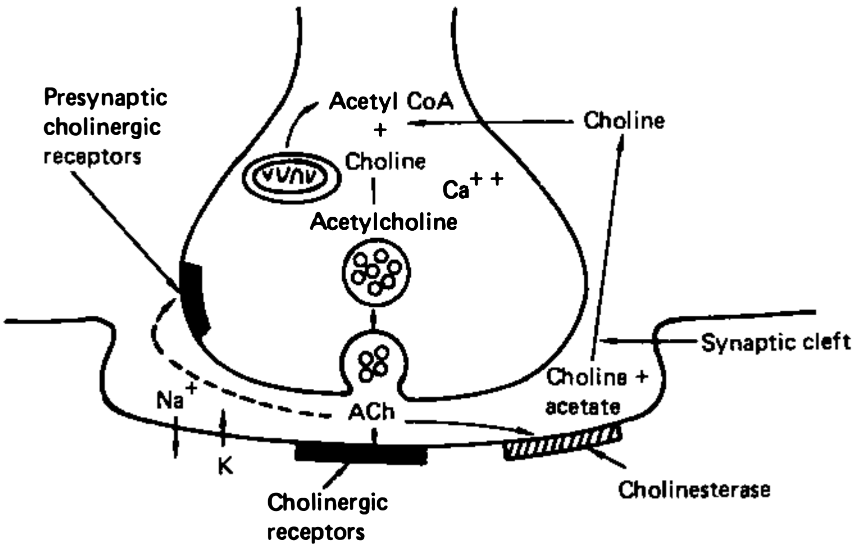


FIGURE 2. Cholinergic nerve terminal. Depicts the synthesis, storage, and release of ACh, hydrolysis by cholinesterase, and action on effector cell and presynaptic cholinergic receptors. (From Goth, 1984.)

release of ACh may be facilitated presynaptically under certain conditions by supplying precursors, such as choline and phosphatidyl choline (lecithin) (Cohen and Wurtman, 1976; Wecker et al., 1978), and releasers, such as aminopyridine (Glover, 1982). Postsynaptically, direct receptor antagonists such as arecoline and oxotremorine can be employed. The most effective mechanism is binding of cholinesterase, which blocks enzymatic degradation of ACh. Physostigmine and tetrahydroaminoacridine (THA) are used for this purpose. Physostigmine has an affinity for cholinesterases that may be 10,000 times greater than that of ACh (Goth, 1984). After combining with these enzymes, it gradually becomes dissociated and is inactivated.

The effects of physostigmine on mental functioning are complex. They are strongly dose related and are influenced by effects on other organ systems. To block the peripheral effects of physotigmine, methscopolamine bromide (which does not cross the blood-brain barrier) is usually administered in a dose of 0.5 to 1.0 mg subcutaneously before physostigmine is administered (Davis et al., 1976).

Administered slowly intravenously or given intramuscularly, physostigmine is a specific antidote for the central and peripheral effects of anticholinergic intoxication (Duvoisin and Katz, 1968). The experiments of Ketchum et al. (1973) suggest doses of about 1 mg for every 3 mg of atropine or 0.33 mg of scopolamine. THA is also effective in doses of 0.5 to 1.5 mg/kg intramuscularly or intravenously (Granacher and Baldessarini, 1975).

ANTICHOLINERGIC DRUGS AND MEMORY

Anticholinergic drugs have been known for centuries to produce delirium, whether taken accidentally (Morton, 1939), as part of religious rites (Benedict, 1959), or recreationally (Goldfrank and Meliek, 1979). In the early years of this century, the specific effects of centrally acting anticholinergics on memory were noted, and in 1906 scopolamine was combined with morphine to produce the twilight sleep of obstetrical anesthesia (Gauss, 1906). More recently, closer attention has been paid to the effects of scopolamine on different aspects of memory that seem distinct from delirium or generalized amnesia. In general, the findings (which are supported by studies in nonhuman primates [Bartus, 1979]) are specific impairment of memory by low doses, with impairment of many aspects of cognitive function at higher doses.

Scopolamine impairs mildly the retention of new information (digit span) but markedly impairs delayed (20-s) recall, even after two rehearsals, at a dose of 10 μg/kg (Safer and Allen, 1971). The effect was not a function of drowsiness, which occurred at low doses, and was supervened by excitement at higher doses. Visual hallucinations occurred in these subjects, who were healthy young men, at total doses as low as 0.53 mg. At a total dose of 0.89 mg, more than half the subjects were hallucinating.

It was later found that small doses of scopolamine (beginning at 5 μg/kg) and atropine (beginning at 32 μg/kg) impaired immediate recall and slightly delayed recall (Ketchum _et al_., 1973). Many subjects in this later study could not repeat short sentences or number sequences. The authors also noted impaired attention, loss of abstracting ability, a tendency to lose track of time, and little awareness by subjects of their degree of impairment. Instead of acknowledging their deficits, the subjects found excuses for their poor performance. Hallucinations, disorientation, and incoherence appeared at a 152-μg/kg dose of atropine and a 20-μg/kg dose of scopolamine. There was a regular progression of effects at these doses. At first, tachycardia and dry mouth were most conspicuous. Later, there was somnolence, restlessness, ataxia, incoordination, hyperreflexia, hyperthermia, and hypertension. Overlapping this stage were disruption of awareness and inability to pay attention, to carry out instructions, to speak coherently, or to interpret stimuli realistically. These effects lasted 6 to 8 h after scopolamine and 10 to 12 h after atropine administration.

Other studies have consistently shown that immediate recall is little impaired (Atkinson and Shiffrin, 1971; Ghoneim and Mewaldt, 1975), but when new material exceeding the capacity of short-term memory is to be learned, performance is markedly impaired (Drachman and Leavitt, 1974; Ghoneim and Mewaldt, 1975). Therefore, the ability to store new information in long-term memory is most disturbed by a decrease in cholinergic activity (Mohs and Davis, 1983). Using large doses of scopolamine (1 mg subcutaneously) accompanied by equal doses of subcutaneously administered methscopolamine, Drachman and Leavitt (1974) also demonstrated that the memory impairment produced in healthy young adults resembles that of mildly memory-impaired elderly subjects: a reduced ability to store new material in long-term memory, with relative sparing of recall and retrieval from long-term memory. Drachman (1978) later showed that drug effects were not due to sedation and were relatively specific to cholinergic transmission. They were reversed by physostigmine, but not by D-amphetamine.

RELATIONSHIP TO ALZHEIMER'S DISEASE

Alzheimer's disease is a dementing illness of unknown etiology, which may occur as early as the fifth decade of life. It does not appear with great frequency until the seventh or eighth decade, during which time the overall risk of dementia is about 1% per year of age (Mortimer et al., 1981a). It is estimated that a person living to 85 years of age has a 1-in-5 risk of severe dementia (Mortimer et al., 1981b). Alzheimer's disease accounts for at least half of all cases of dementia in the latter years (Tomlinson et al., 1970).

Although Alzheimer's disease was formerly thought to be a concomitant of normal aging, it has become apparent that the dementia of this disease is not the result of generalized brain deterioration, but of the loss of specific neurons and neurotransmitters, especially cholinergic transmitters (Bowen and Davison, 1974).

Postmortem studies have shown that Alzheimer's disease patients have greatly reduced CAT activity in the hippocampus and cortex compared with age-matched controls (Davies and Maloney, 1976). A correlation between decreased CAT activity, number of senile plaques, and severity of dementia has also been established (Perry et al., 1978). Reduced spinal fluid ACh has also been demonstrated in Alzheimer's disease patients, the reduction in ACh corresponding to the degree of dementia (Davis et al., 1982b).

An important breakthrough came with the discovery that there is substantial loss of cholinergic cells in the nucleus basalis of Meynert, a cluster of cells in the substantia innominata of the forebrain that projects fibers diffusely to the neocortex and hippocampus (Whitehouse et al., 1981). In addition, it has been found that brain receptors for ACh are normal (White et al., 1977) or only slightly reduced (Reisine et al., 1978) in number. The finding of a dramatic and selective loss of cholinergic cells in patients with Alzheimer's disease together with the known effects of anticholinergic drugs on memory have prompted investigations of the extent to which anticholinergic drugs mimic in humans all of the symptoms of Alzheimer's disease. One recent study (Mohs et al., in press) indicates that, although 0.43 mg of scopolamine substantially impairs storage of information into long-term memory, it has no effect on word-finding ability or constructional praxis, two aspects of cognition known to be impaired in early Alzheimer's disease (Rosen, 1980).

The superficial resemblance of cholinergic dysfunction in Alzheimer's disease to dopaminergic dysfunction in Parkinson's disease led to the suggestion that Alzheimer's disease might be treated by replacing the missing neurotransmitter or its precursor. Although studies have shown that increases in dietary choline or lecithin increase ACh levels in the brain (Haubrich et al., 1975), choline or lecithin administration does not affect memory in young adults, nondemented elders, or persons with Alzheimer's disease (Johns et al., 1983). However, pretreatment with 8 g of choline orally caused a small decrease of scopolamine-induced impairment of a verbal learning task when 0.43 mg of scopolamine was administered subcutaneously (Mohs et al., 1981).

Reasoning that precursors might only increase ACh in presynaptic neurons, it was thought that choline might be more helpful if it was added to a substance that facilitates ACh release, such as piracetam (Wurtman et al., 1981). When lecithin was combined with piracetam, neglible improvement in memory occurred (Friedman et al., 1981).

MEMORY ENHANCEMENT BY CHOLINESTERASE INHIBITION

The cholinomimetic drug physostigmine, in addition to reversing the memory impairment produced by anticholinergics (Granacher and Baldessarini, 1975), also affects memory function in drug-free humans and laboratory animals. The effects are not linear, however, and have been shown in rats to depend on the interval between learning and drug administration and retrieval (Deutsch, 1971) and on finding an optimal dose within a very narrow range (Bartus, 1979).

When normal human subjects were administered doses of 1.5 to 2.0 mg of physostigmine intravenously over 20 min, they became lethargic and expressionless, had decreased spontaneous movements and speech, and experienced difficulty in thinking (Davis et al., 1976). On testing, impairment of short-term memory (measured by digit span forward and backward) was evident compared with that of saline-infused controls. These effects disappeared 90 to 120 min after cessation of physostigmine infusion. No difference was observed between the physostigmine-treated group and the controls on tests of consolidation of short-term memory. In spite of having received prophylactic methscopolamine, almost all subjects became nauseated, and several vomited. Davis et al. (1982a) found that methscopolamine blocked the response of growth hormone to this generalized stress response, but did not block cortisol or vasopressin responses. Of additional interest is that two of the subjects also became depressed and tearful.

Risch *et al*. (1981) administered 0.022 μg of physostigmine per kg (1.5 mg/70 kg after preliminary methscopolamine treatment) by intravenous infusion over 10 min and found that the dose produced profound dysphoria in hospitalized depressives. Doses of 0.033 μg/kg (2.3 mg per 70-kg subject) similarly administered to normals also produced dysphoria, but it was less intense than that experienced by the depressives.

When Davis *et al*. (1978) reduced the test dose of physostigmine to 1.0 mg infused intravenously over 30 min (after pretreatment with methscopolamine, 0.5 mg), the outcome was quite different. Physostigmine significantly enhanced performance on a learning task (recall of words presented 30 min after the infusion began) that involved storing information in long-term memory, marginally increased retrieval from long-term memory (recall of words presented 30 min before infusion), and had no effect on short-term memory (digit span and memory-scanning task). Although no subject showed decreased performance after receiving physostigmine, there was great variability of response between subjects. Thus, in humans, enhancement of memory by physostigmine occurs at low doses but is not uniform.

Sitaram *et al*. (1978) found a similar response to the muscarinic antagonist arecoline in healthy young adults. Using a 4-mg intramuscular dose after pretreatment with methscopolamine, they found improvement in categorized serial learning (learning a fixed sequence of 10 words belonging to a familiar category such as vegetables or fruits). The improvement was most marked in persons who had previously shown the poorest performance. This group was also the most sensitive to impairment by scopolamine. It was found that 6 mg of arecoline was required to reverse the effects of 0.5 mg of scopolamine, suggesting that arecoline is about one-fourth as potent as physostigmine in reversing the central antimuscarinic effects of scopolamine.

There are significant problems with administering physostigmine and arecoline. They are so rapidly metabolized that at present they must be given by intravenous infusion or in frequent doses by mouth. In addition, the memory-enhancing dose is very close to the dose at which confusion and lethargy appear.

SUMMARY AND CONCLUSIONS

The cholinergic system is but one of many transmitter systems in the central nervous system. Studies of central nervous system cholinergic blockade suggest that the cholinergic system may be important in the regulation of mood, in attention, and in learning and remembering. Cholinomimetic drugs block the central effects of anticholinergics. In normal persons, they enhance memory at low doses, and at high doses lethargy, dysphoria, and impaired memory occur.

Because of the finding of reduced cholinergic function in Alzheimer's disease, various strategies have been used to increase cholinergic function, including the use of ACh precursors, muscarinic antagonists, acetylcholinsterase inhibitors, and metabolic enhancers. Some positive effects have been noted with acetylcholinesterase inhibitors.

It seems clear that the cholinergic system is not the only system involved in the encoding, storage, and recall of information. For example, vasopressin may have slight ameliorating effects on the memory of Alzheimer's disease patients (Legros et al., 1978). On the other side of the coin, Solomon et al. (1983) have shown that memory impairment occurs with the use of methyldopa and propanolol in the treatment of essential hypertension. In that study, there was impairment of immediate and delayed recall of verbal memory items, but visual memory was relatively unaffected. Thus, there are clear signs that systems other than the cholinergic system will require investigation if we are to better understand the mechanisms underlying learning and memory.

REFERENCES

Atkinson, R.C., and R.M. Shiffrin. 1971. The control of short-term memory. Sci. Am. 225:82-90.

Bartus, R.T. 1979. Physostigmine and recent memory: effects on young and aged nonhuman primates. Science 206:1087-1089.

Benedict, R. 1959. Patterns of Culture. Pp. 85-86. Houghton Mifflin, Boston.

Bowen, D.M., and A.N. Davison. 1974. Biochemical changes in the cholinergic system of the aging brain and in senile dementia. Psychol. Med. 10:315-318.

Cohen, E.L., and R.J. Wurtman. 1976. Brain acetylcholine: control by dietary choline. Science 191:561-562.

Cooper, J.R., F. Bloom, and R.H. Roth. 1982. The Biochemical Basis of Neuropharmacology, 4th Ed. Oxford University Press, New York.

Davies, P. 1978. Loss of choline acetyltransferase activity in normal aging and in senile dementia. Adv. Exp. Med. Biol. 113:251-256.

Davies, P., and A.J.F. Maloney. 1976. Selective loss of central cholinergic neurons in Alzheimer's disease. Lancet ii:1403.

Davies, P., and R.D. Terry. 1981. Cortical somatostatin-like immunoreactivity in cases of Alzheimer's disease and senile dementia of the Alzheimer type. Neurobiol. Aging 2:9-14.

Davis, B.M., and K.L. Davis. 1979. Acetylcholine and anterior pituitary hormone secretion. Pp. 445-448 in K.L. Davis and P.A. Berger, eds. Brain Acetylcholine and Neuropsychiatric Disease. Plenum Publishing Corp., New York.

Davis, K.L., L.E. Hollister, J. Overall, A. Johnson, and K. Train. 1976. Physostigmine: effects on cognition and affect in normal subjects. Psychopharmacology 51:23-27.

Davis, K.L., R.C. Mohs, J.R. Tinklenberg, A. Pfefferbaum, L.E. Hollister, and B.S. Kopell. 1978. Physostigmine improvement of long-term memory processes in normal humans. Science 201:272-274.

Davis, B.M., G.M. Brown, M. Miller, H.G. Friesen, A.J. Kastin, and K.L. Davis. 1982a. Effects of cholinergic stimulation on pituitary hormone release. Psychoneuroendocrinology 7:347-354.

Davis, K.L., J.Y.-K. Hsieh, M.I. Levy, T.B. Horvath, B.M. Davis, and R.C. Mohs. 1982b. Cerebrospinal fluid acetycholine, choline and senile dementia of the Alzheimer type. Psychopharmacol. Bull. 18:193-195.

Deutsch, J.A. 1971. The cholinergic synapse and the site of memory. Science 174:788-794.

Drachman, D.A. 1978. Central cholinergic system and memory. Pp. 651-662 in M.A. Lipton et al., eds. Psychopharmacology: A Generation of Progress. Raven Press, New York.

Drachman, D.A., and J. Leavitt. 1974. Human memory and the cholinergic system. Arch. Neurol. 30:113-121.

Drachman, D.A., G. Glosser, P. Fleming, and G. Longenecker. 1982. Memory decline in the aged: treatment with lecithin and physostigmine. Neurology (New York) 32:944-950.

Duvoisin, R.C., and R. Katz. 1968. Reversal of central anticholinergic syndrome in man by physostigmine. J. Am. Med. Assoc. 206:1963-1965.

Ferris, S.H., B. Reisberg, T. Crook, E. Friedman, M.K. Schneck, P. Mir, K.A. Sherman, J. Corwin, S. Gershon, and R.T. Bartus. 1982. Pharmacologic treatment of senile dementia: choline, L-Dopa, piracetam, and choline plus piracetam. Pp. 281-286 in S. Corkin et al., eds. Alzheimer's Disease: A Report of Progress. Raven Press, New York.

Fibiger, H.C. 1982. Organization and some projections of cholinergic neurons in the mammalian forebrain. Brain Res. Rev. 4:327-388.

Friedman, E., K.A. Sherman, S.H. Ferris, B. Reisberg, R.T. Bartus, and M.K. Schneck. 1981. Clinical response to choline plus piracetam in senile dementia: relation to red cell choline levels. N. Engl. J. Med. 304:1490-1491.

Gauss, C.J. 1906. Geburten in Kuntslichem Dammerschlaff. Arch. Gynaekol. 78:579-631.

Ghoneim, M.M., and S.P. Mewaldt. 1975. Effects of diazepam and scopolamine on storage, retrieval and organizational processes in memory. Psychopharmacologia (Berlin) 44:257-262.

Giurgea, C., M.G. Greindl, S. Preat, and J. Puigdevall. 1982. Piracetam compensation of MAM-induced behavioral deficit in rats. Pp. 281-286 in S. Cohen et al., eds. Alzheimer's Disease: A Report of Progress. Raven Press, New York.

Glover, W.E. 1982. The aminopyridines. Gen. Pharmacol. 13:259-285.

Goldfrank, L., and M. Meliek. 1979. Locoweed and other anticholinergics. Hosp. Physician 8:13-39.

Goth, A. 1984. Medical Pharmacology. C.V. Mosby, St. Louis.

Granadcher, R.P., and R.J. Baldessarini. 1975. Physostigmine: its use in acute anticholinergic syndrome and with antidepressant and antiparkinson drugs. Arch. Gen. Psychiatry 32:375-380.

Grancher, R.P., and R.J. Baldessarini. 1976. The usefulness of physostigmine in neurology and psychiatry. Pp. 63-79 in H.L. Klawans, ed. Clinical Neuropharmacology. Vol. 1. Raven Press, New York.

Haubrich, D.R., P.F.L. Wang, D.E. Clody, and P.W. Wedeking. 1975. Increase in rat brain acetylcholine induced by choline or deanol. Life Sci. 17:975-980.

Johns, C.A., B.S. Greenweld, R.C. Mohs, and K.L. Davis. 1983. The cholinergic treatment strategies in aging and senile dementia. Psychopharmacol. Bull. 19:185-197.

Ketchum, J.S., F.R. Sidell, E.B. Crowell Jr., G.K. Aghajanian, and A.H. Hayes Jr. 1973. Atropine and scopolamine and ditran: comparative pharmacology and antagonists in man. Psychopharmacologia (Berlin) 28:121-145.

Legros, J.J., P. Gilot, T. Seron, J. Claessens, A. Adams, J.M. Moeglen, A. Audibert, and P. Berchier. 1978. Influence of vasopressin on learning and memory. Lancet i:41-42.

Mohs, R.C., and K.L. Davis. 1979. Cholinomimetic drug effects on memory in young and elderly adults. Pp. 154-158 in A.I.M. Glen and L.J. Whalley, eds. Alzheimer's Disease. Churchill-Livingstone, Edinburgh.

Mohs, R.C., and K.L. Davis. 1983. Psychopharmacology, neurotransmitter systems, and cognition in the elderly. Pp. 305-316 in R.F. Walker and R.L. Cooper, eds. Experimental and Clinical Interventions in Aging. Marcel Dekker, New York.

Mohs, R.C., K.L. Davis, J.R. Tinklenberg, L.E. Hollister, J.A. Yesavage, and B.S. Kopell. 1979. Choline chloride treatment of memory deficits in the aged. Am. J. Psychiatry 136:1275-1277.

Mohs, R.C., K.L. Davis, and M.I. Levy. 1981. Partial reversal of anticholinergic amnesia by choline chloride. Life Sci. 29:1317-1323.

Mohs, R.C., Y. Kim, C.A. Johns, D.D. Dunn, and K.L. Davis. In press. Assessing change in Alzheimer's disease: memory and language tests. In L.W. Poon, ed. Clinical Assessment of Older Adults. American Psychological Association, Washington, D.C.

Mortimer, J.A., L.M. Schuman, and L.R. French. 1981a. Epidemiology of dementing illness. Pp. 3-23 in J.A. Mortimer and L.M. Schuman, eds. The Epidemiology of Dementia. Oxford University Press, New York.

Mortimer, J.A., L.M. Schuman, and L.R. French. 1981b. Epidemiology of dementia: overview and prospects. Pp. 173-182 in J.A. Mortimer and L.M. Schuman, eds. The Epidemiology of Dementia. Oxford University Press, New York.

Morton, H.G. 1939. Atropine intoxication: its manifestation in infants and children. J. Pediatr. 14:755-760.

Perry, E.K., B.E. Tomlinson, G. Blessed, K. Borgmann, P.H. Gibson, and R.H. Perry. 1978. Correlation of cholinergic abnormalities with senile plaques and mental test scores in senile dementia. Br. Med. J. 2:1457-1459.

Reisberg, B., S.H. Ferris, and S. Gershon. 1981. An overview of pharmacologic treatment of cognitive decline in the aged. Am. J. Psychiatry 138:593-600.

Reisine, T.D., H.I. Yamamura, E.D. Bird, E. Spokes, and S.J. Enna. 1978. Pre- and postsynaptic neurochemical alterations in Alzheimer's disease. Brain Res. 159:477-481.

Risch, S.C., N.H. Kalin, and D.S. Janowsky. 1981. Cholinergic challenges in affective illness: behavioral and neuroendocrine correlates. J. Clin. Psychopharmacol. 1:186-192.

Rosen, W.G. 1980. Verbal fluency in aging and dementia. J. Clin. Neuropsychol. 2:135-146.

Safer, D.J., and R.P. Allen. 1971. The central effects of scopolamine in man. Biol. Psychiatry 3:347-355.

Sagales, T., S. Erill, and E. Domino. 1969. Differential effects of scopolamine and chlorpromazine in NREM sleep in normal male subjects. Clin. Pharmacol. Ther. 10:522-529.

Seiden, L.S., and L.A. Dykstra. 1977. Psychopharmacology: A Biochemical and Behavioral Approach. Van Nostrand Reinhold, New York.

Simon, J.R., and M.T. Kuhar. 1975. Impulse-flow regulation of high affinity choline uptake in brain cholinergic nerve terminals. Nature 255:162-163.

Sitaram, N., H. Weingartner, and J.C. Gillin. 1978. Human serial learning: enhancement with arecoline and impairment with scopolamine correlated with performance on placebo. Science 201:274-276.

Solomon, S., E. Hotchkiss, S. Saravay, C. Bayer, P. Ramsey, and R. Blum. 1983. Impairment of memory function by antihypertensive medication. Arch. Gen. Psychiatry 40:1109-1117.

Tomlinson, B.E., G. Blessed, and M. Roth. 1970. Observations in the brains of demented old people. J. Neurol. Sci. 11:205-242.

Wecker, L., W.D. Dettbarn, and D.E. Schmidt. 1978. Choline administration: modification of the central actions of atropine. Science 199:86-87.

White, P., C.R. Hiley, J.J. Goodhardt, L.H. Carrasco, J.P. Keet, I.E.I. Williams, and D.M. Bowen. 1977. Neocortical cholinergic neurons in elderly people. Lancet i:668-671.

Whitehouse, P.J., D.L. Price, R.G. Struble, A.N. Clark, J.T. Coyle, and M.R. DeLong. 1981. Alzheimer's disease and senile dementia: loss of neurons in the basal forebrain. Science 215:1237-1239.

Wurtman, R.J., S.G. Magil, and D.K. Reinstein. 1981. Piracetam diminishes acetylcholine levels in rats. Life Sci. 28:1091-1093.

COGNITIVE EFFECTS OF NOISE, HEAT, AND OTHER STRESSORS

Andrew Smith

One way of approaching the topic of the cognitive effects of noise, heat, and other stressors would be to draw up a list of studies that have examined the effects of various stressors on different aspects of cognitive performance and try to define the safe level of exposure to each stressor. This atheoretical approach has typified the study of some stressors, and many of the early results have been summarized by Poulton (1970). This approach will now be illustrated by considering the effects of heat on cognitive performance.

HEAT AND COGNITIVE PERFORMANCE

Before describing some experimental results, I must point out a few obvious problems. What should be considered cognitive performance? It is often difficult to separate cognitive processing from other types of processing. This problem often arises in examining the effects of heat on performance. A general conclusion from early heat research is that extremes of hot or cold have a large influence on motor performance. This in turn leads to an effect of these conditions on cognitive performance, due to the subjects' making motor responses in the cognitive tasks.

Several studies have been concerned with defining tolerable temperatures for performing cognitive tasks. Sharma et al. (1983) found that above 32°C, mental alertness, associative learning, reasoning, and dual-task efficiency were impaired. However, some studies have shown that different tasks often show maximum impairment at different temperatures (Wyon et al., 1979). Defining tolerable limits is of obvious practical value, but it is difficult to compare different stressors because we do not know what level of one stressor is equivalent to a certain level of another (e.g., how many decibels are equivalent to 32°C?), and this means that comparison will always be imprecise.

We also know that specific features of a stressor are likely to be crucial. Researchers have tried to determine whether the dry air temperature or the humidity level is more important. Experiments have produced conflicting results on this issue, with Wing and Touchstone (1965) showing that a change in dry air temperature alone can impair recall and Allen and Fischer (1978) finding that an increase in humidity is necessary to impair paired-associate learning.

The length of exposure to heat has also been shown to be important in determining performance effects. Short exposures to hot conditions have often been considered arousing, whereas longer exposures decrease arousal. Hancock (1982) showed that there was an earlier heat stress-related decrement in tasks that required a greater response complexity. However, many of the effects of moderate heat stress have been shown to be variable, and this has made it difficult to define tolerable limits. Wyon *et al*. (1979) have also shown that heat-induced impairment can often be counteracted by conscious effort.

This brief description of some of the effects of heat shows that there are problems with the atheoretical approach. In the next section I will try to provide a detailed profile of the effects of another stressor, noise, and show that this approach has more to offer than one that merely aims to define the boundaries within which a particular stressor has little effect on performance.

EFFECTS OF MODERATE-INTENSITY NOISE ON PERFORMANCE

This review is limited to the effects of moderate-intensity noise (70 to 90 dBA) on cognitive performance. Regular exposure to intense noise should be avoided because of the risk to hearing, and there is little point in considering the effects of these high levels on performance. In general, moderate-intensity noise has little effect on sensory-motor tasks (see Broadbent, 1979), and interest has recently focused on the effects of this noise level on cognitive performance. Most of the studies I will mention were laboratory studies of short duration. However, the effects of prolonged exposure have been studied in real-life settings, and an excellent example of this approach is the study by Cohen *et al*. (1980) on the effects of aircraft noise on schoolchildren.

The majority of these experiments have used continuous white noise. The nature of the noise is obviously very important, as can be seen in a study comparing the effects of different types of noise on working memory and semantic memory tasks (Smith, 1985a). Conglomerate, meaningful noise impaired both tasks, intermittent white noise impaired the semantic memory task but not the working memory task, and continuous white noise had no effect on either task. It has also been shown that speech sounds produce much greater impairment of recall of visually presented words than meaningless white noise does (Salame and Baddeley, 1982). This effect does not depend on the meaning of the speech, and it occurs even when speech sounds in another language are used (Colle, 1980).

One of the obvious effects of noise is that it masks auditory signals. In cognitive tasks subjects often use subvocal rehearsal, and it has been suggested that noise may influence performance by interfering with this subvocal articulation. However, results from experiments by Wilding and Mohindra (1980) argue against the idea that noise interferes with internal speech. They required subjects to memorize a list of letters and found that noise improved recall, but only if the letters were acoustically similar and only if internal speech was allowed (in one condition, subvocal articulation was suppressed by having the subject repeat the word "the" over and over). If noise had masked internal speech, a harmful effect on memory would have been expected, and this should then have disappeared when articulatory suppression was added. Millar (1979) also studied the effects of noise and articulatory suppression, and his results do not support the idea that noise masks internal speech.

Jones et al. (1979) studied the effects of noise on the Bakan vigilance task. In this task single digits are presented on a screen, and subjects have to detect particular sequences, such as three successive odd digits. It has been shown that subjects say the digits to themselves, and if noise masks internal speech, this task should be impaired by noise. Such a result was obtained. However, the exact nature of the noise effect depended on specific features of the experiment, and this made it difficult to interpret the results solely in terms of masking. In other experiments with the Bakan task, no differential effects of noise were found for items that were rehearsed either together or in separate groups. This also shows that noise does not interfere with subvocal articulation. These results may be summarized by saying that although tasks involving internal speech are susceptible to the effects of noise, there are often improvements rather than deteriorations, and the exact pattern of results appears to depend on the specific features of the experiment.

There is evidence that noise may impair comprehension. Hockey (1979) found that persons subjected to loud noise had better recall of names from a passage they had just read but poorer comprehension. Smith (cited in Weiskrantz, 1982) gave subjects the cognitive failures questionnaire (Broadbent et al., 1982) and asked them whether any cognitive failures were likely to increase in frequency in the presence of noise. Eighty-two percent of the subjects reported that "reading something but failing to retain the meaning" was a cognitive failure that was increased by noise. This result received support from an experimental study of proofreading under noisy conditions (Jones and Broadbent, 1979). They found that subjects reading a message out loud in the presence of noise had difficulty understanding and recalling the text. There is other evidence that semantic processing may be impaired by noise. Studies of clustering in free recall have shown that persons subjected to noise often make less use of semantic relationships to aid recall than do subjects in quiet conditions (Daee and Wilding, 1977; Hormann and Osterkamp, 1966; Smith et al., 1981).

Daee and Wilding (1977) found that subjects in noisy conditions often relied more on order information as an aid to recall than on semantic relationships. Hamilton (1972) also found that noise aided recall involving order information, and Dornic (1975) suggested that noise leads to increased use of a lower-level storage mechanism that depends on order information. Schwartz (personal communication) found that noise improved recall of phonemically related words but had no effect on recall of semantically related material. In another experiment he found that noise had a greater detrimental effect on the recall of normal sentences (assumed to contain a large amount of semantic information) than on the recall of anomalous strings, random words, and anagram strings. Weinstein (1974) found that subjects checking visual text for spelling and grammatical errors were as good at detecting spelling errors in noise as in quiet, but deteriorated on the grammatical errors in noise.

This levels-of-processing view of the effects of noise on memory can be criticized on three points. First, direct tests involving manipulation of the encoding processes have shown no interaction between noise and encoding conditions (Smith and Broadbent, 1981). Second, there are studies showing that noise has no effect in direct tests of semantic processing (Eysenck and Eysenck, 1979; Smith, 1985a; Wilding and Mohindra, 1983), and other studies have shown that noise produces more-organized recall of categorized lists (Smith et al., 1981). Third, there are results showing that recall of order information is not always better in noise than in quiet (e.g., Smith, 1983a; Wilding and Mohindra, 1980). Overall, these results argue that noise effects cannot be generally explained in terms of a passive shift in the level of processing.

Broadbent (1971) suggested that noise increases the probability of sampling information from dominant or high-probability sources at the expense of nondominant ones. This enhanced selectivity has been demonstrated using memory tasks. Hockey and Hamilton (1970) examined the effects of noise on intentional and incidental learning. Noise improved recall of order information (the intentional task) but impaired recall of location (the incidental task). Smith (1982) has shown that a similar effect can be obtained by using high- and low-priority tasks rather than intentional and incidental learning. Eysenck (1975) found that noise facilitated retrieval of dominant instances of categories but increased the difficulty of retrieving nondominant instances.

Unfortunately, the effects of noise on selectivity in memory and attention have also been shown to be variable (Smith, 1982; Smith and Broadbent, 1982) and to depend on specific features of the experiment. How can this variability in the effects of noise on cognitive performance be explained? One possibility is that noise effects are easily modified by factors such as the time of testing or the personality of the subject. Loeb et al. (1982) examined the effects of noise on the speed of self-paced mental arithmetic. In male subjects, the speed of work was reduced by noise in the morning, but this effect was reversed in the afternoon. When the data for female subjects were added to the analysis, the clear effect of noise found with the men was eliminated. Smith (in press) has shown that such interactions between noise, time of day, and sex of subject occur but that the exact pattern depends on the tasks used, and the interactions are also unreliable.

Many of the tasks that have shown effects of moderate-intensity noise used verbal materials, and it was therefore initially attractive to think in terms of an effect of noise on internal speech. However, verbal tasks often present the subject with several methods of performing them.

If noise influences specific processes or stages of processes, certain results will be obtained when these processes are used a great deal and others will be obtained when they are not used or are used to a lesser extent. Hence, noise effects may be thought of as hidden defects that are revealed by some experimental techniques but not others. This view will be supported by considering the effects of noise and memory load on running memory.

Hamilton et al. (1977) suggested that information is processed faster but short-term storage is reduced under conditions of noise. This view was supported by the results of a running memory experiment in which it was found that noise improved recall of the most recent items but impaired recall of items earlier in the list. In the running memory task the noise-induced defect is reduced storage capacity. This suggests that the effects of noise on running memory will depend on the number of items that have to be recalled, and indeed, Smith (1983b) has shown that the results of Hamilton et al. (1977) are only obtained in high-memory-load conditions.

Smith (1982) has put forward an alternative view of the effects of noise on performance. When subjects carry out a task that can be performed in several different ways, noise may lead to the adoption of certain strategies in preference to others. Daee and Wilding (1977) showed that subjects in noisy conditions may adopt a maintenance rehearsal strategy rather than an elaboration strategy. There is also evidence that subjects in noise may adopt different recall strategies than subjects in quiet. Smith et al. (1981) found that in certain experiments noise reduced the amount of clustering in free recall due to recall under conditions of noise consisting of fewer words in initial clusters and greater subsequent recall of individual words. Smith (1985b) has shown that noise may influence which aspects of complex stimuli are successfully recalled. In quiet, subjects showed a bias in favor of recalling global features, whereas subjects in noise showed a preference for local detail.

In many tasks it is obvious that one strategy should be used in preference to others. Recent work by Smith (1982) and Wilding et al. (1982) has shown that noise may reinforce the use of the dominant strategy. Smith (1982) has argued that in noise the allocation of effort moves towards the operation that appears to best repay the investment of more effort. Verbal memory tasks are particularly sensitive to noise effects because they offer a variety of strategies, and shifts of dominance or preference can occur more easily than in perceptual motor tasks.

Other effects of noise may be due to changes in the efficiency of control. Rabbitt (1979) has shown that the effects of high-intensity noise on the five-choice serial reaction time task (an increase in errors or gaps but little effect on the average reaction time) can be explained in terms of the noise producing inefficient control of the speed-error trade-off function. Inefficient control can also be seen in the noise-induced bias in resource allocation. Smith (1983c) examined resource allocation under conditions of noise by studying performance on pairs of tasks. The difficulty of each task was varied, as were the priority and probability of having to perform the task. In one experiment, a cognitive vigilance task and a running memory task were used. When there were no priority instructions, noise impaired the vigilance task. However, when the vigilance task was given high priority, the impairment disappeared. Other noise effects were also altered by changes in task parameters. However, this result was only found when the 2 tasks competed for common resources and it did not occur when one of the tasks could be performed in a passive, automatic way.

The effects of noise on the efficiency of the control processes may reflect initial coping with the task and disappear when practice shows the best method. However, the effects of noise may be long lasting and may continue even when the noise is stopped (see Cohen, 1980). Studies by Dornic and Fernaeus (1981) and Smith (1985c) have shown that once a strategy is chosen, subjects in noise may continue with this mode of response even though it is not always appropriate. This suggests that subjects in noise can be considered rather inflexible, and further studies are required to see how noise influences their ability to adapt to rapidly changing tasks.

A major effect of noise that has not yet been mentioned is on motivation. Noise reduces subjects' tolerance for frustration (Glass and Singer, 1972), reduces their level of aspiration (Krenauer and Schonpflug, 1980), regulates their effort (Schulz and Schonpflug, 1982) and influences their perception of competence (Jones, 1983). The reader is referred to Schonpflug (1983) for an excellent account of the relationship between cognition, motivation, task demands, and performance.

The length of the noise exposure often changes the effects of noise. For example, Smith (1983a) showed that recall of order information was initially better in noise than in quiet, but this effect was reversed in the second half of the experiment. Smith and Broadbent (1985) exposed subjects to noise for 20 min and then required them to name colors and read color names printed in black ink. The noise exposure changed the color/word ratio, with subjects being relatively quicker at naming the colors than at reading words after being exposed to noise. This effect did not occur when the color-word test was carried out at the beginning of the noise exposure. Smith and Broadbent have explained their results in terms of an effect of noise on the focusing of attention. Subjects carry out the word-reading task using some of the established techniques of reading, whereas such strategies are not used in naming colors. In reading, the eyes are often focusing on words that are well ahead of the word that is spoken. If noise produces a focusing of attention, the reading of words will be slowed down, and the narrowed attention might also benefit the naming of single items, such as color patches. These results are, therefore, similar to the "funnel vision" effects described in detail by Broadbent (1979). Many of these studies (e.g., Hockey, 1970) showed that funneling of attention does not occur until the subject has been in noise for some time, and this fits in with the absence of an initial effect of noise on the naming of colors and reading of words.

The effects of noise on cognitive performance may be summarized in the following way.

We now have performance tests that are sensitive to levels of noise frequently encountered in everyday life.

The nature of the noise is very important, and speech sounds seem to be particularly disruptive even when the task involves reading rather than listening.

There are effects of noise related to internal speech, but these are complex and cannot be accounted for by a simple masking theory.

Noise often alters the motivation of the subjects, and there is a need to study the relationship between motivation and cognitive performance.

Many noise effects are variable and can be changed by altering features of the tasks. However, some general effects have been isolated and these are listed below.

- Noise may lead to the choice of certain strategies in preference to others.
- Noise often reinforces the use of the dominant strategy.
- Noise produces biases in the allocation of processing resources.
- Noise influences the efficiency of the control processes that track and change performance. Subjects in noise are often rather inflexible and unable to change strategy as easily as subjects in quiet.
- Noise may produce focused attention, but this effect appears to depend on the subject's being in noise for some time.

The variability of the effects of noise can be explained by the hidden-defect and strategy choice approaches. At the moment it is unclear whether the effects of noise on cognitive performance are due to changes in state, such as changes in the level of arousal, or are produced by changes in cognitive motivation (see Jones, 1984; Schonpflug, 1983).

COMPARISON OF THE EFFECTS OF VARIOUS STRESSORS ON COGNITIVE PERFORMANCE

Hockey and Hamilton (1983) have described two strategies for stress research. The first, which they have called the narrowband approach, compares the effects of a number of stressors on a single task. An example of this can be seen in the Cambridge studies with the five-choice serial reaction time task. It is difficult to apply this approach to the effects of stressors on cognitive performance because no single task has been used in the same way to study many different stressors. It can be done in a very limited way, and this will be illustrated by considering the effects of different variables on syntactic reasoning.

Baddeley (1968) developed a syntactic reasoning test for use in stress research. Subjects score differently on this test after breathing nitrogen under pressure (Baddeley et al., 1968), at different times of day (Folkard, 1975), and under conditions of conglomerate noise (Smith, 1985a). Interestingly, test results are not affected by alcohol (Baddeley, 1981), continuous or intermittent white noise (Smith, 1985a), or breathing oxygen-helium mixtures (Lewis and Baddeley, 1981). However, it is not known whether the differential sensitivity of the test reflects some general property of the variables or the specific levels of the variables used in the experiments.

Many different variables have been compared as either decreasing or increasing arousal level (or both, depending on exposure or dose). Performance has been related to arousal by an inverted-U function, and more-complex tasks are assumed to show a performance peak at lower levels of arousal than more simple tasks do. However, it is misleading to consider variables such as noise, sleep loss, time of day, and use of amphetamines and alcohol in terms of a single dimension of arousal. Broadbent (1971) showed that the variables must be subdivided into at least two groups. He classified variables by the effects they had on the five-choice serial reaction time task. Some variables (e.g., alcohol use, time of day, barbiturate use, and the personality dimension of introversion or extroversion) influenced the average speed of response, whereas others primarily affected the proportion of errors or blocks in performance (e.g., sleep loss, noise, and amphetamine and chloropromazine use). In each group there were variables that might be considered stimulating and depressing. The former reduce or cancel the effect of the latter, provided they are from the same group (e.g., the effect of sleep loss is reduced by noise). A variable from one class has no consistent effect on variables in the other (e.g., there will be few interactions between noise and time of day). Thus, instead of a single dimension of arousal, Broadbent proposed two mechanisms. One, the lower level, corresponded to the traditional concept of arousal, and the other, the upper level, monitored and controlled the system that attempted to compensate for the departure of arousal from the optimum level. If the upper level is in good shape, the state of the lower level has little effect on performance, but when the upper level is ineffective, performance is affected by over- or underarousal. This theory has not been seriously tested and is probably wrong in detail; the point is that a theory at least this complex is required.

Hockey and Hamilton (1983) contrast the narrow-band research strategy with the broad band strategy, which is concerned with profiles of individual stressors, and the previous account of the effects of noise is an example of this approach. At the moment it is difficult to compare the profiles of other stressors with that of noise because their effects have not been studied in the same detail. To compare different stressors, one has to look at general indicators rather than specific tasks. A weakness of this approach is that the variation that occurs within a stressor or condition is often neglected in order to make between-stressor comparisons. For example, Hockey (1983) argues that noise increases alertness and selectivity and has little effect on speed, but impairs accuracy and short-term memory. However, prolonged noise exposure can decrease alertness (see Hartley, 1973), and studies of the Stroop task have shown that the effects of noise on selectivity vary (see Smith and Broadbent, 1985, for a review of this area). As already pointed out, the effects of noise on short-term memory depend on factors such as the memory load (Smith, 1983b), the nature of the material to be remembered (Wilding and Mohindra, 1980), and the nature of the noise (Salame and Baddeley, 1982). Smith (1985c) has shown that moderate-intensity noise does not increase errors in a serial reaction time task, in agreement with Broadbent's (1979) conclusion that the noise intensity must be greater than 95 dB to get such effects. Although it is agreed that it is valuable to compare the effects of different stressors, it is argued that this should not be carried out at the expense of detailed information about individual stressors. To obtain detailed profiles of performance, the microstructure of responding must be examined; for example, the strategies used to perform the task, not just gross indicators such as mean reaction time or number of errors. Suppose, for example, that subjects in one condition responded like the tortoise (slowly but steadily), whereas in another they were like the hare (quick bursts followed by long rests). If only the average response times are examined, the two conditions might seem equivalent, whereas a detailed examination of the strategies used would reveal differences between the two. I would recommend Hockey's approach of drawing up profiles of performance for each stressor. However, these profiles must be at least as detailed as the one already outlined for noise, and the profiles of performance should be accompanied by detailed profiles of physiological changes and subjective protocols.

There is a great need to examine the interaction of stressors in detail. Many of the early studies examining interactions used perceptual-motor tasks (e.g., the five-choice serial reaction time task), and there are few data on the effects of combined variables on cognitive tasks. The rationale behind interaction studies now has to be considered in more detail. When variables were linked together in terms of a single dimension of arousal, predicting when an interaction would occur was straightforward. When the subject had a low level of arousal (e.g., when in a sleep-deprived state), the addition of an arousing stimulus (such as noise) would improve performance because the subject was brought nearer to the optimum arousal level. Broadbent's (1971) classification suggested that variables will only interact consistently if they affect the same level. It is quite likely that the combined effects of two variables will also depend on the nature of the task. However, it is still unclear whether variables only interact if they have opposite effects on a task or whether some more complicated pattern may be found.

CONCLUSIONS

This account has shown that the effects of stressors on cognitive performance can be studied in several different ways. The empirical approach, which has typified the study of stressors such as heat, has been shown to be of practical importance but of little use for comparing stressors. Interpretations of the effects of stressors within a single framework, such as arousal theory, have also been shown to be inadequate. It is suggested that detailed profiles of the effects of each stressor be produced and comparisons made on the basis of such profiles rather than on more general indicators. An example of a detailed performance profile showed that the microstructure of responding must be analyzed and motivational and strategic changes must be examined. To make comparisons on the basis of more general indicators is to neglect the variety of different effects produced by each stressor and also to ignore the importance of specific parameters of the stressor. Even with detailed performance profiles for each stressor, comparisons will still be difficult because equivalent levels of different variables are not known.

REFERENCES

Allen, M.A., and G.J. Fischer. 1978. Ambient temperature effects on paired associate learning. Ergonomics 21:95-102.

Baddeley, A.D. 1968. A three-minute reasoning test based on grammatical transformation. Psychonom. Sci. 10:341-342.

Baddeley, A.D. 1981. The cognitive psychology of everyday life. Br. J. Psychol. 72:257-269.

Baddeley, A.D., L. Di Giguerlo, W. Hawkeswell, J.W. Curtis, and A.N. Williams. 1968. Nitrogen narcosis and underwater performance. Ergonomics 11:157-167.

Broadbent, D.E. 1971. Decision and Stress. Academic Press, London. 522 pp.

Broadbent, D.E. 1979. Human performance and noise. Pp. 17-1 - 17-20 in C.S. Harris, ed. Handbook of Noise Control, 2nd Ed. McGraw-Hill, New York.

Broadbent, D.E., P.F. Cooper, P. FitzGerald, and K.R. Parkes. 1982. The Cognitive Failures Questionnaire (CFQ) and its correlates. Br. J. Clin. Psychol. 21:1-16.

Cohen, S. 1980. Aftereffects of stress on human performance and social behavior: a review of research and theory. Psychol. Bull. 88:82-108.

Cohen, S., G.W. Evans, D.S. Krantz, and D. Stokols. 1980. Physiological, motivational and cognitive effects of aircraft noise on children: moving from the laboratory to the field. Am. Psychol. 35:231-243.

Colle, H.A. 1980. Auditory encoding in visual short-term recall: effects of noise intensity and spatial location. J. Verb. Learn. Verb. Behav. 19:722-735.

Daee, S., and J. Wilding. 1977. Effects of high intensity white noise on short-term memory for position in a list and sequence. Br. J. Psychol. 63:335-349.

Dornic, S. 1975. Some studies on the retention of order information. Pp. 230-240 in P.M.A. Rabbitt and S. Dornic, eds. Attention and Performance, Vol. 5. Academic Press, London.

Dornic, S., and S.E. Fernaeus. 1981. Type of processing in high-load tasks: the differential effects of stress. Report No. 576. Department of Psychology, University of Stockholm, Sweden.

Eysenck, M.W. 1975. Effects of noise, activation level, and response dominance on retrieval from semantic memory. J. Exp. Psychol. Hum. Learn. Mem. 1:143-148.

Eysenck, M.W., and M.C. Eysenck. 1979. Memory scanning, introversion-extroversion and levels of processing. J. Res. Pers. 123:305-313.

Folkard, S. 1975. Diurnal variation in logical reasoning. Br. J. Psychol. 66:1-8.

Glass, D.C., and J.E. Singer. 1972. Urban Stress: Experiments on Noise and Social Stressors. Academic Press, New York.

Hamilton, P., G.R.J. Hockey, and J.G. Quinn. 1972. Information selection, arousal and memory. Br. J. Psychol. 63:181-189.

Hamilton, P., R. Hockey, and M. Rejman. 1977. The place of the concept of activation in human performance processing theory: an integrative approach. Pp. 463-486 in S. Dornic, ed. Attention and Performance, Vol. 6. Academic Press, London.

Hancock, P.A. 1982. Task categorization and the limits of human performance in extreme heat. Aviat. Space Environ. Med. 53:778-784.

Hartley, L. 1973. Similar and opposing effects of noise on performance. Pp. 379-388 in W. Dixon Ward, ed. Proceedings of the International Congress on Noise as a Public Health Problem. U.S. Environmental Protection Agency, Washington, D.C.

Hockey, G.R.J. 1970. Signal probability and spatial location as possible bases for increased selectivity in noise. Q. J. Exp. Psychol. 22:37-42.

Hockey, G.R.J. 1979. Stress and the cognitive components of skilled performance. Pp. 141-177 in V. Hamilton and D.M. Warburton, eds. Human Stress and Cognition: An Information Processing Approach. John Wiley & Sons, Chichester, U.K. 396 pp.

Hockey, R., and P. Hamilton. 1970. Arousal and information selection in short-term memory. Nature 226:866-867.

Hockey, R., and C. Hamilton. 1983. The Cognitive Patterning of Stress States. Pp. 331-332 in G.R.J. Hockey, ed. Stress and Fatigue in Human Performance. John Wiley & Sons, Chichester, U.K.

Hormann, H., and J. Osterkamp. 1966. Über den Einfluss von kontinuierlichem Lärm auf die Organization von Gedächtnisinhalten. Z. Exp. Angewandte Psychol. 13:31-38.

Jones, D.M. 1984. Performance effects. Pp. 155-184 in D.M. Jones and A.J. Chapman, eds. Noise and Society. John Wiley & Sons, Chichester, U.K.

Jones, D.M., and D.E. Broadbent. 1979. Side effects of interference with speech by noise. Ergonomics 22:1073-1081.

Jones, D.M., A.P. Smith, and D.E. Broadbent. 1979. Effects of moderate intensity noise on the Bakan vigilance task. J. Appl. Psychol. 64:627-634.

Krenauer, M., and W. Schonpflug. 1980. Regulation und Fehlregulation im Verhalten. III. Zielsetzung und Ursachenbeschreilbung unter Belastung. Psychol. Beitr. 22:414-431.

Lewis, V.J., and A.D. Baddeley. 1981. Cognitive performance, sleep quality and mood during deep oxyhelium diving. Ergonomics 24:773-793.

Loeb, M., D.H. Holding, and M.A. Baker. 1982. Noise stress and circadian arousal in self-paced computation. Motiv. Emotion 6:43-48.

Millar, K. 1979. Noise and the 'rehearsal-masking' hypothesis. Br. J. Psychol. 70:565-577.

Poulton, E.C. 1970. Environment and Human Efficiency. Charles C. Thomas, Springfield, Ill. 162 pp.

Rabbitt, P.M.A. 1979. Current paradigms and models in human information processing. Pp. 115-140 in V. Hamilton and D.M. Warburton, eds. Human Stress and Cognition. John Wiley & Sons, Chichester, U.K.

Salame, P., and A.D. Baddeley. 1982. Disruption of short term memory by unattended speech: implications for the structure of working memory. J. Verb. Learn. Verb. Behav. 21:150-164.

Schonpflug, W. 1983. Coping efficiency and situational demands. Pp. 299-330 in R. Hockey, ed. Stress and Fatigue in Human Performance. John Wiley & Sons, Chichester, U.K.

Schulz, P., and W. Schonpflug. 1982. Regulatory activity during states of stress. Pp. 51-73 in W. Krohne and L. Laux, eds. Achievement, Stress and Anxiety. Hemisphere, Washington, D.C.

Sharma, V.M., G. Pichan, and M.R. Panwar. 1983. Differential effects of hot-humid and hot-dry environments on mental functions. Int. Arch. Occup. Environ. Health 52:315-327.

Smith, A.P. 1982. The effects of noise and task priority on recall of order and location. Acta Psychol. 51:245-255.

Smith, A.P. 1983a. The effects of noise and time of task on recall of order information. Br. J. Psychol. 74:83-89.

Smith, A.P. 1983b. The effects of noise and memory load on a running memory task. Br. J. Psychol. 74:439-445.

Smith, A.P. 1983c. The effects of noise on strategies of human performance. Pp. 797-808 in G. Rossi, ed. Proceedings of the Fourth International Congress on Noise as a Public Health Problem. Turin, Italy.

Smith, A.P. 1985a. The effects of different types of noise on semantic processing and syntactic reasoning. Acta Psychol. 58:263-273.

Smith, A.P. 1985b. The effects of noise on the processing of global shape and local detail. Psychol. Res. 47:103-108.

Smith, A.P. 1985c. Noise, biased probability and serial reaction. Br. J. Psychol. 76:89-95.

Smith, A.P. In press. Circadian variation in stress: the effects of noise at different times of day. In C. Spielberger and J. Strelau, eds. Stress and Anxiety, Vol. 12. McGraw-Hill, New York.

Smith, A.P., and D.E. Broadbent. 1981. Noise and levels of processing. Acta Psychol. 47:129-142.

Smith, A.P., and D.E. Broadbent. 1982. The effects of noise on recall and recognition of instances of categories. Acta Psychol. 51:257-271.

Smith, A.P., and D.E. Broadbent. 1985. The effects of noise on the naming of colours and reading of colour names. Acta Psychol. 58:275-285.

Smith, A.P., D.M. Jones, and D.E. Broadbent. 1981. The effects of noise on recall of categorized lists. Br. J. Psychol. 72:299-316.

Weinstein, N.D. 1974. Effect of noise on intellectual performance. J. Appl. Psychol. 59:548-554.

Weiskrantz, L. 1982. Relation between task parameters and allocation of effort in noise. Final Report to Social Science Research Council on Grant HR 6384.

Wilding, J., and N. Mohindra. 1980. Effects of subvocal suppression, articulating aloud and noise on sequence recall. Br. J. Psychol. 71:247-261.

Wilding, J., and N. Mohindra. 1983. Noise slows phonological coding and maintenance rehearsal: an explanation for some effects of noise on memory. Pp. 759-770 in G. Rossi, ed. Proceedings of the Fourth International Congress on Noise as a Public Health Problem. Turin, Italy.

Wilding, J., N. Mohindra, and K. Breen-Lewis. 1982. Noise effects in pre recall with different orienting tasks. Br. J. Psychol. 73:479-486.

Wing, J.F., and R.M. Touchstone. 1965. The effects of high ambient temperature on short-term memory. Wright-Patterson Air Force Base Report AMRL-TR-65-103. Aerospace Medical Research Laboratories, Dayton, Ohio.

Wyon, D.P., I.B. Anderson, and G. Lundquist. 1979. The effects of moderate heat stress on mental performance. Scand. J. Work Environ. Health 5:352-361.

COGNITIVE EFFECTS OF NEUROTOXIC AGENTS

David A. Eckerman

Neurotoxicology seeks to characterize and remedy the deleterious effects of chemicals and other agents brought into the body. Neurotoxic effects may well mimic the effects of inappropriate diet, one reason to review toxic effects on cognition in the context of this conference. A second reason is perhaps more urgent. Although effects resulting from dietary deficiencies are typically highlighted in the nutrition literature, food is also a major source of toxicants. The fields of neurotoxicology and nutrition are thus unfortunately intimate. In the present report, I will review several issues addressed in the neurotoxicity literature and approaches that have been taken to evaluate the effects of neurotoxicants, in hopes that our struggles will be informative to those seeking to evaluate the related question of the effects of nutrition. Cognition is interpreted here in the broadest sense of acquiring and using information about the environment and thus is a topic quite compatible with various accounts of the human condition, including radical behaviorism. The effects of neurotoxicants on other aspects of behavior will also be noted in passing, but the focus will be on impairment of cognition by toxicants. The report is divided into three sections: a review of issues that seem common to both endeavors, then a review of several approaches to evaluating neurotoxic effects, and lastly a review of some exemplary data from our field. In preparing this review, I have relied heavily on several recent reviews: Anger (1985), Anger and Johnson (1984), Bornschein *et al*. (1980), Feldman *et al*. (1980), Johnson and Anger (1983), and Weiss (1983).

ISSUES COMMON TO NEUROTOXICITY AND NUTRITIONAL ASSESSMENT

Our Problems Are Largely Man-Made and Are Increasing

Some of the toxic agents we ingest are not man-made. Were these our only toxins, our job would be relatively simple. We have evolved defenses against many of them, through natural avoidance, rapid conditionability, and lore. Furthermore, attention can often minimize any permanent damage. Unfortunately, most of the toxic agents we ingest, absorb, or inhale are very much man-made, and we are unprepared by evolution or social history to detect and deal with their sequelae. That carbon monoxide from incomplete combustion and lead from industrial products had been identified as highly toxic more than 2,000 years ago and are still major sources of toxicity demonstrates our inability to resolve these issues (Anger, 1982). We have not solved old problems, and new ones are accumulating at an increasing pace. New chemicals are developed at a dramatic rate. That most of the industrial chemicals known by 1977 to be neurotoxic were identified as such only after 1960 (Anger, 1982) shows the danger. Each year we introduce hundreds of chemicals with no strategy for evaluating their long-term toxic effects. Many of these agents find their way into our food. The recent determinations that ethylene dibromide levels in grains have grown unexpectedly and that imported fruits and vegetables come to to us with pesticides banned for use in this country, although they are produced here, emphasize our inability to keep these chemicals out of our food. And we purposely introduce relatively untried chemicals into our food. Weiss (1983) points out that food additives are "probably the most ubiquitous products of modern chemistry." How these additives interact with the unintended ones is another threat.

Many of our early warnings of toxicity have come from industry, where opportunities for exposure are many and often prolonged. Of the 100,000-plus chemicals found in the workplace, 850 are now known to be neurotoxic (Anger and Johnson, 1984). Thus, there are many toxicants in the workplace even though they are but a small proportion (about 1%) of the chemicals present. The prevalence of exposure worsens the situation. Of the 197 industrial chemicals to which over 1 million workers are exposed in U.S. worksites, 65 (33%) are known neurotoxicants. An estimated 7.7 million U.S. workers are exposed full-time to these neurotoxic agents (Anger, in press). Since worksite practices in other countries are seldom better than those in the United States, the problem is magnified when viewed worldwide. When extended to the population at large, the problem of toxic exposure comes to its full account. Individuals are exposed through air, food, and drinking water. In 1982, I attended a conference on evaluation of hazardous waste effects at which it was noted that research on toxic effects was limited by our inability to find a truly unexposed control population in the United States.

Assessment of Impairment Is Very Difficult

Neurotoxic effects are often subtle, as are the effects of inadequate nutrition. Bornschein *et al.* (1980), while evaluating the neurotoxic effects of lead poisoning, however, noted that subtle does not in this case mean small. Large but hard to evaluate effects and effects for which it is difficult to find a cause are also characterized as subtle. Xintaras and Burg (1980) note that for organophosphate pesticide poisoning, "the onset of some effects, in fact probably most of the effects, is so insidious that the worker may not realize that any change is taking place." Cognitive impairment is one such subtle kind of toxic impairment. Difficulty in remembering and concentrating is one of the most common neurotoxic complaints (e.g., Anger and Johnson, 1984).

Because toxic effects other than those involving cognition are easier to measure, the most documented neurotoxic impairment is in the sensory-motor functions (see Johnson and Anger, in press; Shaumburg et al., 1981). In his evaluation of industrial chemicals, Anger (1984) lists the 36 most common categories of documented neurotoxic impairment (Table 1). To these, I have added four categories from Anger and Johnson (1984). Only 7 of these 40 are cognitive difficulties, and sensory-motor difficulties make up the majority, 17. General complaints and emotional and personality effects constitute an additional category of 16 effects. What lesson might nutrition researchers learn from Table 1? One is that they should study the effects of nutrition on sensory and motor functions as well as on cognitive functions, because determining impairment in these functions may be somewhat easier. A second lesson might be to evaluate general, emotional, and personality effects as well, even though documentation on these effects will be even more difficult than on cognitive effects. A last lesson is that effects on learning, memory, and attention processes seem to be the most common cognitive effects. Nutritional deficits and dietary toxicants might be expected to take their toll on these processes as well.

It may be instructive that the cognitive effects found with neurotoxicants often involve a slight but general slowing or "dulling" of performance. Neurotoxicants seem to be blunt instruments of impairment in contrast to the sharp, transmitter-specific alterations in performance produced by pharmacological agents. Neurotoxicants, after all, are not selected on the basis of a narrow effect on behavior. Their effects are discovered as side effects, usually long after the substance is in common use.

Pharmacological agents, on the other hand, are selected expressly for a specific effect on particular functions, relative freedom from side effects being a necessary precondition to their release. Therefore, although we should be on the alert for function-specific toxic effects, we are likely to identify an agent as toxic on the basis of a generalized, subtle decrement in performance. The goal of neurotoxic assessment is to ensure that such subtle effects are detected. Nutritional assessment seems similar. Perhaps both fields should model their evaluations after those demonstrating the effects of age on performance (e.g., Hunt and Hertzog, 1981), since aging also seems to exact a general, subtle, and often denied toll.

Common Methodological Issues

What Subjects to Use? To carry out a planned study with human subjects, much must already be known about the effects! The effects

TABLE 1. Symptoms reported for neurotoxicity

Function	Anger (1984), Anger and Johnson (1984)	Weiss (1983)	Valcuikas *et al.* (1979)
Motor	Activity changes		
	Incoordination	X	X
	Paralysis	X	
	Pupil size change		
	Reflex abnormal		
	Weakness	X	X
	Ataxia		
	Convulsions	X	
	Tremor/twitching	X	
Sensory	Auditory disorder		
	Irritability		
	Equilibrium change	X	X
	Olfaction disorder	X	
	Pain disorder		
	Pain, feelings	X	
	Tactile disorder	X	X
	Vision disorder	X	X
Cognitive	Confusion	X	
	Memory problems		
	Speech/reading impairment		
	Alertness/concentration problem		
	Mental retardation	X	
	Indecisiveness/judgment		
	Learning impairment		
General	Anorexia	X	
	Autonomic dysfunction		
	Cholinesterase increase		
	Central nervous system depression		
	Narcosis/stupor	X	X
	Peripheral neuropathy	X	
	Fatigue/lethargy	X	X
		Headache	
Affect/ Personality	Apathy		
	Delirium		
	Depression	X	X
	Excitability		
	Hallucinations		
	Irritability	X	
	Nervous/tense	X	X
	Restless		
	Sleep disturbances	X	
		Psychiatric signs	

must be demonstrably reversible. Furthermore, since considerable resources are required to carry out an adequate study of subtle effects on human subjects, great urgency must be felt in obtaining the result. Studies with nonhuman subjects may well be needed to answer questions about reversibility and importance. As noted by Weiss (1983), "The difficulties posed by studies of human populations make well-designed animal experiments that much more crucial." A review is provided by Cabe and Eckerman (1982) (see Miller and Eckerman, in press, for an updated review) for much of the work on toxic effects on the complex behavior of nonhumans. Study of nonhuman model sytems is instructive even though they are only analogous to human processes. As an example, the radial-arm maze provides a rapidly learned complex rodent performance. In this maze, the rat avoids alleys it has already entered that day, showing good accuracy even for delays of 4 to 6 h (Beatty and Shavalia, 1980). Their performance is analogous to list learning and spatial orientation in humans and is sensitive to toxic impairment (e.g., Miller _et al_., 1982) and subject to pharmacological enhancement, as had been found in recent work with humans (e.g., Gallagher _et al_., 1983). Their performance is not, however, directly analogous to human memory tasks, as shown by its lack of disruptibility by retroactive interference (Maki _et al_., 1979). The effects of cholinergic manipulations seem to be analogous to those seen in humans (Godding _et al_., 1982).

For both scientific and political reasons, the greatest impact on decision making is made by a careful study with human subjects. Questions of generalization and relevance to human problems are simplified somewhat when the subjects are humans. Studies seeking to characterize cognitive effects must be even more cautious when generalizing across species, although humans are not, I would hazard, as unique in problem-solving processes as some proclaim. We must interpret work with nonhuman subjects cautiously, since we will ultimately slow progress by too quickly accepting an analogous nonhuman system as homologous. We should also cautiously interpret work in human subjects, since the issue of generalization is not simplified just because the subjects are human. Generalization from one subpopulation of humans to another should also be made cautiously (e.g., from allergic to nonallergic, male to female, young to old, culture A to culture B, etc.).

Individual or Group Analyses? Weiss (1980) advised skepticism in conventional analyses:

> The enormous variation in susceptibility to toxic substances, including lead, that prevails in the human population is widely recognized. The extraordinary polymorphism of our species guarantees that the ideal Gaussian distribution of responses to such substances will almost never be observed in practice. Instead, we are likely to see skewed distributions, often marked by multiple modes as well. The distributions derived from actual rather than hypothetical populations probably also reflect multiple mechanisms of susceptibility

rather than variations arising from different degrees of susceptibility to a single mechanism. We are all flawed in different ways. Even if one grants only moderate credibility to this proposition, it still contributes enough distortion to most statistical assumptions that conventional analyses based on group designs have to be regarded with some skepticism.

What implications has this statement for an evaluation of the effects of toxicants? It is suggested that allergy is a good model for toxic effects--some individuals may be peculiarly sensitive to allergens while the group average may suggest they have no effect. For a practicing clinician, an individual analysis is clearly called for. But the usual goal of a neurotoxicity assessment is to determine health risk for the population. Are individual analyses called for? It would appear that we must at least look beyond the group mean, measuring increases in variability, number of outliers, or incidence of subcriterion performance. In an exemplary series of studies, Brown et al. (1981) (see below) determined the number of individuals tested who fell below the 15th percentile. This approach may represent the best way to answer population-oriented questions while acknowledging individual sensitivities.

Role of Retrospective Studies. Although planned studies have great power, they are limited even when ethically defensible. Generalization from planned studies to situations about which decisions must be made may be limited by several factors: selection factors (volunteers are seldom representative), limited duration of exposure (acute and long-term effects may well differ), and special conditions of the laboratory (e.g., attention given the subjects, eagerness to comply with experimenter and to "meet the challenge" of the tests, etc.). With disturbing frequency, we are provided with populations of individuals accidentally exposed to toxicants (and inadequate diet). Such situations are almost always directly connected to problems that decision makers are called on to solve and thus are directly relevant. Were there a way to evaluate these unfortunate "natural experiments," our research power would be multiplied considerably. Such field studies constitute much of the toxicity assessment literature and will be pointed out below. The many potentially confounding factors in such studies are, however, very difficult to control.

An instructive example of the dangers found in field studies can be found in the thorough follow-up on the accidental introduction of polybrominated biphenyls (PBB) into the food chain in Michigan in 1973. An extensive field study was undertaken by the Mount Sinai Environmental Sciences Laboratory (Valcuikas et al., 1978b). The approach included extensive biological and environmental measurement of PBB and other indicators, evaluation of neurological symptoms, and performance measures that had been found to be sensitive in prior work in evaluating lead toxicity (Valcuikas et al., 1978a). Evidence of neurological impairment was found by comparing measures for Michigan farm residents with those for Wisconsin farm residents. Symptoms included nervousness, parasthesias, somnolence, dizziness, loss of balance, and headaches (Valcuikas et al., 1978b). In following up on these evaluations, Brown and Nixon (1979) evaluated memory problems reported by several of the exposed individuals, finding impairments in prose recall and in short-term retention of words. They were cautious in interpreting these differences, however, because the memory impairment was correlated with measures of anxiety and depression on the Minnesota Multiphasic Personality Inventory (MMPI) but was not correlated with PBB concentrations. They therefore undertook a laboratory evaluation with chemical workers who had been exposed to higher levels of PBB (Brown et al., 1981). Whereas farm residents had a mean PBB concentration of 3.9 ppm, the chemical workers had a mean concentration of 9.3 ppm. Although the chemical workers studied did not differ in educational, occupational, or cultural variables, their performance on six memory tests was equal to or better than that of the farm residents. Brown et al. concluded that depression and other "factors affecting psychological functioning may be more important causes of memory complaints and low scores on objective tests of memory than the level of PBB contamination." It is clear from this progression that field evaluations are fraught with potential confoundments. The best research strategy would undoubtedly utilize field studies, but would provide considerable laboratory backup to support the interpretation of findings.

Single or Repeated Measures? In some situations only one testing session can be arranged for a subject. This session is unlikely to produce optimal performance, since several things are simultaneously being learned--the laboratory situation, the responses required, the range of stimuli that can be expected, the dimensions that distinguish alternative stimuli, the feedback provided, and, finally, the task itself. Of course, performance at any point during this practice might be used to evaluate the effect of a treatment. Performance may even be more subject to impairment during early learning stages and thus provide more interesting results. Yet performance early in practice will vary more between as well as within subjects, making measurements less precise. Also, it is likely that preexperiment differences between subjects (practice, educational level, etc.) will confound treatment effects to a greater extent early in practice. Despite these drawbacks of brief testing, we have been developing a battery of tasks to offer a brief field test of neurotoxicity (Eckerman et al., 1985), since there are situations in which only one testing session can be hoped for. To optimize our measurements, we have worked very hard to make the instructions and tasks compatible and to sequence the tasks so as to minimize practice effects on individual tasks. I will describe our battery more fully below.

Repeated tests offer greater stability of measurement, although not all material is appropriate for repeated testing. For example, if words are to be learned, different forms must be used from day to day and comparability of forms must be ensured, a difficult matter.

Furthermore, repeated testing differs from single testing in a number of ways. Several aspects of task performance may change due to practice. A thorough study of task stability has been made by the Office of Naval Research Performance Evaluation Testing for Environmental Research (PETER) program. This project considered over 140 performance tests. Tasks that did not quickly show good correlation across days were excluded. For a bibliography of published work from this program, see Harbeson et al., 1983. Only 30% of the tasks have been found to be stable enough for inclusion in a repeated test battery (a proposed battery is described by Kennedy et al., 1984, and Bittner et al., 1984).

Desjardins et al. (1982) offer an interesting repeated measure of learning that derives from the operant conditioning tradition and includes themes from paired-associate learning. One task included in the testing sessions required learning a new sequence of stimulus-response associations each session, whereas the other task required learning the same chain of associations each day. Tasks were alternated each 15 min throughout each session. Typically, the acquisition task was more sensitive to drug impairment than the performance (same day to day) task. This procedure sets an excellent standard for stability of measurement in a continuous learning task for both human and nonhuman subjects. Yet caution seems appropriate in interpreting results from this and perhaps many other repeated acquisition procedures. What is learned each day? In the Desjardins et al. procedure, the subject learns which sequence is appropriate. Such learning may be quite unlike the initial acquisition of stimulus control or other ways of learning what to do.

Laboratory or Field Studies? Greater control of conditions and measurements is often afforded when we bring subjects into our domain, the laboratory, clinic, or hospital. But we are more likely to obtain a representative sample of subjects and obtain important additional information if we go to the worksite, the home, etc. Thus, Milar et al. (1980) and Hawk and Schroeder (1983) were able to make greater sense of the complex effects of lead on intelligence scores when they visited and characterized home environments. We can attempt to represent the field situation by obtaining survey data through the mail, yet it is rare for such a survey to produce quick and unbiased data. Sterman and Varma (1983) bemoan the fact that they had only mail-in survey responses and not door-to-door interview data when characterizing the incidence of neurological complaint resulting from invasion of the pesticide aldicarb into the water supply of Suffolk County, Long Island. Their 3-year effort to obtain data through the mail produced only 204 usable questionnaires from 1,500 households surveyed. These results suggested that a timely door-to-door survey would have provided clear evidence of impairments. Of course, field research has special requirements for which the laboratory scientist might be unprepared. Anger (1985b), for example, reviews special needs that must be met when recruiting subjects for and following up on worksite evaluations.

Survey or Performance Measures? We do not evaluate our own cognitive abilities well. It is good counsel to suggest that a spouse or other close associate confirm the individual's self-description when assessing dementia. Objective measurement is preferred over mere report. Sunderland et al. (1983) had head-injured, formerly head-injured, and control patients complete memory complaint questionnaires, keep checklists of memory failures for 7 days, and also complete a battery of nine cognitive tests. Relatives also completed the memory failure questionnaire and kept a checklist. Head-injured patients showed significant memory loss in all but one of the objective tests (immediate word recognition) and in relatives' evaluations and checklists. Memory complaint questionnaires filled out by the patients themselves, however, did not distinguish head-injured from control patients (and patients' checklists were only marginally different). The importance of objective measurement is thus confirmed. Yet it should be stressed that the objective measurements were not all indicative of the relatives' view of the individual's distress. When objective test scores were correlated with relatives' assessments, continuous recognition, face recognition, and immediate word recognition were unrelated to the estimated memory problem. Story recall and paired-associate recall were best related ($r = 0.35$ to 0.72). Objective test measures thus do not all capture the aspects of memory loss that are believed to be problems.

There is a place for subjective measurement in toxic evaluation, emphasized by the need to assess the personality changes commonly reported in cases of toxic exposure (irritability, depression, loss of libido). These shifts are difficult to document except through self-report. Proper construction makes symptom surveys less subject to misinformation given by malingerers (e.g., inclusion of lie scales and faking scales, as on the MMPI). I know of no instrument other than the arduous MMPI, however, that uses such an approach and is appropriate for neurotoxic evaluations.

Test for Crystallized as well as Fluid Intelligence? The acquisition and retrieval of new material, requiring what has been called fluid intelligence, is more likely to show impairment than is retrieval of very well-learned semantic and cultural information and problem-solving strategies, said to require crystallized intelligence (Horn, 1976). Thus, verbal subtests from the Wechsler Adult Intelligence Scale (WAIS), especially vocabulary, might be less sensitive to toxic impairment occurring during adulthood than the other subtests are. Also, the performance IQ might decrease in response to toxic impairment before the verbal IQ does. Such a shift is found for brain-damaged (all but left-brain-damaged subjects in Russell, 1979, 1980) and for normally aging subjects (see Hunt and Hertzog, 1981). It is instructive that a different pattern is seen for the long-term nutritional and toxic conditions represented in Korsakoff patients. These individuals appear to lose much crystallized intelligence as well and thus are distinguishable from dementia patients (Weingartner et al., 1983). For testing subclinical toxicity, measuring crystallized intelligence might be a useful way of matching exposed and nonexposed populations for premorbid intellectual capacity. Of course, such a difference would not be expected when impairment began in childhood.

Objective Scoring or Clinical Judgment? Administering psychological tests requires establishing rapport. Many tests, however, require considerable clinical training and judgment in scoring as well. For example, drawing ability, handwriting, and other measures of visual-motor coordination have been found to be sensitive to toxic impairment (Johnson and Anger, 1983). But scoring these performances is difficult and may not be reliable among all judges. To "go to court" with such data, then, requires that the person who scored the tests be present to testify and face potentially aggressive cross-examination, perhaps years after the test was administered. We have chosen objective measures for our test battery to simplify the evaluation of test results, realizing that we may be losing some sensitivity in doing so.

Complex or Simple Tasks? Simple reaction time may be slowed when an individual is exposed to certain toxicants (e.g., carbon disulfide [Tuttle et al., 1976] and fuel mixtures [Knave et al., 1978]). It is frequently recommended, however, that complex tasks are more sensitive to the effect of toxicants than are simple tasks. Why would that be? Simple tasks would seem to have the advantage in ease of training and stability of performance. Also, they would seem to be more easily aligned with particular component skills and thus offer better diagnostic information. One account of the advantage of using complex tasks is that subjects who have fewer processing resources are less able to simultaneously coordinate and execute a sequence of simple activities even though they are capable of executing each of the component tasks. Thus, in a dual-task arrangement, less attention can be assigned to a secondary task of peripheral visual signal monitoring by an impaired subject whose primary task is to keep a dot positioned over a moving target. Such a dual task has in fact been found to be a good indicator of carbon monoxide impairment in the laboratory (Putz, 1979) and in exposed toll booth operators (Johnson et al., 1974).

A second account of the advantage of complex over simple tasks is that more skills must be simultaneously intact to allow good performance--impairment of any of the component abilities would yield a decrement in performance. A performance that requires that several component skills be intact is called an apical test (Cabe and Eckerman, 1982). The hope when seeking an apical test is expressed in the left-hand panel of Fig. 1. Were any element in this chain broken, the performance would fall. An alternative arrangement of abilities is possible whenever a number of elements are brought together--when a task can be carried out in a number of ways, the loss of one ability may be compensated for by a shift in the solving strategy, and the impairment may thereby be minimized. Inasmuch as a complex task can be solved in several ways, the individual can use available processing

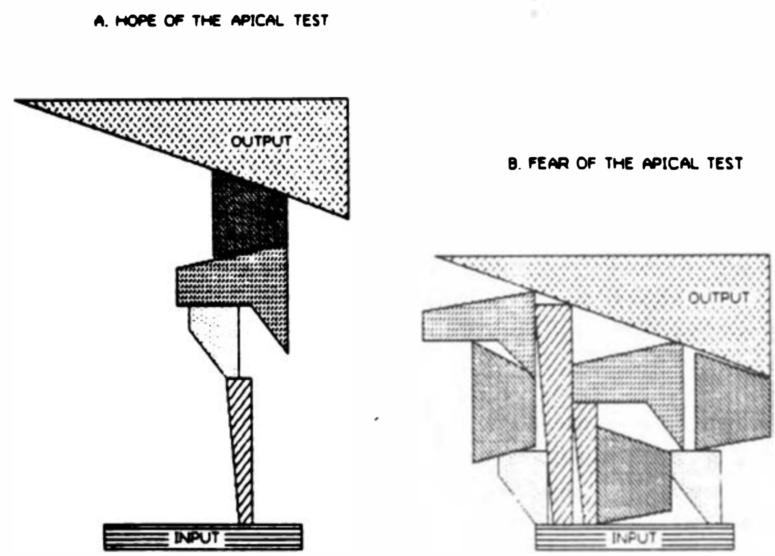


FIGURE 1. Global measures (apical tests) of behavior would be more sensitive as indicators of toxicity if processing were serial (A) but not if processing were parallel and redundant (B).

resources to make up for decreased ability. This possibility is reflected in the right-hand panel of Fig. 1. For example, Hunt and Davidson (cited in Hunt, 1983) detected an age-related decrease in the ability to visualize only by closely measuring strategy. Subjects over 40 years old shifted to the less efficient verbalization strategy when asked to solve a picture-sentence verification task. Simply looking at task accuracy and roughly estimating performance time would not have revealed this difference; close attention was required to detect the shift and determine its source. Thus, Ekel concludes (p. 435) that "only [by] investigating relatively basic, elementary components of complex behavior...is one likely to detect small, toxin-induced impairments." Smith and Langolf (1981) suggest (p. 701) that "the need for precise, ability-specific tests arises because gross performance measures may not pose a sufficiently severe test to detect the impaired system." Are there good apical tests of cognitive ability? Although complex tasks may be sensitive, it is not clear that their sensitivity is derived from an interdependence of functions.

Can Recognition Be Used Instead of Recall? More is required of subjects when they are asked to produce the correct word, drawing, melody, etc., rather than merely recognize which alternative is correct. One might then expect that recall would be a more sensitive indicator of impairment than recognition. A number of studies have found this to be true (e.g., in detecting memory changes with normal aging [Branconnier, 1985]; in detecting the effects of amitriptyline on memory [Branconnier *et al.*, 1985]). This great sensitivity is obtained at some cost, for it is difficult to automate the recording and scoring of extended responses such as spoken words or drawn pictures. Is there a way to get as much sensitivity by using recognition instead of recall? If registering or encoding information is impaired, then recognition would seem to be as good a test as recall. Williams and Rundell (1983) found that secobarbital impaired subsequent recognition as well as recall of nouns when no special encoding aids were given at word presentation. When special care was taken to ensure that the subject adequately encoded the word, recognition was less affected than recall. Birnbaum and Parker (1977) and Jones and Jones (1977) also found that recognition as well as recall was affected. In their studies, alcohol intoxication prevented adequate encoding of the stimuli. Bruce and Cofer (1965) point out that recognition may be made more difficult by the presentation of highly similar alternatives (e.g., similar in meaning). When the subject was required to choose between similar alternatives, memory impairment was more easily demonstrated. By using recognition procedures, the power of signal detection analyses to reveal response bias versus true loss of information (sensitivity) becomes available (e.g., Miller and Lewis, 1977; Ferris et al., 1980). With these demonstrations of sensitivity and the additional power of analysis, we have used recognition measures in our screening battery (described in Table 2, below). Were we required to sort out types of memory impairment that recognition procedures do not distinguish, sophisticated analyses of recall versus recognition are available that provide considerable power (e.g., Block and Berchou, 1984; Gonheim et al., 1984a, 1984b), but we anticipate that such analyses might be made after an impairment is indicated through screening.

CONSTRUCTING A BATTERY OF TASKS TO EVALUATE NEUROTOXIC IMPAIRMENT

Battery Used at the Institute of Occupational Health, Helsinki, Finland

Hannenin and colleagues have been evaluating the effects of occupational exposure to toxins since 1964. On the basis of their considerable experience, they have selected a battery of tests that they find sensitive and informative. These tests are listed in column 1 of Table 2, with tasks organized by the function they primarily measure. It should be noted that several of these tests require considerable training in administration and scoring, skills that are readily learned in the context of the institute. Reference to the extensive studies of this group may be found in Johnson and Anger (1983).

The WHO Core Battery

In May 1983, the National Institute for Occupational Safety and Health (NIOSH) and the World Health Organization (WHO) jointly sponsored a conference of international experts to recommend an approach to neurotoxicity testing that might be taken by Third World

TABLE 2. Batteries for testing neurotoxicity

Function[a]	Finnish Battery (Hanninen and Lindstrom, 1979)	WHO Battery[b] (Seppalainen et al., in press)	U.S. Field Evaluation (Johnson and Anger, 1983)	Baker-Letz Battery (Baker et al., 1985)	Microtox Battery[c] (Eckerman et al., 1985)
Sensory	Figure identification Kuhnburg figure		Flicker fusion Iconic memory Embedded figure Figure rotation Visual acuity Color vision Peripheral vision Hearing Touch/graphesthesia		Visual masking (1) Choice reaction time (distraction) (2) Mueller-Lyer (4)
Motor reaction	Finger tapping	Santa Ana dexterity (core)	Simple reaction time	Simple reaction time	Simple auditory
	Simple reaction time Choice reaction time	Aiming (core) Simple reaction time (core)	Choice reaction time Grip strength		time (2) Choice auditory
reaction	Santa Ana dexterity Mira Symmetry drawing	Finger tapping (ad) Flan. coordination (ad) Choice reaction time (ad) Pegboard (ad) Rate of manipulation (ad)	Tremor Balancing Tapping Pegboard Eye-hand coordination		time (2)
Cognitive attention	Bourdon-Wiersma Digit-symbol	Digit-symbol (core) Bourdon-Wiersma (ad) Letter search (ad) Test d-2 (ad) Auditory serial attention (ad) Continuous performance (ad) Stroop interference (ad)	Letter search Vigilance Stroop interference Digit-symbol Dual track/detect	Digit-symbol Continuous performance	Switching attention (3) Stroop interference (5)
Memory	Digit span Benton visual reproduction Logical memory, WMS Associate memory, WMS	Benton retention (core) Digit span (core) Visual retention, WMS (ad) Benton repetition Logical memory (ad) Associate memory (ad)	Memory scanning Memory span Pair association Delayed pair association Memory distractor test Logical memory Associate memory	Pair associate learning Pair associate recall Digit span Visual retention Memory scanning	Memory scanning (6) Simon sez (6) Digit span (6) Supra-span learning (6) Objects in house (6) Which first/last? (6) Continuous recognition (6) Items in categories (6)
Other	Similarities Block design Picture completion	Block design (ad) Similarities (ad)	Block design Mental arithmetic	Mill Hill vocabulary Armed Forces qual.	"Fuzzy set" bounds (3)

(TABLE 2 continues)

TABLE 2. Batteries for testing neurotoxicity

Function[a]	Finnish Battery (Hanninen and Lindstrom, 1979)	WHO Battery[b] (Seppalainen et al., in press)	U.S. Field Evaluation (Johnson and Anger, 1983)	Baker-Letz Battery (Baker et al., 1985)	Microtox Battery[c] (Eckerman et al., 1985)
Affective	Rorschach	Profile of mood states (core) Eysenck personality inventory (ad)	Adjective checklist Feeling tone checklist MMPI Clinical analysis questionnarie Edwards personality preference Manifest anxiety scale	Mood scales	

[a]The sorting of tasks by functions is somewhat arbitrary, as tasks usually tap several functions. The reader is encouraged to regroup the tasks where appropriate.

[b]The WHO Battery is divided into core tests (making up a brief, 40- to 50-min battery), additional tests (ad) requiring more time or resources and recommended where a fuller account is desired, and tests in development (requiring still more elaborate apparatus or training or as yet not validated for toxic assessments). This listing includes the core and additional tests.

[c]The cognitive tasks paradigm (Carroll, 1980) (see Table 3), which the task represents, is indicated in parentheses.

countries. The recommended core battery of tasks is listed in the second column of Table 2. It was the goal of that conference to provide a brief (40 to 45 min) test battery that offered a comprehensive and evenly distributed assessment for a broad range of abilities. They also sought to provide a battery that could be administered without highly specialized equipment or training, since their recommendations were especially directed to developing nations. These tests are appropriate for individuals with a broad range of educational backgrounds. The tasks included in the core battery were felt to be "variable enough to keep the subject interested and motivated to cooperate." A principal reason for including a task was that in prior field studies, it had "proven to be sensitive to a variety of neurotoxic agents" (Seppalainen _et al_., in press). The authors note, however, that "the reliability and sensitivity of some of these short tests are not optimal, and so this minimum battery may be unable to detect subtle changes if the assessed groups are small." Additional tests were indicated, which could be added to the evaluation to increase sensitivity or as follow-up evaluation for individuals who appeared to be impaired on the basis of core test performance. Some of these additional tests require special equipment. The authors also note the following cognitive factors as missing from their battery but of interest: word fluency, reasoning, spatial visualization, and tactile sensitivity. They also recommended word knowledge as a useful measure to indicate preexposure capacity.

Batteries of Tests Used in U.S. Field Evaluations of Neurotoxicity

Johnson and Anger (1983) review the many field evaluations of neurotoxicity carried out in the United States, for the most part under the direction of NIOSH. Table 2, column 3, lists the tests that were included in these studies. In general, these tests show somewhat more influence from the psychological laboratory and less influence from the neuropsychological clinic. The tests are typically simpler and more given to objective scoring, but effects have not been as consistently found in these studies as in those carried out in other countries.

Automated Test Batteries Still in Development

In October 1983, a workshop was held by the Environmental Protection Agency (EPA) in Rougemont, N.C., on approaches to the field testing of neurotoxic impairment in human populations. A number of investigators demonstrated batteries of tests they had implemented on microcomputer systems. Columns 4 and 5 of Table 2 show tasks under evaluation for inclusion in two of the field test batteries being developed specifically to aid in neurotoxicity assessments. The battery being developed by Baker and colleagues at the Harvard School of Public Health is influenced heavily by the field study tradition represented in columns 1 and 2 of Table 2. The battery being developed at the University of North Carolina by Eckerman and colleagues is somewhat more influenced by the laboratory traditions that have influenced other U.S. field evaluations. J.B. Carroll, a collaborator in the North Carolina-EPA battery, had proposed a typology of elementary cognitive task paradigms. This analysis has strongly influenced the selection of tasks for that battery, and the paradigm that the tasks included represent is indicated in parentheses. Carroll's task paradigms are shown in column 1 of Table 3, with representative tasks for each of these paradigms indicated in column 2. By factor-analyzing data from several studies that included representative tasks from two or more of these paradigms, Carroll determined that individual differences in cognitive task performance can be fairly well accounted for by assuming that there are 10 identifiable cognitive factors that differ among individuals. Descriptions of the factors he found are listed in column 3 of Table 3. Note that the map between tasks (columns 1 and 2) and factors (column 3) is not clearly drawn--each task really requires that several factors be utilized. Yet by judicious selection of tasks, it would appear that the various factors could be characterized for an individual. The tasks that Carroll suggests for assessment of the various factors are underlined in column 2.

The North Carolina-EPA effort represents a first step in an attempt to select a battery of tests to assess component functions indicated by a thorough theory of human cognitive processes. Candidate theories might be found in the developing literature on the nature of intelligence (e.g., Carroll, 1980; Hunt, 1983). This approach seeks to find clusters of related tests (e.g., through factor analysis) and then to include marker tests for these factors. Hunt proposes that one such factor is language use (vocabulary, paragraph comprehension), a second is visualization (examination and mental manipulation of visual patterns), a third is lexical access (recognizing that a set of letters is a word, that two stimuli belong to the same concept), and another is reasoning or fluid intelligence (analogical reasoning, series completion, detection of patterns in visual displays), which does not correlate with the visualization or lexical access factors. Interrelating these factors with those proposed by Carroll and others (e.g., Sternberg, 1977) would occupy more space than appropriate here.

Electrophysiological Measures of Cognition

Electrophysiological measures of cognitive processes have been found to be sensitive in detecting toxic impairment (e.g., Otto, 1983; Seppalainen *et al*., 1980; Seppalainen, in press). A portable field-testing battery for electrophysiological assessment is being developed by EPA. This battery includes measures of cognitive as well as sensory functioning (Otto *et al*., 1985).

Batteries Developed for Assessing Drug Effects

The drug evaluation literature should provide good counsel for both neurotoxic and nutritional assessment. In Table 4, I list those tasks

TABLE 3. Carroll's elementary cognitive tasks and cognitive factors[a]

Paradigm	Example tasks[b]	Factor
1. Perceptual apprehension	Recognizing objects Single letter threshold <u>Visual duration threshold</u>	1. Monitor process (e.g., instructions, strategies)
2. Reaction time and movement	Simple reaction time <u>Choice reaction time</u> Fitt's tapping task	2. Attention process (focus, stimuli expected)
3. Evaluation/decision	<u>Word or nonword</u> Dichotic order judgment Picture or nonpicture	3. Apprehension (sensory buffer representation)
4. Stimulus matching/comparison	Letter search Letter matching Letter rotation <u>Same/different pictures</u>	4. Perceptual integration (matching with prior buffer representations)
5. Naming/reading/association	Stroop task Word naming Naming pictures	5. Encoding process (formation of a mental representation, abstraction)
6. Episodic memory read-out	<u>Memory span</u> <u>Probed memory scanning</u> <u>Running recognition</u>	6. Comparison of mental representations
7. Analogical reasoning	People piece analogy Miller analogies test	7. Corepresentation formation (a new representation of a pair of representations)
8. Algorithmic manipulation	Mental arithmetic Sunday + Tuesday Raven's matrices	8. Corepresentation retrieval from memory
		9. Transformation of representation (e.g., rotation, transposition)
		10. Response execution planning (cognitive process preceding the overt response)

[a]Data from Carroll, 1980.
[b]Underlined tasks are included in Carroll's proposed experiments to estimate process parameters (Carroll, 1980, pp. 71-73).

TABLE 4. Test batteries used in example drug evaluations

Function[a]	Weingartner, Sitram, and Caine (Sitram et al, 1977, 1978; Weingartner et al., 1979, 1981, 1983)	Mohs and Davis, 1980	Branconnier (1985); Branconnier et al. (1982)	Ferris, Crook, and McCarthy (Ferris et al., 1977, 1980, 1982; McCarthy et al., 1981)	Squire et al., 1980
Sensory	Detect tone in noise Detect gap of silence Tonal order judgment		Iconic memory		
Motor				Finger tapping Simple reaction time Visual search (digit) Visual search (word)	Porteus maze
Attention	Detect repeated words	Time production Visual search Divided attention (digit transform, detect letter)			Visual search (pairs) Visual search (word) Digit-symbol Delayed target detection
Learning	Serial learning Selective reminding	Selective reminding	Selective reminding	Shopping list learning First/last names	Number learning Pair associates
Episodic memory	Distracted forgetting Consistent recall Visual pattern recall Delayed recall (10 min) Free vs category-cued recall, given different encodings	Supra-span digits Memory scanning Digit span (F/B) Delayed recall Story recall Consistent recall	Memory scanning Cued recall, semantic and phonetic Recognition memory, learned/incidental	Digit span (dialing) Running digit span Face recognition Misplaced objects Paragraph recall Category retrieval Design recall Pair-associate recall	Recognition of items (immediate) Recognition of items (10-min delay) Paragraph recall Design recall Self-ratings for memory complaints
Semantic memory	Recall words by first letter Recall words by category		Squire television test Category instance fluency		Remote event recall
Other cognitive		Trail map Serial sevens subtraction Picture arrangement	Spatial rotation time Simple visual reaction time National adult reading Road map test		Trails A and B
Affective		Bunny-Hamburg scale	NIMH mood scales Mood scale—elderly Psychic/somatic complaint		

[a]The sorting of tasks by function is somewhat arbitrary, as tasks usually tap several functions.

that five notable research groups have included in their drug evaluations. The work of Weingartner, Sitram, and colleagues (Sitram et al., 1977, 1978; Weingartner et al., 1979, 1981, 1983) offers an especially thorough coverage of memory effects (column 1), and the work of Davis, Mohs, and colleagues (Davis et al., 1976, 1978; Mohs et al., 1978a, b, 1980) (column 2) provides assessment of a somewhat broader range of functions. This group also focused on the memory-scanning tasks, which provide a measure of memory processing speed rather than mere capacity. For another instructive use of this task in drug toxicity evaluation, see MacCleod et al. (1982). The set of tasks used by Branconnier et al. (1985) are notable because they are commercially available as a portable hardware-software package. The tasks selected by Ferris and Crook (see also McCarthy et al., 1981) are notable, as they represent tasks that derive from common laboratory measures of cognitive processing and yet have high face validity for both young and old subjects from a broad range of the population. By arranging the learning of grocery lists, phone numbers, faces, and names, they are able to recruit subject cooperation and motivation more readily than by using nonsense syllables. Lastly, the study of Squire et al. (1980) broadly samples cognitive performance.

Assessing Unusual Working Environments and Conditions

Table 5 lists tasks selected for inclusion in several batteries developed to assess the effects of work environments and conditions on behavior. These include tasks selected to be sensitive to work cycle and shift duration variables (Walter Reed Performance Assessment Battery, column 1), tasks currently included in the automated battery being evaluated by the Navy for testing unusual work environments (column 2) (see Harbeson et al., 1983, for a bibliography of the evaluation studies of this group), a battery under development by the Air Force for use in selecting recruits with good potential for high-level training (column 3), and lastly the classic battery of Alluisi et al. (1973), here referred to through a study on the "toxic" effects of viral disease on performance (column 4). In general, these tasks have been chosen to tap more complex skills than those represented in the toxicity and drug evaluation literatures.

EXAMPLES OF NEUROTOXICITY EFFECTS

A Psychometric Approach to Measuring Lead Poisoning

One of the relatively firm results in what has become a complex literature on the effects of low levels of lead on behavior is found in a Science report of an evaluation of 90 secondary-lead smelter workers by investigators from the Mount Sinai Environmental Sciences Laboratory (Valcuikas et al., 1978a). It is often difficult to adequately characterize the body burden of lead. The level of lead in blood may merely reflect current levels and not the history of exposure. The concentration of lead in other tissues, e.g., bone, would be a better

TABLE 5. Batteries for testing unusual work environments

Function[a]	Walter Reed PAB (Thorne *et al.*, 1985)	Navy Biodynamics (Irons and Rose, 1985)	Air Force[a] (Allen and Morgan, 1985)	Alluisi Battery (Alluisi *et al.*, 1973)
Sensory/motor	Four-choice reaction time	Choice reaction time (auditory/visual) Spoke test		
Attention	Two-letter search Six-letter search	Pattern comparison		Warning lights Blinking lights
Learning/memory	Missing digit (scrambled) Pattern recognition (two levels of difficulty)	Memory scanning Auditory digit span Recognize letters	Coded messages (12 pairs) Emergency procedures (ordered set) Time check (compare analog and digital clocks) Security check (spatially ordered pair associates) Communication control (hierarchically structured verbal material) Strategic decision (object-number pairs and mental arithmetic) Direction judgment (map learning)	
Other cognitive	Letter-map coding Mental arithmetic, two-column addition and serial add/subtract Logical reasoning	Code substitution Serial addition Mental arithmetic (four functions) Manikin (figure rotation) Grammatical reasoning (visual, auditory)		Mental arithmetic Code-lock solving Target identification
Affect	Mood-activation scale			

[a]The sorting of tasks by function is somewhat arbitrary, as tasks usually tap several functions.

indicator of long-term accumulation, but measurement is inappropriately invasive. In some cases, lead in baby teeth has been measured to represent the exposure history of young children, but such a measurement is generally not possible with older individuals. Valcuikas et al. (1978a) determined body burden by using a relatively long-term measure--zinc protoporphyrin (ZPP) levels in the blood. This measure represents a 4-month averaging of lead burden.

Taking the age and educational background of the subjects into account is necessary to make an adequate evaluation. As most cognitive functions show a shift during the working years (see Hunt and Hertzog, 1981), age is an important variable to control. Educational level and mood (e.g., depression, irritability) are also important. Valcuikas et al. (1978a) determined and corrected for an age effect by using control as well as exposed group data. They then determined the correlation between ZPP level and age-corrected score on each of five measures--the WAIS block design subtest, the WAIS digit-symbol subtest, an embedded figures test developed by Valcuikas, and the dominant hand and both hands on the Santa Ana dexterity test. The tests taken from the WAIS are all from the performance IQ, which is more sensitive to brain insult. All tests are from the clinical testing tradition and are useful in making clinical decisions. Significant correlations were found for the first three measures (see Fig. 2). Inspection of the data shows that there was considerable individual variability and that the low values for individuals with extremely high levels of lead in the blood were an important determinant of the correlations. As Valcuikas et al. (1978a) point out, "the scatter of scores due to individual variability greatly exceeds the effect that ZPP levels have on scores for the population studied here. It is then impossible to draw conclusions about an individual's ZPP level or lead intoxication from his test scores alone." Yet the relation between ZPP level and performance is clear, representing an observable correlation that persisted to ZPP levels of less than 170 μg/dl.

An Experimental Laboratory Approach to Measuring Mercury Poisoning

Smith and Langolf (1981) used a task with a long-standing laboratory tradition and a growing clinical evaluation tradition in evaluating the effects of mercury on short-term memory. An assessment was made of 26 workers in two mercury cell chlor-alkali plants. Their study showed that mercury slows memory-scanning efficiency. In their binary classification task, subjects were asked to memorize a short list of digits (two, three, or five digits long). A test digit was then shown. The subject indicated as quickly as possible whether the test digit had or had not been included in the previously memorized list. The time required for this decision was the primary measure, as accuracy is generally quite high. This reaction time typically increases in a linear fashion as a function of set size, with the slope of the function taken to represent efficiency of memory scanning and the extrapolated reaction time for a list of zero length being taken as

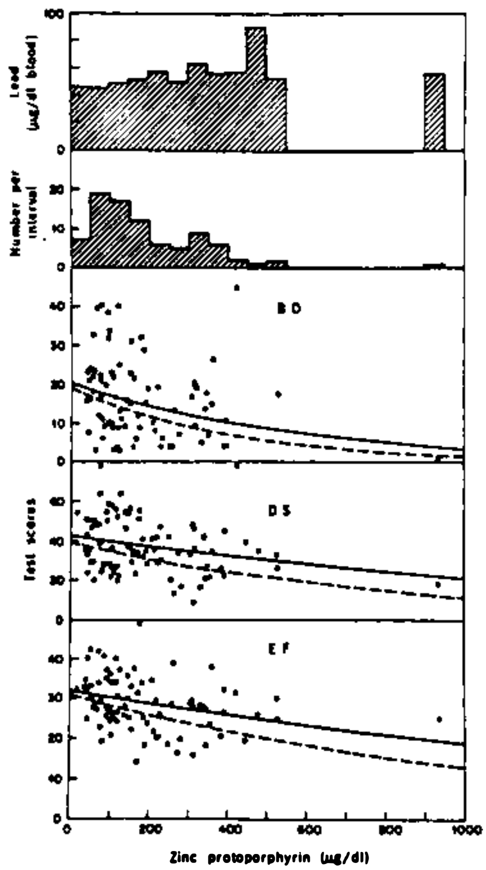


FIGURE 2. Age-corrected scores for 90 lead smelter workers on three performance tests (BD, DS, and EF) as functions of their individual ZPP levels. Solid curve, exponential (Eq. 2) fitted to these points. Dashed curves, exponentials fitted to the uncorrected scores; they show that the correction is not large. Hatched histograms give the distribution of ZPP levels for the test population and the blood lead levels corresponding to 50-µg/dl concentrations of ZPP. (Figure taken from Valcuikas et al., 1978. See that reference for an explanation of equation 2 and statistical parameters.)

an index of stimulus registration and response output time. A difference in the function for trials in which the test probe had not been included in the set is taken to measure response bias. To establish mercury levels, mercury in the urine was measured routinely every 3 months, and an average of several measurements was taken. Data were analyzed by forward, stepwise regression, with mercury level, the three performance measures (slope, intercept, and bias), and biographical indexes entered. Mercury levels and biographical indexes were not significantly correlated, reducing the likelihood of confounding variables. Scanning time and mercury level were found to be related (see Fig. 3). The intercept was not related to mercury level, and response type was only marginally related, an effect that the authors did not consider interpretable. The effect of a small increase in mercury in the urine was not only statistically but also clinically significant, an increase in scanning time of 12 ms per item for an increase of 0.1 mg/liter. This is roughly equivalent to 9 years of aging. This decrease is also accompanied by a decrease in digit span (Smith et al., 1985).

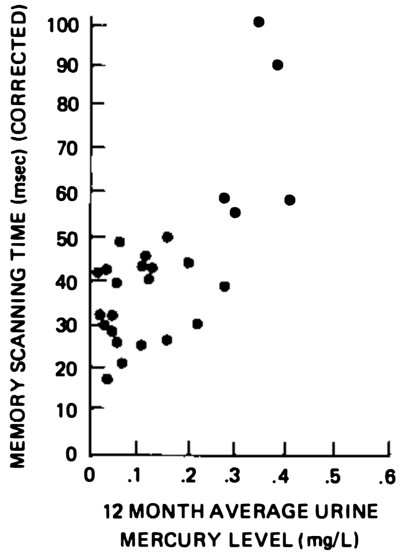


FIGURE 3. Relationship between memory-scanning time and mercury levels in urine. (Figure taken from Smith and Langolf, 1981.)

Two Approaches to Testing the Effects of Food Additives

Consecutive articles in a 1980 issue of Science dealt with the effects of food additives on children. In the first study (Swanson and Kinsbourne, 1980), 40 children were studied during a 5-day stay in a hospital clinic. All subjects had been referred with behavioral symptoms suggesting hyperactivity; half had high Conner's Rating Scale scores (average, 16.2) and had previously shown a favorable response to stimulant medication (these were designated hyperactive children), and half had lower Conner's scores (average, 12.3) and had shown an adverse response to stimulant medication (designated nonhyperactive children). The day before hospitalization and for the first 3 days of their hospital stay, any medication was stopped and an additive-free (Feingold) diet was strictly maintained. "On days 4 and 5, the children received, orally at 10:00 a.m., capsules that contained either a blend of nine food dyes [either 100 or 150 mg] or placebo (sugar), with dye and placebo order counterbalanced across subjects" (Swanson and Kinsbourne, 1980). Paired-associate learning tests were administered on each of the 5 days at 9:30, 10:30, and 11:30 a.m. In this test, pictures of animals (different each day) were used as stimuli, and numbers were used as responses. The stimuli were presented sequentially until a perfect recitation was obtained (10 s was allowed for a response, order was different in each presentation). The list length was adjusted over days 1 to 3 for each child, to a level which produced about 20 errors per test. The measure taken was the total number of errors made. For hyperactive children, there was a clear increase in errors after dye challenge (see Fig. 4). The effect reached a maximum at 1.5 h, consistent with the notion of a pharmacological or toxic mechanism. For nonhyperactive children, no clear effect of the dye challenge was seen, and errors were consistently at or below the levels seen for the hyperactive children. The apparent increase in errors across time of testing in this group was not commented on, except to note that challenge and testing time did not interact. The data thus demonstrate a toxic effect of food additives for some children, here identifiable by their common hyperactivity and reaction to stimulant therapy.

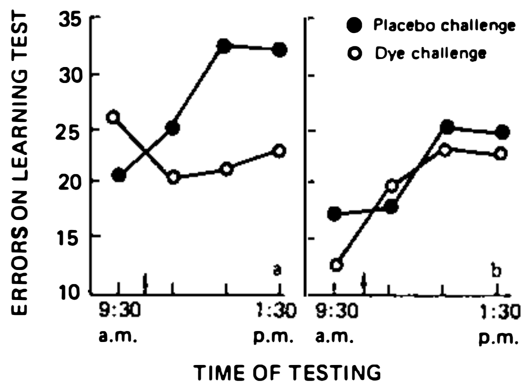


FIGURE 4. Responses of 20 hyperactive children (a) and 20 nonhyperactive children (b) challenged with a food dye blend or a placebo. (Figure taken from Swanson and Kinsbourne, 1980.)

Weiss et al. (1980) evaluated 22 children who were enrolled in the Kaiser-Permanente Health Maintenance Organization. None of these children had been diagnosed as hyperkinetic, and none had clinically significant medical or psychiatric problems. Each, however, had been reported as improved in behavior "when kept on a diet that excluded artificial colors and flavors for at least the 3-month period preceding the study." The study was carried out over 77 observation days, using a highly structured observation schedule that was highly individualized for each child. Two 15-min observation periods were conducted by a parent each day, with any occurrence of several target behaviors chosen for the child being recorded. A global estimate of the frequency and severity of each target behavior and several other general reactions to the day were also obtained. Daily telephone interviews with a parent and weekly home visits helped maintain stable criteria for observations. The children were kept on the additive-free diet throughout the study. At a specified time on each day, however, the child consumed a bottle of soft drink containing either caramel or cranberry coloring (placebo) or a freeze-dried monoblend of seven dyes plus the cranberry coloring (challenge). The challenge drink was given on 8 days, with both child and parent blind to which days. Whereas 20 of the 22 children "displayed no convincing evidence of sensitivity to the color challenge" (Weiss et al., 1980, p. 1488), two children demonstrated convincing sensitivity. One 34-month-old girl showed especially clear effects across several target behaviors (Fig. 5), with the consistency of the effects strikingly confirmed through statistical tests. It is clear that individual sensitivities need to be taken into account to adequately characterize the effects of food additives. Group mean statistics drastically underestimate the effects for sensitive individuals.

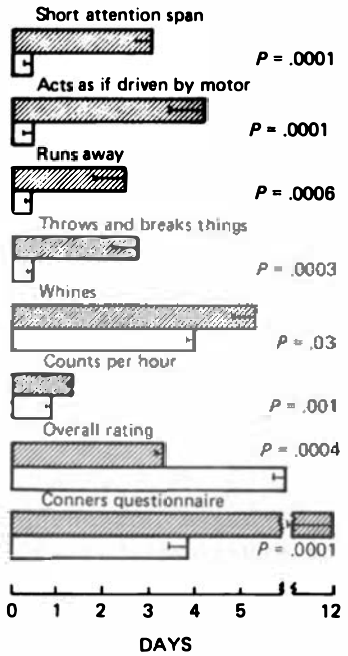


FIGURE 5. Mean total day ratings on target behaviors for one subject plus corresponding global scores. Open bars, mean % SE for placebo days; hatched bars, mean % SE for challenge days. The 15-min tallies recorded ca. 3 h after drink consumption were significantly higher on challenge than on control days for behaviors 2 (P = 0.003) and 3 (P = 0.016). (Figure taken from Weiss et al., 1980.)

REFERENCES

Allen, G.L., and B.B. Morgan. 1985. Assessment of learning abilities using complex experimental learning tasks. Neurobehav. Toxicol. Teratol. 7:355-358.

Alluisi, E.A., W.R. Beisel, P.H. Bartelloni, and G.D. Coates. 1973. Behavioral effects of tularemia and sandfly fever in man. J. Infect. Dis. 128:710-717.

Anger, W.K. 1982. Neurobehavioral testing in the workplace. Paper presented at the American Association for the Advancement of Science meeting, 3-8 January 1982, Washington, D.C.

Anger, W.K. In press. Workplace exposures. In Z. Annau, ed. Behavioral Toxicology. Johns Hopkins University Press, Baltimore, MD.

Anger, W.K., and B.O. Johnson. 1984. Chemicals affecting behavior. Pp. 51-149 in J.L. O'Donoghue, ed. Neurotoxicity of Industrial and Commercial Chemicals. CRC Press, Boca Raton, Fla.

Baker, E.L., R. Letz, A.T. Fidler, S. Shalat, D. Plantamura, and M. Lyndon. 1985. A computer-based neurobehavioral evaluation system for occupational and environmental epidemiology: methodology and validation studies. J. Neurobehav. Toxicol. Teratol. 7:369-378.

Beatty, W.W., and D.A. Shavalia. 1980. Spatial memory in rats: time course of working memory and effect of anesthetics. Behav. Neur. Biol. 28:454-462.

Birnbaum, I.M., and E.S. Parker. 1977. Acute effects of alcohol on storage and retrieval. In I.M. Birnbaum and E.S. Parker, eds. Alcohol and Human Memory. Lawrence Erlbaum Associates, Hillsdale, N.J.

Bittner, A.C., Jr., R.C. Carter, R.S. Kennedy, M.M. Harbeson, and M. Krause. 1984. Performance evaluation tests for environmental research (PETER): the good, bad, and ugly. Paper presented at Human Factors Society Meeting, San Antonio, Tex., October 1984.

Block, R.I., and R. Berchou. 1984. Alprazolam and lorazepam effects on memory acquisition and retrieval processes. Pharmacol. Biochem. Behav. 20:233-241.

Bornschein, R., D. Pearson, and L. Reiter. 1980. Behavioral effects of moderate lead exposure in children and animal models. I. Clinical studies. Crit. Rev. Toxicol. 8:43-99.

Branconnier, R.J. 1985. Dementia in human populations exposed to neurotoxic agents: a portable microcomputerized dementia screening battery. J. Neurobehav. Toxicol. Teratol. 7:379-386.

Branconnier, R.J., D.R. DeVitt, J.O. Cole, and K.F. Spera. 1982. Amitriptyline selectively disrupts verbal recall from secondary memory of the normal aged. Neurobiol. Aging, Vol. 3.

Brown, G.G., and R.K. Nixon. 1979. Exposure to polybrominated biphenyls: some effects on personality and cognitive functioning. J. Am. Med. Assoc. 242:523-527.

Brown, G.G., R.C. Preisman, M.D. Anderson, R.K. Nixon, J.L. Isbister, and H.A. Price. 1981. Memory performance of chemical workers exposed to polybrominated biphenyls. Science 212:1413-1415.

Bruce, D., and C.N. Cofer. 1965. A comparison of recognition and recall in short-term memory. Proc. Am. Psychol. Assoc. 20:81-82.

Cabe, P.A., and D.A. Eckerman. 1982. Assessment of learning and memory dysfunction in agent-exposed animals. Pp. 133-198 in C.L. Mitchell, ed. Nervous System Toxicology. Raven Press, New York.

Caine, E.D., H. Weingartner, C.L. Ludlow, E.A. Cudahy, and S. Wehry. 1981. Qualitative analysis of scopolamine-induced amnesia. Psychopharmacology 74:74-80.

Carroll, J.B. 1980. Individual different relations in psychometric and experimental cognitive tasks. L.L. Thurstone Psychometric Laboratory Report No. 163. NTIS Doc. No. AD-A086-057. The University of North Carolina, Chapel Hill, N.C. 315 pp.

Crook, T., S. Ferris, and M. McCarthy. 1979. The misplaced-objects task: a brief test for memory dysfunction in the aged. J. Am. Geriat. Soc. 27:284-287.

Crook, T., S. Ferris, M. McCarthy, and D. Rae. 1980. Utility of digit recall tasks for assessing memory in the aged. J. Consult. Clin. Psychol. 48:228-233.

Crook, T., S. Ferris, G. Sathananthan, A. Raskin, and S. Gershon. 1977. The effect of methylphenidate on test performance in the cognitively impaired aged. Psychopharmacology 52:251-255.

Crook, T., J.G. Gilbert, and S. Ferris. 1980. Operationalizing memory impairment for elderly persons: the Guild Memory Test. Psychol. Rep. 47:1315-1318.

Davis, K.L., L.E. Hollister, J. Overall, A. Johnson, and K. Train. 1976. Physostigmine: effects on cognition and affect in normal subjects. Psychopharmacology 51:23-27.

Davis, K.L., R.C. Mohs, J.R. Tinklenberg, A. Pfefferbaum, L.E. Hollister, and B.S. Kopell. 1978. Physostigmine: improvement of long-term memory processes in normal humans. Science 201:272-274.

Desjardins, P.J., J.M. Moerschbaecher, D.M. Thompson, amd J.R. Thomas. 1982. Intravenous diazepam in humans: effects on acquisition and performance of response chains. Pharmacol. Biochem. Behav. 17:1055-1059.

Eckerman, D.A., J.B. Carroll, D. Foree, C.M. Gullion, M. Lansman, E.R. Long, M.B. Waller, and T.S. Wallsten. 1985. An approach to brief field testing of neurotoxicity. J. Neurobehav. Toxicol. Teratol. 7:387-394.

Feldman, R.G., N.L. Ricks, and E.L. Baker. 1980. Neuropsychological effects of industrial toxins: a review. Am. J. Ind. Med. 1:211-227.

Ferris, S.H., G. Sathananthan, S. Gershon, and C. Clark. 1977. Senile dementia: treatment with deanol. Am. Geriat. Soc. 25:241-244.

Ferris, S.H., T. Crook, E. Clark, M. McCarthy, and D. Rae. 1980. Facial recognition memory deficits in normal aging and senile dementia. J. Gerontol. 35:707-714.

Ferris, S.H., B. Reisbert, T. Crook, E. Friedman, M.K. Schneck, P. Mir, K.A. Sherman, J. Corwin, S. Gershon, and R.T. Bartus. 1982. Pharmacologic treatment of senile dementia: choline, L-dopa, piracetam, and choline plus piracetam. Aging 19:475-481.

Gallagher, M., R.A. King, and N.B. Young. 1983. Opiate antagonists improve spatial memory. Science 221:975-976.

Ghoneim, M.M., S.P. Memwaldt, and J.V. Hinrichs. 1984a. Dose-response analysis of the behavioral effects of diazepam. I. Learning and memory. Psychopharmacology 82:291-295.

Ghoneim, M.M., S.P. Memwaldt, and J.V. Hinrichs. 1984b. Dose-response analysis of the behavioral effects of diazepam. II. Psychomotor performance, cognition, and mood. Psychopharmacology 82:296-300.

Godding, P.R., J.R. Rush, and W.W. Beatty. 1982. Scopolamine does not disrupt spatial working memory in rats. Pharmacol. Biochem. Behav. 12:919-923.

Hanninen, H., and K. Lindstrom. 1979. Behavioral test battery for toxicopsychological studies. Used at the Institute of Occupational Health in Helsinki. 2nd revised ed. Reviews 1. Institute of Occupational Health, Helsinki. 51 pp.

Harbeson, M.M., A.C. Bittner Jr., R.S. Kennedy, R.C. Carter, and M. Krause. 1983. Performance evaluation tests for environmental research (PETER): bibliography. Percept. Motor Skills 57:283-293.

Hawk, B., and S.R. Schroeder. 1983. Factors interactive with IQ and blood lead levels in children. Paper presented at the Gatlinburg Conference on Mental Retardation and Developmental Disabilities, Gatlinburg, Tenn., 15 June 1983.

Horn, J.L. 1976. Human abilities: a review of research and theory in the early 1970's. Annu. Rev. Psychol. 27:437-485.

Hunt, E. 1983. On the nature of intelligence. Science 219:141-146.

Hunt, E., and C. Hertzog. 1981. Age related changes in cognition during the working years. Final Report, Office of Naval Research Contract N00014-70-C-0345. [Unpublished]

Irons, R.C., and P.V. Rose. 1985. Some automated tests developed at the Naval Biodynamics Laboratory. J. Neurobehav. Toxicol. Teratol. 7:395-398.

Johnson, B.L., and W.K. Anger. 1983. Behavioral toxicology. Pp. 329-350 in W.N. Rom, ed. Environmental and Occupational Medicine. Little, Brown & Co., Boston.

Johnson, B.L., H.H. Cohen, R. Struble, J.V. Setzer, W.K. Anger, B.D. Gutnik, T. McDonough, and T. Hauser. 1974. Field evaluation of carbon monoxide exposed toll collectors. P. 306 in C. Xintaras, B.L. Johnson, and I. de Groot, eds. Behavioral Toxicology: Early Detection of Occupational Hazards. DHEW Publication No. 74-126, National Institute for Occupational Safety and Health, Department of Health, Education, and Welfare, Washington, D.C. 306 pp.

Jones, B.M., and M.K. Jones. 1977. Alcohol and memory impairment in male and female social drinkers. In I.M. Birnbaum and E.S. Parker, eds. Alcohol and Human Memory. Lawrence Erlbaum Associates, Hillsdale, N.J.

Kennedy, R.S., A.C. Bittner Jr., M.G. Smith, and M.M. Harbeson. 1984. Development of a portable performance assessment system for behavioral toxicology. Paper presented at the Behavioral Toxicology Society meeting, Toronto, Canada, 30 August 1984.

Knave, B., B.A. Olson, S. Elofson, F. Gambrale, A. Isaksson, P. Mindus, H.E. Persson, G. Struwe, A. Wennberg, and P. Westerholm. 1978. A cross-sectional epidemiologic investigation on occupationally exposed industrial workers with special reference to the nervous system. Scand. J. Work Environ. Health 4:19-45.

MacCleod, C.M., A.S. Dekaban, and E. Hunt. 1978. Memory impairment in epileptic patients: selective effects of phenobarbital concentration. Science 202:1102-1104.

Maki, W.S., S. Brokofsfy, and B. Berg. 1979. Spatial memory in rats: resistance to retroactive interference. Anim. Learn. Behav. 7:25-30.

McCarthy, M., S.H. Ferris, E. Clark, and T. Crook. 1981. Acquisition and retention of categorized material in normal aging and senile dementia. Exp. Aging Res. 7:127-135.

Milar, C.R., S.R. Schroeder, R. Muchak, J.L. Dolcourt, and L.D. Grant. 1980. Contributions of the care-giving environment to increased lead burden of children. Am. J. Mental Deficiency 84:339-344.

Miller, D.B., and D.A. Eckerman. In press. Learning and memory measures. In Z. Annau, ed. Behavioral Toxicology. Johns Hopkins University Press, Baltimore, Md.

Miller, D.B., D.A. Eckerman, M.R. Krigman, and L.D. Grant. 1982. Chronic neonatal organotin exposure alters radial-arm maze performance in adult rats. Neurobehav. Toxicol. Teratol. 4:185-190.

Miller, E., and P. Lewis. 1977. Recognition memory in elderly patients with depression and dementia: a signal detection analysis. J. Abnorm. Behav. 86:84-86.

Mohs, R.C., and K.L. Davis. 1980. Choline chloride effects on memory: correlation with the effects of physostigmine. Psychiatry Res. 2:149-156.

Mohs, R.C., J.R. Tinklenberg, W.T. Roth, and B.S. Kopell. 1978a. Slowing of short-term memory scanning in alcoholics. J. Studies Alcohol 39:1908-1919.

Mohs, R.C., J.R. Tinklenberg, W.T. Roth, and B.S. Kopell. 1978b. Methamphetamine and diphenhydramine effects on rate of cognitive processing. Psychopharmacology 59:13-19.

Mohs, R.C., J.R. Tinklenberg, W.T. Roth, and B.S. Kopell. 1980. Sensitivity of some human cognitive functions to effects of methamphetamine and secobarbital. Drug Alcohol Dependence 5:145-150.

Otto, D.A. 1983. The application of event-related low brain potentials in occupational medicine. In R. Gilioli, ed. Neurobehavioral Methods in Occupational Health. Pergamon Press, New York.

Otto, D.A., S.B. Bauman, G.S. Robinson, E. Donchin, and E. Heffley. 1985. Application of a portable microprocessor-based system for electrophysiological field testing of neurotoxicity. Neurobehav. Toxicol. Teratol. 7:409-414.

Putz, V.R. 1979. The effects of carbon monoxide on dual task performance. Hum. Factors 21:13-24.

Russell, E.W. 1979. Three patterns of brain damage on the WAIS. J. Clin. Psychol. 35:611-620.

Russell, E.W. 1980. Fluid and crystallized intelligence: effects of diffuse brain damage on the WAIS. Percept. Motor Skills 51:121-122.

Schaumburg, H.H., J.C. Arezzo, L. Markowitz, and P.S. Spencer. 1981. Neurotoxicity assessment. Pp. 81-91 in W.W. Lowrange, ed. Assessment of Health Effects at Chemical Disposal Sites. The Rockefeller University, New York.

Seppalainen, A.M., K. Lindstrom, and T. Martelin. 1980. Neurophysiological and psychological picture of solvent poisoning. Am. J. Ind. Med. 1:31-42.

Seppalainen, A.M., H. Hanninen, W.K. Anger, E.L. Baker, R. Carter, N. Cherry, R. Dick, R. Feldman, R. Letz, D. Otto, and D. Vinson. In press. Report from the work group on recommended test methods on prevention of neurotoxic illness in working populations. National Institute for Occupational Safety and Health, Department of Health and Human Resources, Washington, D.C.

Sitram, N., H. Weingartner, E.D. Caine, and J.C. Gillin. 1977. Choline: selective enhancement of serial learning and encoding of low imagery words in man. Life Sci. 22:1555-1560.

Sitram, N., H. Weingartner, and J.C. Gillin. 1978. Human serial learning: enhancement with arecholine and choline and impairment with scopolamine. Science 201:274-276.

Smith, P.J., and G.D. Langolf. 1981. The use of Sternberg's memory-scanning paradigm in assessing effects of chemical exposure. Hum. Factors 23:701-708.

Smith, P.J., G.D. Langolf, and J. Goldberg. 1983. Effects of occupational exposure to elementary mercury on short-term memory capacity. Br. J. Ind. Med. 40:413-419.

Squire, L.R., L.L. Judd, D.S. Janowsky, and L.Y. Huey. 1980. Effects of lithium carbonate on memory and other cognitive functions. Am. J. Psychiatry 37:1042-1046.

Sterman, A.D., and A. Varma. 1983. Evaluating human neurotoxicity of pesticide aldicarb: when man becomes the experimental animal. Neurobehav. Toxicol. Teratol. 5:493-495.

Sternberg, R. 1977. Intelligence, information processing, and analogical reasoning. Erlbaum, Hillsdale, N.J.

Sunderland, A., J.E. Harris, and A.D. Baddeley. 1983. Do laboratory tests predict everyday memory?: a neuropsychological study. J. Verb. Learn. Verb. Behav. 22:341-357.

Swanson, J.M., and M. Kinsbourne. 1980. Food dyes impair performance of hyperactive children on a laboratory learning test. Science 207:1485-1486.

Thorne, D.R., S.G. Genzer, H.C. Sing, and F.W. Hegge. 1985. The Walter Reed performance assessment battery. J. Neurobehav. Toxicol. Teratol. 7:415-418.

Tuttle, T.C., G.D. Wood, C.B. Grether, B.L. Johnson, and C. Xintaras. 1977. Behavioral and neurological evaluation of workers exposed to carbon disulfide. DHEW Publication No. 77-128. National Institute for Occupational Safety and Health, Department of Health, Education, and Welfare, Washington, D.C. 79 pp.

Valcuikas, J.A., R. Lilis, A. Fischbein, I.J. Selikoff, J. Eisinger, and W.E. Blumber. 1978a. Central nervous system dysfunction due to lead exposure. Science 201:465-467.

Valcuikas, J.A., R. Lilis, M.S. Wolff, and H.A. Anderson. 1978b. Comparative neurobehavioral study of a polybrominated biphenyl-exposed population in Michigan and a nonexposed group in Wisconsin. Environ. Health Perspect. 23:199-210.

Weingartner, H., N. Sitram, and J.C. Gillin. 1979. The role of the cholinergic nervous system in memory consolidation. Bull. Psychonom. Soc. 13:9-11.

Weingartner, H., P. Gold, J.C. Ballenger, S.A. Smallberg, R. Summers, D.R. Rubinow, R.M. Post, and F.K. Goodwin. 1981. Effects of vasopressin on human memory functions. Science 211:601-603.

Weingartner, H., J. Graftman, W. Boutelle, W. Kaye, and P.R. Martin. 1983. Forms of memory failure. Science 221:380-382.

Weiss, B. 1980. Conceptual issues in the assessment of lead toxicity. Pp. 127-134 in H.L. Needleman, ed. Low Level Lead Exposure: The Clinical Implications of Current Research. Raven Press, New York.

Weiss, B. 1983. Behavioral toxicology and environmental health science: opportunity and challenge for psychology. Am. Psychol. 38:1174-1187.

Weiss, B., J.H. Williams, S. Margen, B. Abrams, B. Caan, L.J. Citon, C. Cox, J. McKibben, D. Ogar, and S. Schultz. 1980. Behavioral responses to artificial food colors. Science 207:1487-1488.

Williams, H.L., and O.H. Rundell. 1983. Secobarbital effects on recall and recognition in a levels-of-processing paradigm. Psychopharmacology 80:221-225.

Xintaras, C., and J.A.P. Burg. 1980. Screening and prevention of human neurotoxic outbreaks: issues and problems. Pp. 663-674 in P.S. Spencer and H.H. Schaumburg, eds. Experimental and Clinical Neurotoxicology. The Williams & Wilkins Co., Baltimore, Md.

SESSION III

INTRODUCTION--METHODS OF ASSESSING MILITARY PERFORMANCE

The purpose of this session was to present an overview of the principal methods currently used to assess the performance of military personnel. The methods reviewed are the use of simulators, task analysis of military jobs, field studies, field exercises, and cognitive test batteries. Speakers with first-hand knowledge of each of these methods discussed how they are currently used by the military. Factors that might affect the suitability of each method for studying nutritional variables are reviewed; these include cost, complexity, ability to control extraneous variables, precision of measurement, and validity. Several cognitive test batteries are now being developed to assess military personnel, and a program to develop a battery for assessing drug effects is presented. Three of the batteries previously developed for the military are reviewed since many of the problems encountered in developing and validating these batteries are likely to be encountered in a program of research to assess nutritional effects.

THE USE OF SIMULATORS IN THE MILITARY

Kent A. Kimball

A review of the large volume of literature on simulators and simulation applications in military settings reveals that simulation techniques have been used in the military training environment for many years. In fact, military simulation has been used at least since feudal times, as evidenced by the war games that became popular in Europe about A.D. 1200. Jousts were held to hone combat skills and in effect were simulations of combat. There were also joust simulators. This device, called a quintain, provided instant feedback to the trainee. If it was not struck appropriately, the device rotated quickly to deliver a blow to the head of its unwary assailant (Tunis, 1954). A more recent but still historic example of simulation would be the photogun developed by Eastman Kodak to train aerial gunners during World War I. This device was a spring-wound, motorized camera mounted on the Lewis machine gun. When the gunner pulled the trigger, film was exposed and a constant record of aim-point was obtained. Mock aerial battles could be accomplished, and "kills" could be recorded safely and provided to student aviators for instructive purposes (Stockbridge, 1920).

Since the use of simulation varies widely in the military setting, it is necessary to narrow the scope. First I will focus on a common definition of what is meant by simulation, and then I will review some uses of simulation as applied to man-machine systems. After this discussion, the focus will narrow again to flight simulation, a brief history of its development, and some examples of systems that have served to enhance aircrew training. Finally, I will review the Army's present flight simulation and training program and discuss the potential for similar systems in biomedical human performance research.

A number of years ago, McCormick (1964) provided a working definition of simulation. He stated that simulation consists of some type of reproduction or representation of an actual or conceptual physical object, system, process, or situation or of a theoretical construct. This definition illustrates that the process of simulation is not limited to developing large, complex systems, but can be merely a line drawing or flow diagram of a process or a mathematical model of a concept.

McCormick (1964) delineated five areas in which simulation is applied in the development and function of man-machine systems. Many of us are or have been engaged in human performance research. In the laboratory and the field, examples of how an experimental paradigm involves some simulation of a real-world setting are found readily. The research interest may be human behavior, but how it generalizes to the operational environment is also of interest. Some examples of this type of experimentation would be a simple tracking task, a tracking task with auditory distraction, or a psychological assessment battery. In each case, we are assessing performance in the laboratory, i.e., the simulation, with the ultimate object of predicting human behavior in the real-world setting.

In the second category, simulation can be applied to the development and design phase of specific hardware systems. Again, depending on the design requirements, various levels of simulation can be used. Some examples of these applications are simple line drawings, software mock-ups, and complex prototypes. The application of simulation in this area allows designers to evaluate new systems and hardware before making final designs for production. This approach has often served to decrease design costs and improve ultimate system capability.

The use of simulation in operational problem solving involves situations in which various procedures, scenarios, and organizational structures are manipulated to develop optimal performance strategies. This application is typically oriented to the idea of gaming. McCormick (1964) points to war game simulations as an example of this application.

As systems proceed through the development cycle to completion, there are numerous stages at which simulation techniques can be used. When such systems reach maturity, they are normally tested again for reliability, maintainability, and effectiveness. Simulation normally plays a role in these performance evaluations. Some examples of how simulation has played a part in developmental testing are operational tests of engine longevity, such as those conducted by the Army's Test Activity, and the fly-off for the Army's UH60 Blackhawk. In each case, aircraft are flown against mission profiles normally anticipated during actual operations.

Certainly the most widely acknowledged area in which simulation is used has been the field of training. Although there are tremendous variations in the degree of simulation used, most man-machine systems of substantial complexity have simulator or procedural trainer counterparts to assist operators in acquiring and practicing their skills under controlled conditions. The most obvious examples of the application of simulation to training are a flight simulator and a mission simulator that would emulate the performance characteristics of a military aircraft.

According to Smith (1981), the first flight trainer used in the United States was developed by Ed Link in 1929. Mr. Link's trainer was the first system used to reduce the number of flight hours required to train civilian pilots. Mr. Link apparently had considerable success with his system, for by 1941 he had delivered his Link Trainers to the U.S. Army as well as to 35 other countries. As one might suspect, World War II saw a rapid expansion of the use of trainers for pilot proficiency. Smith states that there were some 30 different types of device in use by 1945.

By 1950, most major aviation weapons systems were provided with counterpart electronic simulators. These devices were typically without visual displays, but were used for training in systems operations, emergency procedures, instrument flight, and fire control. Updates and improvements in aircraft were not made in their simulator counterparts during this period. As a result, the use of simulation for pilot training was not vigorously pursued.

The fuel crisis in the 1970s increased interest in simulation for pilot evaluation and training. Increasing fuel costs, the increasing complexity of military weapons systems, and the potential for enhanced visual flight simulation as demonstrated by civilian airlines all served to place increased emphasis on simulators for flight-training programs for the armed services. This emphasis created a burgeoning simulator industry to meet both civilian and military requirements (Latour, 1978). As of the date of his report, there were no less than 222 simulators of civil aircraft in use in the Western world, most of which were owned by civilian airlines. Although figures are not available, it can be assumed that most military services are using simulators.

I would like now to narrow the focus to the U.S. Army's flight simulation program, particularly as it relates to training. Although simulation is being used elsewhere within the Army for both research and training, I will confine my discussions to flight simulation developments at the Army Aviation Center.

In a recent paper detailing progress in Army helicopter flight simulation, Siering (1983) points out that the Army placed a unique requirement on the simulation industry in that it required advanced rotary wing simulators for training purposes. Specifically, in 1967 the Army required that it be provided simulators to support five major helicopter systems: the UH1 (Huey), the CH47 (Chinook), the AH1 (Cobra), the UH60 (Blackhawk), and the AH64 (Apache) (U.S. Army Aviation Center, 1981).

The first system provided under this plan was the UH1 flight simulator (FS). This device was a fully instrumented motion base system with no outside visual display. This system permits four cockpits to be used simultaneously while a central control panel provides profile programming and in-flight performance data. The main use of this device is for training current aviators. Although these systems allow positive transfer of training (Caro et al., 1975) and skills maintenance (Weitzman et al., 1976), pilot performance measurement was not considered optimum.

The second simulator to be provided to the center was the CH47 FS. This device was equipped with platform motion and an on-board visual system. After considerable on-site prototype development, this system was introduced as a training device in the CH47 aviation qualification course and provides transition pilots with high-fidelity visual cockpit training.

Soon after the CH47 FS was implemented as a training device, the next simulator, the AH1 FS, simulating the Army's attack helicopter, was provided to the center. The AH1 FS comprised a motion system and an on-board camera system similar to the CH47 FS terrain board, except that it included a system to permit weapons use. Two model board displays were provided. This allowed separate visual scenarios for the pilot and the copilot gunner. This device is now being used for both aircraft qualification training and weapons currency training.

Most recently, the UH60 FS has become operational at the center. This simulator was delivered with two cockpits having separate but equivalent motion systems. The cockpits had two different visual systems: one a computer-generated image system and the other a camera model system. Both cockpits have been shown to be effective, providing good transfer of training for the UH60 aviator qualification course (Luckey et al., 1982).

The last aircraft system that is to have a counterpart simulator is the YAH64, the Army's advanced attack helicopter; the prototype is presently being constructed. It is anticipated that it will be delivered to the Aviation Center in 1985.

The flight simulation training program at the Army Aviation Center has achieved substantial success in enhancing proficiency and reducing aircraft qualification time for the Army aviator. Cost savings have been realized by using these devices (Drackett et al., 1982; Hopkins, 1979). How well do these devices assess flight proficiency or, more specifically, overall aviator performance? These devices were primarily designed to assist in training; as a consequence, emphasis was placed on fidelity as viewed by the instructor pilot and student. Thus, the data collection and analysis packages attached to the output side of these systems are somewhat constrained by both available hardware and software flexibility. However, these devices can be used effectively for crew performance studies. At present, the accomplishment of this objective in a cost-effective manner would require a "strap-on" data collection and management system.

For a number of years, personnel at the U.S. Army Aeromedical Research Laboratory have been collecting in-flight data with the UH1 helicopter (Huffman et al., 1972; Jones et al., 1983). A similar system has been developed and interfaced with a low-fidelity helicopter trainer in the laboratory. This system permits certain parameters to be measured in real time. Although the simulation device does not provide the fidelity desired for crew performance studies, we are presently using it for extended studies on aviator performance under the influence of antidotes and pretreatment therapies for chemical defense. Our future plans include the acquisition of a UH60 FS for these crew performance studies. We hope that the lessons we have learned in developing our more modest performance assessment procedures will permit us to provide comprehensive and effective performance assessment systems for the future.

REFERENCES

Caro, P.W., R.N. Isley, and O.B. Jolley. 1975. Mission suitability testing of an aircraft simulator. Human Resources Research Organization Technical Report 75-12. Human Resources Research Organization, Alexandria, Va. 69 pp.

Drackett, G.F., J.E. LaVerne, and L.D. Southard. 1982. UH-60 flight simulator (UH60FS) training development study. TRANSANA TDS-19-82. U.S. Army TRADOC Systems Analysis Activity. White Sands Missile Range, N.Mex.

Hopkins, F.C. 1979. Cost and training effectiveness analysis (CTEA) of the AH-1 flight simulator (AH1FS). U.S. Army Aviation Center, Fort Rucker, Ala.

Huffman, H.W., M.A. Hoffman, and M.R. Sleeter. 1972. Helicopter in-flight monitoring system. Report No. USAARL 72-11. U.S. Army Aeromedical Research Laboratory, Fort Rucker, Ala. 52 pp.

Jones, H.B., J.A. Lewis, and A.A. Higdon Jr. 1983. Helicopter in-flight monitoring system--second generation (HIMS II). Report No. USAARL 83-13. U.S. Army Aeromedical Research Laboratory, Fort Rucker, Ala. 51 pp.

Latour, C. 1978. Flight simulation. NATO's Fifteen Nations 23:76-81.

Luckey, J.R., W.R. Bickley, W.G. Maxwell, and E.A. Cerone. 1982. Operational test II, phase I, of the UH60A BLACKHAWK flight simulator (OTII, Phase I, UH60FS). Final test report. U.S. Army Aviation Board, Fort Rucker, Ala. 424 pp.

McCormick, E.J. 1964. Human Factors Engineering, 2nd Ed., pp. 599-613. McGraw-Hill, New York.

Siering, G.D. 1983. Progress in Army helicopter flight simulation. Pp. 615-620 in Richard S. Jensen, ed. Proceedings of the Second Symposium on Aviation Psychology, 26-27 April 1983, Columbus, Ohio. Ohio State University Press, Columbus.

Smith, J. 1981. Experience with flight simulators--training effectiveness and future developments. AFHRL-TP-81-14. Air Force Human Resources Laboratory, Brooks Air Force Base, Ariz. 28 pp.

Stockbridge, F.P. 1920. Yankee Ingenuity in the War. Harper and Brothers, New York. 320 pp.

Tunis, E. 1954. Weapons: A Pictorial History, pp. 46-47. World Publishing, New York.

U.S. Army Aviation Center. 1981. Revised training device requirement (TDR) 0027 for synthetic flight training system (rotary wing). U.S. Army Aviation Center, Fort Rucker, Ala.

Weitzman, D.O., M. Fineberg, H. Ozkaptan, and G.L. Compton. 1976. Evaluation of the synthetic flight training system (device 2B24) for maintaining IFR proficiency among experienced pilots. Pp. 63-68 in Proceedings of the Ninth NTEC/Industry Conference. Naval Training Equipment Center, Orlando, Fla.

HOW DOES NUTRITION RESEARCH IMPROVE MILITARY READINESS?

Stanley F. Bolin

A general officer opened the workshop on cognitive testing methodology for nutrition research by asking for help. He asked the participants to help defend the nutrition research budget by looking for ways to show how nutrition made a practical difference in military readiness. This challenge was not made lightly. In the continuing political struggles over the defense budget, many people may easily be convinced that nutrition research funds might better be spent on weapons or not spent at all.

The challenge was not met during the workshop. The general officer did not expect to be satisfied immediately, and the participants were not prepared to address such a large issue. This paper was prepared after the workshop with some appreciation of the concerns of nutrition research scientists and a belated understanding of why it had been invited in the first place. This paper will also fail to satisfy the challenge, but it will provide some information and suggest some steps that may be helpful in doing so.

Task analysis performed by the Army in support of training may serve nutrition research as a starting point for further analysis to identify sensitive measures organized by some theory of nutrition and behavior. Specifically, the tasks described in Army soldier's manuals can satisfy the need to document military importance, but considerable further research is needed to establish sensitive measures relevant to nutrition theory.

A direct effort to show an empirical connection between nutrition or any other behavioral variable and readiness training performance is not recommended. "Steel-on-target" and combat exchange ratios are influenced by too many variables, and there are too many controversies about the meaning of combat field exercises. Military history is too full of instances when ill-fed armies prevailed over well-fed opponents. Consider the American Revolution and the Korean action and the arguments suggested by such examples.

An alternative to direct impact evaluation in the real world is the demonstration of theoretical consequences that make a plausible difference against realistic contingencies in combat scenarios. With the hubris of one innocent of experience in nutrition research, I suggest theory building and research to show that well-nourished soldiers fight better and longer than they would if less well nourished. The outcome of the battle is another story and quite intractable from a research viewpoint.

SYSTEMS APPROACH TO TRAINING

A systems approach to training (SAT) was adopted by the Army (TRADOC, 1982) to ensure that all military training can be considered relevant and critical for mission accomplishment. SAT is based on the Interservice Procedures for Instructional Systems Development (TRADOC, 1975), a five-volume manual on the application of instructional system design (ISD) to military training. The ISD model is widely used in training technology (Gagné and Briggs, 1974; Kearsley, 1984). ISD moves through five phases: analyze, design, develop, implement, and evaluate. Each phase feeds information to every other phase in a network of system loops so that the whole system should be self-correcting and demonstrate progressive improvement and efficiency.

The first phase begins with task analysis oriented to supporting the entire process of ISD. The importance of task performance in unit missions and individual duty positions is examined. Training needs analysis is conducted, and the feasibility of training is considered. The possibilities of developing job aids to preclude training are explicitly recognized and programmed for the development phase. The products of task analysis are performance specifications and measures and terminal learning objectives for critical tasks. These products are published in soldier's manuals for almost all military occupational specialties.

The task analysis methods developed in and for the Army are remarkably thorough in documenting the judgments involved. The Job and Task Analysis Handbook (TRADOC, 1979) supplements the ISD volumes with how-to-do-it flowcharts, forms, and detailed guidelines, but there are several missing chapters. These missing elements are "Soft Skill Analyis," "Collective/Individual Analysis Interface," and "Life Cycle Systems Management/Training Developments Interface." These topics need not be covered in detail to support the current SAT regulation because the ISD model is nonregulatory. These seeming holes in the task analysis methods are being satisfied by delegation to Army schools and responsible commanders. The implication for research based on soldier's manual information is that the task lists may not include some complex and difficult to analyze tasks, or if listed, such tasks may be treated summarily.

The significant feature of SAT task analysis for nutrition research is the nature of the critical task information. The SAT regulation (TRADOC, 1982, p. 29) describes it well:

Critical task identification. This activity entails identifying collective and individual tasks that are considered essential to the accomplishment of a mission or duty, and then determining those which require training, i.e., critical tasks for training. The products of critical task identification are critical task lists accompanied by reports on the methodology utilized.

Critical task analysis. This analysis consists of the following two activities:

Develop task performance specifications (TPS). TPS are developed by analysts with the assistance of subject matter experts. The specifications consist of detailed descriptions of the overall activities (to include cues and performance elements and steps), the conditions of collective and individual task performance in a mission environment or duty position, and the desired results and standards. TPS enable the development of task performance measures.

Develop task performance measures (TPM). In many cases and for a number of reasons, it is not possible to reproduce in the training situation the actual conditions of task performance in a mission environment or duty position. To overcome this problem, TPM are developed. A TPM is an analyst's prescription of most feasible performance measure. . . . There is no requirement to develop TPM for those tasks for which the actual performance and conditions in the mission environment or duty position can be replicated in the training situation. In those cases, the terminal learning objectives (TLO), developed during the 'design' phase, can be derived directly from the TPS.

SOLDIER'S MANUALS AND DATA SOURCES

Soldier's manuals can be obtained through military publication channels or from the Government Printing Office. They are listed as field manuals (FM) with number series specific to each military occupational specialty. Since SAT is nonregulatory and decentralized, there are variations in practice by the various Army school commands responsible for producing and updating these manuals and related reading materials. Some soldier's manuals are very recently published; some are several years old. All contain tasks, conditions, and standards of performance presented to enlisted personnel and their superiors to guide self-study and training.

The idea of ordering a complete set of all soldier's manuals is impractical unless one is prepared to handle a library. There are about 400 military occupational specialties (MOS), and many MOS have two or three soldier's manuals to cover skill levels. The soldier's manual of common tasks, generic to all soldiers, and a small set of MOS in fields of research interest are suggested to gain familiarity with this source.

Each of some 26 Army schools has a director of training developments and staff responsible for pursuing SAT. Task analysis details and evaluation results are not centrally aggregated, to my knowledge, so inquiry to particular schools may be necessary. Field visits will probably be necessary since SAT was adopted to reduce the burden of paperwork required by ISD.

The Individual Training Evaluation Directorate, Army Training Support Center, Ft. Eustis, Va., is a useful source of information about skill qualification testing (SQT). Written SQTs are now used to support personnel management and evaluate individual training insofar as it can be evaluated with written tests. Although SQT no longer includes hands-on performance testing, these tests of essential task knowledge have been validated, and the scores are moderately correlated with soldier quality measures such as mental category and civilian education. This directorate can also provide current information on the nature of score distributions and their reliability (Bolin et al., 1981).

When the SQT included hands-on performance testing and job site performance measures (before SAT), score distributions for these two kinds of measures were commonly observed to be highly skewed in the direction of high to very high success rates. These observations were expected because criterion-referenced testing for training promotes mastery and because job site measures are subject to the inflation of supervisory ratings. The written tests of SQT then provided the most variance (Bolin, 1981; Harman, 1981).

FURTHER TASK ANALYSIS NEEDS

Task analysis in the context of ISD is never finished; it is merely stopped or frozen temporarily to get on with the other phases of ISD. The processes of training and evaluation are expected to inform the task analysts of initial errors. Corrections in design and development phases then lead to another run through implementation, training, and evaluation. Iteration through the system continues until some point of diminishing returns is reached. That ISD was ever attempted on a full scale by a military organization is simply amazing. That the necessary iterations could not be accomplished because of personnel turnover and other factors is a common observation among participants in the effort. Yet the resulting soldier's manuals and the concepts of SAT provide a starting point and framework for further research. The first need is for more and better task analysis.

Task analysis for ISD is ad hoc; it serves the purposes of training and instruction. There is no real technology of task analysis as a separate discipline or generalized procedure because every kind of task analysis is tailored to its purpose and its audience (Fine and Wiley, 1971; Mager, 1972; McCormick, 1979; Pajer, 1979). Task analysis for wage and salary administration works at the level of job descriptions to develop rating factors for clusters of duty positions similar enough to be called a job; the objective is to ensure that similar jobs get similar pay and that a linear relationship holds between pay and job levels within an organization. Task analysis for training requires what Gagné (1977) calls learning analysis to identify the domains of human capability: intellectual skill, cognitive strategy, verbal information, motor skill, and attitudes. Gagné's concept of cumulative learning suggests that cognitive strategies are learned by building on the other four domains in problem solving given unique or unusual conditions.

If military commanders want soldiers who fight smart, in the sense of adapting doctrine and weapons systems to the circumstances of mission, enemy, and terrain, then cognitive strategies for tactical problem solving are high-value training goals.

Problem solving of this kind is not likely to be found in soldier's manuals because by definition problem solving under unique conditions cannot be broken down into set steps. Complex tasks with large variations in conditions and solving steps were generically called soft skills during the development of the job and task handbook for ISD. Included among these neglected skills were leadership and interpersonal relations (making judgment and problem solving seem even softer to some minds). Behavioral analysis was proposed and developed as a potential technique for coping with these nonprocedural tasks, but task analysts rejected the approach on the grounds of complexity.

An extended task analysis procedure (ETAP) was developed specifically to deal with this problem (Reigeluth and Merrill, 1980). ETAP is designed to provide a learning analysis for the whole range of tasks. It begins with procedural analysis and moves to knowledge analysis. Soft skills are handled by introducing two kinds of transfer of training ideas as cognitive strategies, called factor transfer and principle transfer. Problems presenting variations in conditions and constraints on performance may be solved by weighting factors differently in different situations. Problems requiring variations in general procedures or unusual applications of old principles may be solved by principle transfer. Finally, ETAP suggests that mixed cases require combining procedural analysis with factor and principle transfer. The method has not been tested beyond the developmental work, but it seems promising and was acceptable to task analysts at TRADOC.

If further task analysis is conducted in support of nutrition research, it would be natural to tailor the work to specific research interests. A theory of the relationship between nutrition and behavior would guide selection of MOS and tasks within that MOS. Physical fitness and endurance might well be important. In view of the efforts being made in selection, training, and human engineering to reduce the skill demands of new weapon systems, the most important task of future soldiers may be decision making or tactical problem solving. From this optimistic perspective on the development of smart weapons, I suggest that nutrition research might seek to show that well-fed soldiers fight smarter and longer in the physical sense than less well fed soldiers.

REFERENCES

Bolin, S.F. 1981. A task-based approach to soldier training. Defense Manage. J. 18:24-28.

Bolin, S.F., R. Eastman, J. Sova, J. Harman, J. Kessler, and D. Macpherson. 1981. Skill qualification testing--an evolving system. Pp. 1355-1390 in Proceedings of the 23rd Annual Military Testing Association Conference. Military Testing Association, Arlington, Va.

Fine, S.A., and W.W. Wiley. 1971. An Introduction to Functional Job Analysis. W.E. Upjohn Institute for Employment Research, Kalamazoo, Mich.

Gagné, R.M. 1977. The Conditions of Learning. Holt, Rinehart & Winston, New York.

Gagné, R.M., and L.J. Briggs. 1974. Principles of Instructional Design. Holt, Rinehart & Winston, New York.

Harman, J. 1981. Pretesting the skill component. Infantry 71:35-38.

Kearsley, G. 1984. Training and Technology. Addison Wesley, Reading, Mass.

Mager, R.F. 1972. Goal Analysis. Fearon, Belmont, Calif.

McCormick, E.J. 1979. Job Analysis. AMACOM, New York.

Pajer, R.G. 1979. How to develop job-related personnel selection procedures. Publication OIPP, 152-1140. Office of Personnel Management, Washington, D.C.

Reigeluth, C.M., and M.D. Merrill. 1980. Extended Task Analysis Procedure. Army Training Development Institute, Ft. Monroe, Va.

TRADOC (Training and Development Command). 1975. Interservice Procedures for Instructional Systems Development. Pamphlet 350-30. TRADOC, Ft. Monroe, Va.

TRADOC (Training and Development Command). 1979. Job and Task Analysis Handbook. Pamphlet 351-4. TRADOC, Ft. Monroe, Va.

TRADOC (Training and Development Command). 1982. A Systems Approach to Training. Regulation 350-7. TRADOC, Ft. Monroe, Va.

FIELD EXERCISES

Terry Michael Rauch

Military field exercises generally provide large-scale testing and training of soldiers and Army systems by simulating combat environments. The mission objectives of such exercises are based on conventional and unconventional warfare scenarios likely to be employed by U.S. forces. Operational and environmental components of field exercises place realistic demands on the soldier-materiel system interface that are unparalleled by laboratory or field experiments for realism and thus provide a unique test of military performance. The primacy accorded field data in social-behavioral research is most significant and based on the unvarying testimony of proponents with diverse methodological views. Although there is little unanimity about the analysis and interpretation of field data, there is a consensus about the need to examine and understand the assumptions of field data and their implications in determining the performance of the individual soldier, the military unit, and the interfacing materiel systems. I will discuss methodological perspectives of measuring soldier performance, focusing on the strengths and weaknesses of research designs.

Methods of research on soldier performance may be organized into three major technique categories: laboratory experiments, field-chamber experiments, and field studies. Classification depends on distinctions between experimental versus nonexperimental approaches and between laboratory versus field settings. The characteristics and criteria of laboratory experiments, field-chamber experiments, and field studies (field exercises are synonymous with field studies) designed to measure military performance need to be compared for their relative merits. Operationally defined, military performance is viewed as a function of the individual soldier or unit interfacing with a dependent materiel system; hence, individual soldier or unit performance is perceived as a functional entity deserving independent measurement and analysis. Emphasis is on the performance parameters of the individual soldier or unit. However, inferences may be applied to the soldier-machine-environment interface, although such designs become complex and it becomes difficult to generalize from them.

Laboratory experiments of soldier performance are characterized by criteria in which the variance of all (or nearly all) the influential independent variables not relevant to a dependent measure(s) is purposefully minimized. This type of variance is reduced by isolating research in a physical environment away from routine processes and by systematically manipulating one or more independent variables under rigorously controlled conditions. Hence, laboratory experiment designs have the inherent virtue of relatively complete control of extraneous factors, reducing error variance on dependent measures. Laboratory designs use random assignment and manipulation of one or more independent variables to minimize error variance. Moreover, precise, accurate, unambiguous experimental procedures result in less error variance. With exact specification of materials and methods, the risk that subjects may respond equivocally and introduce random variance is lessened. However, the most significant weakness of laboratory tests of soldier and unit performance is that independent variables have little effect on dependent variables, an attribute of the artificiality of such designs. Yet when experimental designs are developed to exclude (control) environmental noise, implying systematic manipulation of one or few independent variables, it becomes very difficult to relate the results to the real world. Laboratory designs incorporate high internal validity at the expense of external validity. Characteristically, internal validity implies a causal X-Y relationship. For example, the relationship between X, a pharmacological agent, and Y, heart rate, is a simple paradigm. Let us assume that the relationship is linear, so that increases in dose X result in significant increases in heart rate Y. However, extraneous factors deliberately excluded or controlled for will affect heart rate in a similar way in the real environment. Hence, many extraneous variables such as fear, noise, vibration, dust, wind, or any stressor affecting autonomic outflows will confound the principal independent variable. Obviously, external validity is limited in laboratory designs of soldier or unit performance parameters. A significant value of laboratory research is determination of the relationships and mechanisms underlying X-Y to derive a functional form, Y = (f)X, so that empirical values can be seen to agree with or deviate from predicted values. The basic logic of research planning should use the relationships derived in the laboratory for validation in field situations.

Field experiments include chamber, obstacle course, classroom, and, to a certain extent, open-field settings. Their primary purpose is to simulate a limited military operation in which one or more independent variables are systematically manipulated and extraneous noise is controlled as much as the physical resources permit. Whereas laboratory experimentation has maximum control, field experiments operate with less control as a trade-off for realism. Chamber studies permit the manipulation of a wide spectrum of environmental variables and simulate the operational military environment to a limited degree, the primary constraint being physical resource parameters. For example, in an environmental chamber subjects may be evaluated on several cardiopulmonary parameters as a function of workload, induced by a treadmill, at a specified temperature, humidity, amount of light, and windspeed, while wearing various clothing ensembles. Because independent variables can be varied and randomization used, the criterion of control is satisfied within limits. Control of the experimental field is rarely as tight as in the laboratory; nevertheless, intrinsic properties of the physical experimental situation generally give independent variables a stronger effect, so that the more realistic the research environment, the stronger the variables. Realism not only increases the strength of variables but also contributes to external validity. Realistic field experiments offer more valid generalizations to the real military operational environment.

Designs developed in field experiments are appropriate for investigating physiological parameters, but are also essential to understanding complex social influences, processes, and dynamics in military unit cohesiveness and psychological performance. The principal weakness of field experiments is the awkwardness of clearly quantifying and manipulating independent variables in addition to the use of randomization. Although there is no theoretical reason why randomization cannot be used in many open-field designs, difficulties such as unwillingness to break up units or allow soldiers in small elite or special units to be assigned to experimental and control groups at random occur frequently. Random assignment may be employed in such designs, but the independent variable may still be confounded because the effect of treatments cannot be isolated from other effects.

Another weakness inherent in designs of this nature is lack of precision, resulting in systematic and random noise. Hence, to measure the effect of X-Y in a field experiment, it becomes necessary to maximize variance attributed not only to the manipulated or assigned variables, but also to quantify and measure the dependent variable precisely. Yet, all things considered, in the military field situation extraneous independent variables are numerous and measures of dependent variables may not be sensitive enough to detect variance resulting from the effects of independent variables.

Field studies are *ex post facto* efforts characterized by an inability to gain direct control of independent variables because their effects have already been manifested in some variance format or because they are inherently not manipulable. Inferences about relations between variables are discerned, without direct intervention, from concomitant variation of independent and dependent variables. The experimenter works on the premise that X is quantified and varied to observe Y and determine whether concomitant variation expected or predicted from manipulation of X occurs. Hence, to achieve control, randomization and manipulation of X can lead to the assumption that $Y = (f)X$. In *ex post facto* research, Y is observed and X, or several X's, is also observed, either before, after, or concomitant with observations of Y. The logical validity of the structure of experimental and *ex post facto* designs is the same: to establish the empirical validity of a fundamental conditional statement, i.e.,

$\underline{Y} = (f)\underline{X}$. However, the basic difference is that of direct control by manipulation of $\underline{X}$ and secondarily the use of randomization in experimental as opposed to ex post facto designs. Field studies incorporate a real military operational situation or exercise to study relationships among performance factors, both psychological and physiological, of individual soldiers and their units in a true sense of realism. Normally, independent variables are not manipulated. Field designs of this type have traditionally focused on significant factors in the field, relationships among factors, and establishing a framework for hypothesis testing.

Research efforts with military field exercises are strong in realism, significance, strength of variables, and heurism. Variables used in field designs exhibit a large degree of variance compared with laboratory designs. Consider the relative strength of cohesiveness as a dependent variable in the laboratory where subjects are asked about their desire to remain in a group dependent on some reward versus a similar measure of unit cohesion for soldiers deployed in a simulated combat field exercise operating toward a major objective. Although the strength of field variables is great, field designs are plagued by the noise of extraneous variables, adding to the difficulty of separating confounding variables from independent variables. The most serious weakness of field studies is their ex post facto character. Hence, statements of relationships are weaker than in experimental studies, in addition to the inherent plethora of confounding variables and variance.

Indeed, measurement methods for soldier performance, aside from field validations of fundamental relationships derived in the laboratory, involve a trade-off between how much internal validity is sacrificed for external validity and what degree of control the investigator is willing to exercise as a function of the relative merits of laboratory, chamber, and field designs.

COGNITIVE TESTING IN MILITARY PERFORMANCE RESEARCH

Jared B. Jobe and Louis E. Banderet

Cognitive performance testing in the military has been a challenge for military psychologists for many years. Psychologists in many different laboratories have assessed cognitive performance with many different test batteries; there seems to be little agreement on the best approach.

This paper will review some varied approaches to the assessment of cognitive performance in the military under stressful conditions. A comprehensive review of all cognitive performance batteries is beyond the scope of this paper; instead, a few representative approaches are covered.

MULTIPLE-TASK PERFORMANCE BATTERY

The multiple-task performance battery (MTPB) was developed primarily under Air Force contract to assess the performance of aerospace crews operating under different work-rest schedules and periods of partial and total sleep deprivation (cf. Chiles et al., 1968). Criteria were that the task must measure functions performed by the operator in advanced systems; the task must measure, from the subject's viewpoint, something valid and important; the test-retest and mechanical reliabilities of the tasks must be satisfactory; the task measures must be orthogonal; the task must be sensitive to various manipulations; and asymptotic performance must be achievable within reasonable training time. Five tasks were selected by these criteria. (1) Monitoring static processes consisting of warning lights and auditory vigilance: the operator is required to scan dials, lights, or displays while looking or listening for a discrete binary change in the system. (2) Monitoring dynamic processes consisting of monitoring dials or displays that continually change and determining whether a prescribed range was maintained. (3) Stimulus discrimination or target identification: the operator makes simultaneous or successive comparisons of stimuli in order to render some judgment according to one or more parameters of the stimuli. Targets consisted of a 6 x 6 matrix of lights used to present contoured figures. (4) Arithmetic computation: the operator adds two three-digit numbers and then subtracts a third three-digit number from the sum.

The problems are presented at a rate of three per minute. (5) Procedural performance: group problem-solving tasks that require discovering the sequence in which five buttons must be pushed to solve a code-lock problem (Alluisi, 1967, 1969; Alluisi et al., 1964; Chiles et al., 1968; Morgan and Alluisi, 1972).

An important feature of the MTPB was that the tasks were performed on a time-sharing basis. Tasks were presented for 2-h blocks consisting of monitoring alone or a combination of two or three tasks for 15- to 30-min periods. In all studies subjects were confined to the experimental crew compartments for the duration of the study.

Regarding practice effects, the tasks either took extensive practice to reach asymptote or were at high levels initially. Asymptotic levels took longer to attain if time-sharing demands were high. For the percentage of correct arithmetic problems index, performance was asymptotic from the beginning. In terms of speed, 15 h of practice was estimated as the time required to reach asymptote for the arithmetic problems. For pattern discrimination, improvement occurred across 45 h of practice (Adams and Chiles, 1961), and the test was replaced with a target identification task that required 9 to 12 h of practice to achieve asymptote (Alluisi et al., 1963). Performance on the code-lock task reached asymptote after only 3 h of practice (Chiles et al., 1968). Probability-monitoring performance was initially high for both percent correct and detection time. Code-lock solving did not achieve asymptote for a considerable period of time, especially when time-sharing demands were high. Performance on the auditory vigilance task showed almost no learning effect, starting and remaining at 80 to 90% correct detections.

Reliability coefficients for the test battery were generally quite high (0.674 to 0.968), except for target identification (0.275 to 0.283). Intercorrelations among tasks were generally low (0.003 to 0.534), except for auditory vigilance, which correlated with arithmetic performance. The reason for this was hypothesized to be a factor common to both tasks, i.e., alertness or motivation.

Performance on the MTPB has generally been shown to be sensitive to a variety of conditions. Thirteen studies investigating circadian rhythms, sleep loss, and work-rest schedules were summarized by Chiles et al. (1968). Significant circadian variation was found on probability monitoring, arithmetic performance, and warning-light monitoring. Comparing 6-h-on, 2-h-off (6/2) and 4/2 work-rest schedules for 4 days yielded inconclusive results. Performance on a 4/4 schedule for 30 days was superior to that on a 4/2 schedule for all but the arithmetic task. In a subsequent study, a 4/4/4/12 schedule resulted in performance equal to that for the 4/4 schedule of the earlier study on all tasks. Further studies were conducted in which periods of sleep deprivation were superimposed on the work-rest schedules. Results showed that performance on the 4/2 schedule was superior to that on the 4/4 schedule, except for the target identification task, both before and during sleep deprivation.

The MTPB was used in a study assessing the effects of 36 h of sleep deprivation superimposed on a 4/4/4/12 schedule (Coates et al., 1974). Performance decrements ranged from 11 to 33% depending on the time of day the sleep loss period began (data from individual tasks were not presented). Similar data were presented by Morgan and Coates (1974) and Morgan et al. (1974); sleep loss and circadian effects were found after 36 to 48 h of sleep loss.

The effects of infectious diseases on MTPB performance were also assessed (Alluisi, 1969; Alluisi et al., 1971; Thurmond et al., 1968, 1971). In one experiment (Alluisi et al., 1971), eight subjects infected with *Pasteurella tularensis* (rabbit fever) and two uninfected controls worked a 4/4/4/12 schedule for 12 successive days, with exposure to the disease occurring on day 5. Average efficiency fell about 25% during the illness and was still 10 to 15% below that of the controls after recovery. Performance on active tasks was worse than on watch-keeping tasks, but recovery was less nearly complete on the watch-keeping tasks. The study was replicated during an extended 15-day study, and performance and recovery on the watch-keeping tasks were better than on the active tasks (Thurmond et al., 1968). The effects of a similar disease, sandfly fever, have also been studied (Alluisi et al., 1973).

PERFORMANCE EVALUATION TESTS FOR ENVIRONMENTAL RESEARCH

Another test battery, Performance Evaluation Tests for Environmental Research (PETER), was developed in the late 1970s at the Naval Biodynamics Laboratory in New Orleans, La. This ongoing program has been guided by theoretical and statistical issues associated with repeated performance assessment and represents approximately 7 years of development. A 90-citation bibliography (Harbeson et al., 1983) summarized the program's efforts.

This program has made many contributions. Initially, a bibliography of tests from environmental, information processing, and neuropsychological batteries that were affected by various environmental stressors was compiled (Dixon et al., 1978). Next, the PETER program studies varied test media and tests (Bittner et al., 1983, 1984a, 1984b; Carter and Sbisa, 1982; Jones et al., 1981; Kennedy et al., 1981a, 1981b, 1982; Smith et al., 1983). Paper-and-pencil tests, video games, computer-administered tests, and tests as diverse as contrast sensitivity, the Stroop test, reaction time, spoke test, time estimation, Baddeley's grammatical reasoning (Baddeley, 1968), and pattern comparison were evaluated. Methods and statistical procedures were specified or identified for description and quantification of test stability and reliability (Bittner and Carter, 1981; Krause and Woldstad, 1983). These methods provide information on practice requirements, number of subjects required, testing duration, and the relative stability and reliability of various tests and test media. Papers describing the advantages of repeated-measures designs and stable tests have also been written (Bittner and Carter, 1981). A transition to automated testing was evident when electronic and video games were demonstrated as a viable off-the-shelf test medium (Jones et al., 1981; Kennedy et al., 1981a, 1981b, 1982). Shortly thereafter, a publication listing FORTRAN programs showed explicitly how eight paper-and-pencil tests and their alternative forms could be created by computer (Carter and Sbisa, 1982). Such programs could also generate tests for computer administration.

Many different tests and measurement indexes (e.g., number correct, proportion correct, percent change) have also been evaluated (Bittner et al., 1984a). Each test or measurement index was examined for the time required to achieve total stability of means, variances, and intertrial correlations; the time required for differential stability of intertrial correlations; and the estimated reliability for a 3-min administration. Each of the 112 measures in the report was given one of four ratings; only 30 were recommended, 15 were acceptable but redundant, 35 were marginal, and 32 were unacceptable. The 30 recommended measures and other descriptive information are summarized in Table 1. It is interesting that a test's statistical properties are influenced by the measurement index chosen. For example, the manikin test, using log latency measures, is a recommended test, whereas if accuracy measures are used, it is a marginal test. Also, Sternberg item recognition for either positive target set 1 or 4 is a recommended test. When slope or intercept measures are used, the Sternberg test is rated unacceptable. It is also noteworthy that the practice times to achieve total test stability (1 to 135 min) are much shorter than those required with the MTPB described earlier. It is likely that the time-sharing characteristics and the greater complexity of some of the MTPB tasks produce the markedly different practice requirements.

Such procedures also make it possible to compare modified versions of a test and to compare paper-and-pencil tests with their automated equivalents (Bittner et al., 1984b; Smith et al., 1983). The last work is especially provocative, since it suggests that tests should be automated cautiously. Unfortunately, the statistical properties of automated tests are often inferior to those of their paper-and-pencil counterparts (Bittner et al., 1984a, 1984b; Krause, personal communication; Smith et al., 1983).

One of the most recent contributions from the PETER program was the implementation of an automated portable test system (Bittner et al., 1984b). The hardware selected was a notebook-sized computer, NEC PC8201A, which can be operated from batteries and is very portable. To date, three questionnaires and 15 performance tests have been computerized, and the adaptation of 30 or more tests is anticipated. The test system is being used in a wide range of studies that include vibration, simulator aftereffects, flight tests, hypoxia, and drug

TABLE 1. Thirty recommended tests from the PETER program[a]

Test	Factor[b]	Domain[c]	Administration Time (min)	Administration Type[d]	Time in minutes for total stability (differential stability)	3-min reliability[e]
Aiming	Aiming: fine eye-hand coordination	P, M	2.0	G	30 (30)	0.87
Arithmetic: vertical addition	Number facility	C	4.0	G	48 (8)	0.90
Associative memory: number correct list 1	MA	C	2.5	G	20 (20)	0.65
Atari™ Air Combat Maneuvering	Pursuit tracking	P, M	2.25	I	135 (135)	0.63
Atari™ Antiaircraft	Unknown	P, M	2.25	I	126 (126)	0.67
Choice reaction time:						
One choice	Simple reaction time	P	5.0	I	35 (35)	0.58
Four choice	Choice reaction time	P	5.0	I	50 (50)	0.80
Code substitution	MA, PS	C, P	2.0	G	16 (16)	0.84
Flexibility of closure	CF	P	3.0	G	9 (9)	0.88
Grammatical reasoning	RL	C	1.5	G	18 (18)	0.93
Graphic and phonemic analysis: sense/nonsense	Reading speed	C	8.0	G	16 (16)	0.66
Letter classification: name	Retrieval from LTM and matching	C	12.0	G	84 (84)	0.55
Letter classification: category	Retrieval from LTM and matching	C	11.0	G	121 (121)	0.69
Manikin test: log latency	Spatial transformation	P	7.0	I	14 (14)	0.79
Minnesota rate of manipulation: turning	Manual dexterity	M	2.0–4.0	I	10 (10)	0.64
Pattern comparison: no. correct minus no. incorrect	Spatial ability	P	2.0	G	18 (18)	0.93
Perceptual speed	PS	P	2.5	G	23 (15)	0.86

(TABLE 1 continues)

TABLE 1. Con't.

Test	Factor[b]	Domain[c]	Administration Time (min)	Administration Type[d]	Time in minutes for total stability (differential stability)	3-min reliability[e]
Search for typos in prose: median detection time	Reading speed	P	6.0	I	54 (54)	0.65
Spoke control task	Speed, arm movement	M	0.67	G	1 (1)	0.95
Sternberg item recognition:						
Positive set 1	STM scan	C	3.0	I	18 (18)	0.70
Positive set 4	STM scan	C	3.0	I	15 (9)	0.80
Stroop color words	Mixed	C, P	0.5	G	1.5 (1.5)	0.97
Tracking:						
Critical	Tracking, critical	P, M	1.0	I	100 (100)	0.60
Dual critical	Tracking, critical & dual factor	P, M	1.0	I	100 (100)	0.50
Visual contrast sensitivity: method of increasing contrast	Contrast sensitivity function:					
	1 cpd	P	3.0	I	<1 (<1)	0.51
	2 cpd	P	3.0	I	<1 (<1)	0.52
	4 cpd	P	3.0	I	<1 (<1)	0.74
	8 cpd	P	3.0	I	<1 (<1)	0.75
	16 cpd	P	3.0	I	<1 (<1)	0.53
Word fluency	Word fluency	C	3.0	G	<1 (<1)	0.79

[a]From Bittner *et al.*, 1984a.
[b]Abbreviations: MA, memory, associative; PS, perceptual speed; CF, flexibility of closure; RL, logical reasoning; LTM and STM, long- and short-term memory; cpd, cycles per degree.
[c]Abbreviations: P, perceptual; M, motor; C, cognitive.
[d]Abbreviations: G, group; I, individual.
[e]Estimated 3-min administration.

effects. Over 1,000 subjects have been studied with the system in environments such as vibration platforms, ships, and hypobaric and hyperbaric chambers. Suggested procedures and formulas to evaluate automated tests have also been described (Bittner et al., 1984b).

Selected methods from the PETER program have been applied at the U.S. Army Research Institute of Environmental Medicine (USARIEM) in Natick, Mass. Six of the eight tests described (Carter and Sbisa, 1982) were computer generated and printed on an off-line laser copier. Most tests were adapted and made more difficult, e.g., pattern comparison discriminations were made less obvious and two new verbs were added to Baddeley's (1968) grammatical reasoning test. In addition, three new paper-and-pencil tests with alternative forms were developed at USARIEM. The Tower of Hanoi, a popular puzzle and test of logical reasoning (Anzai and Simon, 1979), was adapted as a paper-and-pencil test with discrete, forced-choice responses (Banderet et al., 1984). The computer interaction test (Banderet et al., 1984) requires actions similar to those of military personnel who use computer keyboard and display systems. It evaluates a person's global transactions with a computer system, a 12-digit desk-top calculator. The map-compass orientation test is a paper-and-pencil test with word problems to test application of map, compass, and related spatial concepts.

Different environmental conditions have been studied with tests adapted from PETER or new tests developed at USARIEM. In a double-blind study (Banderet and Jobe, 1984), 0.5- to 2.0-mg doses of atropine were tested. Pattern comparison errors increased the first time atropine (0.5 mg) was given. Decreased performance rates on the coding and grammatical reasoning tests approached statistical significance during replication of the 2-mg dose. It is likely that the dose regimen and the small number of subjects (7) caused an underestimate of the actual drug effects.

In a study of hypoxia and acute mountain sickness, seven tests were administered to 23 subjects at a simulated high (4,600 m) altitude (Banderet and Burse, 1984). Performance on all tests was most impaired on the first administration at high altitude (1 or 6 h) and was 72 to 88% of baseline. After 38 and 43 h at high altitude, performance had recovered fully on all but two of the tests.

In a study of dehydration (room temperature) and cold exposure with restricted fluid intake (Banderet et al., in press), five tests were given to 18 subjects dehydrated 2.5% by body weight at room temperature. The dehydrated subjects' performance on all tests except grammatical reasoning was significantly lower than that of the control subjects. Performance on all tests except grammatical reasoning was impaired when the control subjects wore the protective Arctic uniform in -25°C (4 km/h winds) conditions. Also, dehydration produced impairments comparable to those observed during the cold exposure.

Results from these environmental studies indicate that the tests adapted from the PETER program are sensitive. In addition, these results show that methods from the PETER program can aid in the development and evaluation of new tests (Banderet et al., 1984).

WALTER REED PERFORMANCE ASSESSMENT BATTERY

The Walter Reed Performance Assessment Battery (PAB) is a primarily cognitive test battery developed for computer presentation and control and for sensitivity to ultradian and circadian variations. The PAB uses a 48K Apple II or Apple II+ microcomputer with a single disk drive and requires a California Computer System timer card as well as minor video circuit modifications. Programs are written in Applesoft BASIC under DOS 3.3 (Thorne et al., 1985). The microcomputer system allows ease of presentation, subject feedback, scoring, and statistics. Currently, the battery consists of the following 11 tasks (Thorne et al., 1985). (1) Two-letter search, a visual search and recognition task. Two target letters are presented at the top of the screen, followed by a string of 20 letters in the middle of the screen. The subject must determine whether both target letters are in the string. (2) Six-letter search consists of a six-letter target, otherwise it is identical to the two-letter search. (3) Encoding-decoding consists of translating a series of letters into numerical map coordinates, or vice versa, using a double set of code keys. (4) Two-column addition consists of adding five two-digit numbers which are presented simultaneously in column format. The column disappears with the first entry, and no aids for the carried-over numbers are allowed. This addition task requires a different strategy than the addition task from the PETER battery and is more difficult. (5) Logical reasoning, similar to the PETER task, consists of the presentation of the letter pair AB or BA along with a statement that correctly or incorrectly describes the order of the letters in the pair. (6) Digit recall, a short-term memory task, consists of the presentation of a row of nine digits for 1 s. Following a retention interval of 3 s, eight of the nine digits are redisplayed in a different random order, and the subject enters the missing digit (no digit appears more than twice). (7) Serial addition or subtraction consists of the 50-ms presentation of two random digits and either a plus or a minus sign. The subject performs the operation and enters the least significant digit of the result, adding 10 if the result is negative. (8) Pattern recognition I, a spatial memory task, consists of the presentation for 1.5 s of a random pattern of asterisks in a 14 x 14 matrix, in which each column and row contains a single dot and 13 blanks. After a 3.5-s retention interval, a second pattern is presented, either identical or with three dots in different positions. The subject determines whether the two patterns are identical. (9) Pattern recognition II is a more difficult version, with a 16 x 16 matrix in which either two or no dots change position. The pattern is presented successively, adding a memory component, whereas in the PETER the two patterns are presented simultaneously. (10) Mood-activation scale consists of the presentation of 65 adjectives representing positive or negative effect or activation. (11) Four-choice reaction time uses a box containing four light-emitting diodes in a square array mounted above four corresponding pushbuttons. Single lights are illuminated randomly, and the subject presses the corresponding button as quickly as possible.

Like the two performance batteries reviewed previously, the PAB is designed for use in repeated-measures experiments with subjects being trained to asymptotic levels. In several versions, the battery has been used to study the effects of sleep deprivation (Thorne et al., 1983), sustained performance, jet lag, heat stress, physical fatigue, hypoxia, and sickle cell disorders (cf. Thorne et al., 1985). Eight PAB tasks (except encoding-decoding, mood, and reaction time) were used in a study assessing the effects of a 72-h sleep deprivation (Thorne et al., 1983). The tasks were performed at asymptotic levels before the 72-h period for the two female and six male subjects. Each subject spent a minimum of 30 min each hour on testing, with the PAB given every other hour. Other tests were also given, and subjects were allowed free time. All eight tasks showed similar declines after sleep deprivation and showed circadian effects of approximately 10% of baseline as well. Substantial recovery was demonstrated on all eight tasks after a 4-h nap. Performance was 26% of baseline at 72 h and 77% of baseline after the 4-h nap.

SUMMARY

Three different test batteries, primarily cognitive, were reviewed. The batteries were chosen to represent three different services, Air Force, Navy, and Army, and three different approaches. The MTPB was designed to model real-world, complex operator tasks; the PETER was developed for methodological and statistical properties to be used in a variety of situations; and the PAB was developed for testing ultradian and circadian variations and for ease of presentation and scoring via computer. These batteries, designed to assess different military performance and skills, have all been successful in being sensitive to a wide range of experimental manipulations, including sleep deprivation, environmental extremes, infectious disease, drugs, and restricted fluid intake.

REFERENCES

Adams, O.S., and W.D. Chiles. 1961. Human performance as a function of the work-rest ratio during prolonged confinement. U.S. Air Force WADC Technical Report 61-720. Aerospace Medical Laboratory, November 1961. Wright-Patterson Air Force Base, Ohio.

Alluisi, E.A. 1967. Methodology in the use of synthetic tasks to assess complex performance. Hum. Factors 9:375-384.

Alluisi, E.A. 1969. Sustained performance. In E.A. Bilodeau and I. M. Bilodeau, eds. Principles of Skill Acquisition. Academic Press, New York.

Alluisi, E.A., W.D. Chiles, T.J. Hall, and G.R. Hawkes. 1963. Human group performance during confinement. Aerospace Medical Research Laboratory Report AMRL-TDR-63-87 (NTIS No. AD-426-661). Wright-Patterson Air Force Base, Ohio.

Alluisi, E.A., W.D. Chiles, and T.J. Hall. 1964. Combined effects of sleep loss and demanding work-rest schedules on crew performance. Aerospace Medical Research Laboratory Report AMRL-TDR-64-63 (NTIS No. AD-606 214). Wright-Patterson Air Force Base, Ohio.

Alluisi, E.A., J.B. Thurmond, and G.D. Coates. 1971. Behavioral effects of infectious disease: respiratory _Pasteurella_ _tularensis_ in man. Percept. Motor Skills 32:647-688 (Monogr. Suppl. 1).

Alluisi, E.A., W.R. Beisel, P.J. Bartelloni, and G.D. Coates. 1973. Behavioral effects of tularemia and sandfly fever in man. J. Infect. Dis. 128:710-717.

Anzai, Y., and H. Simon. 1979. The theory of learning by doing. Psychol. Rev. 86:124-140.

Baddeley, A.D. 1968. A three minute reasoning test based on grammatical transformation. Psychonom. Sci. 10:341-342.

Banderet, L.E., and R.L. Burse. 1984. Cognitive performance at 4500 meters simulated altitude. Presented at the Meeting of the American Psychological Association, Toronto, Canada, 24-28 August 1984. [Unpublished.]

Banderet, L.E., and J.B. Jobe. 1984. Effects of atropine upon cognitive performance and subjective variables. Technical Report No. T5/85, November 1984. U.S. Army Research Institute of Environmental Medicine, Natick, Mass. 21 pp.

Banderet, L.E., K.P. Benson, D.M. MacDougall, R.S. Kennedy, and M. Smith. 1984. Development of cognitive tests for repeated performance assessment. Pp. 375-380 in Proceedings of the 26th Annual Meeting, Military Testing Association, Munich, Federal Republic of Germany, 5-8 November 1984, Vol. 1.

Banderet, L.E., D.M. MacDougall, D.E. Roberts, D. Tappan, M. Jacey, and P. Gray. In press. Effects of dehydration or cold exposure and restricted fluid intake on cognitive performance. In Proceedings of the Workshop on Knowledge Needed for the Development of Predictive Models of Military Performance Decrements Resulting from Inadequate Nutrition, Washington, D.C., 22-24 October 1984. National Academy Press, Washington, D.C.

Bittner, A.C., Jr., and R.C. Carter. 1981. Repeated measures of human performance: a bag of research tools. Report No. NBDL-81R011, November 1981. Naval Biodynamics Laboratory, New Orleans, La. 24 pp.

Bittner, A.C., Jr., R.C. Carter, M. Krause, R.S. Kennedy, and M.M. Harbeson. 1983. Performance evaluation tests for environmental research: Moran and computer batteries. Aviat. Space Environ. Med. 54:923-928.

Bittner, A.C., Jr., R.C. Carter, R.S. Kennedy, M.M. Harbeson, and M. Krause. 1984a. Performance evaluation tests for environmental research: evaluation of 112 measures. Report No. NBDL-84R006, September 1984. Naval Biodynamics Laboratory, New Orleans, La. 38 pp.

Bittner, A.C., Jr., M.G. Smith, R.S. Kennedy, C.F. Staley, and M.M. Harbeson. 1984b. Automated portable test (APT) system: overview and prospects. Report No. 70189-0407, November 1984. Essex Corporation, Orlando, Fla. 16 pp. [Also published as "Instruments and Computers", Behav. Res. Methods 17:217-221, 1985.]

Carter, R.C., and H. Sbisa. 1982. Human performance tests for repeated measurements: alternate forms of eight tests by computer. Report No. NBDL-82R003, January 1982 (NTIS AD A11502). Naval Biodynamics Laboratory, New Orleans, La. 55 pp.

Chiles, W.D., E.A. Alluisi, and O.S. Adams. 1968. Work schedules and performance during confinement. Hum. Factors 10:143-196.

Coates, G.D., B.R. Brown, and B.B. Morgan, Jr. 1974. Interaction of continuous work and sleep loss with the effects of circadian rhythm--II. In B.B. Morgan, Jr., and G.D. Coates, eds. Sustained Performance and Recovery During Continuous Operations. ARI Technical Report No. 74-2. U.S. Army Research Institute, Arlington, Va.

Dixon, D.J., M.G. Copeland, and C.G. Halcomb. 1978. Psychomotor battery approaches to performance prediction and evaluation in hyberbaric, thermal, and vibratory environments: annotated bibliographies and integrative review. Report No. NAML-N00203-78-M5020, May 1978. Naval Aerospace Medical Research Laboratory, New Orleans, La. 212 pp.

Harbeson, M.M., A.C. Bittner, Jr., R.S. Kennedy, R.C. Carter, and M. Krause. 1983. Performance evaluation tests for environmental research (PETER): bibliography. Percept. Motor Skills 57:283-293.

Jones, M.B., R.S. Kennedy, and A.C. Bittner, Jr. 1981. A video game for performance testing. Am. J. Psychol. 94:143-152.

Kennedy, R.S., D.A. Andrews, and R.C. Carter. 1981a. Performance evaluation tests for environmental research (PETER): a microcomputer game as a memory test. Paper presented at the Meeting of the Aerospace Medical Association, San Antonio, Texas, May 1981.

Kennedy, R.S., A.C. Bittner, Jr., M.M. Harbeson, and M.B. Jones. 1982. Television-computer games: a "new look" in performance testing. Aviat. Space Environ. Med. 53:49-53.

Kennedy, R.S., A.C. Bittner, Jr., and M.B. Jones. 1981b. Video-game and conventional tracking. Percept. Motor Skills 53:310.

Krause, M., and J.C. Woldstad. 1983. Massed practice: does it change the statistical properties of performance tests? Report No. NBDL-83R005, April 1983. Naval Biodynamics Laboratory, New Orleans, La. 29 pp.

Morgan, B.B., Jr., and E.A. Alluisi. 1972. Synthetic work: methodology for assessment of human performance. Percept. Motor Skills 35:835-845.

Morgan, B.B., Jr., B.R. Brown, and E.A. Alluisi. 1974. Effects on sustained performance of 48 hours of continuous work and sleep loss. Hum. Factors 16:406-414.

Morgan, B.B., Jr., and G.D. Coates. 1974. Synthetic-work methodology for assessing performance under continuous operations. In B.B. Morgan, Jr., and G.D. Coates, eds. Sustained Performance and Recovery During Continuous Operations. Army Research Institute Technical Report No. 74-2, December 1974. U.S. Army Research Institute, Arlington, Va.

Smith, M.G., M. Krause, R.S. Kennedy, A.C. Bittner, Jr., and M.M. Harbeson. 1983. Performance testing with microprocessors: mechanization is not implementation. Pp. 674-678 in Proceedings of the 27th Annual Meeting of the Human Factors Society, Santa Monica, Calif, 1985, Vol. 2. Human Factors Society, Santa Monica, Calif.

Thorne, D., S. Genser, H. Sing, and F. Hegge. 1983. Plumbing human performance limits during 72 hours of high task load. Pp. 17-40 in Proceedings of the 24th DRG Seminar on the Human as a Limiting Element in Military Systems, Toronto, Canada, 1983. Defense and Civil Institute of Environmental Medicine, Toronto.

Thorne, D.R., S.G. Genser, H.C. Sing, and F.W. Hegge. 1985. The Walter Reed performance assessment battery. Neurobehav. Toxicol. Teratol. 7:415-418.

Thurmond, J.B., E.A. Alluisi, and G.D. Coates. 1968. An extended study of the behavioral effects of Pasteurella tularensis in man. Report ITR 68-8. University of Louisville Performance Research Laboratory, Louisville, Ky.

Thurmond, J.B., E.A. Alluisi, and G.D. Coates. 1971. An extended study of the behavioral effects of respiratory Pasteurella tularensis in man. Percept. Motor Skills 33:439-454.

WORKING GROUP DISCUSSION SUMMARIES

The Workshop on Cognitive Testing Methodology in Military Nutrition Research was designed to accomplish three aims: first, to review the state of the art in testing cognitive functions; second, to develop guidelines for selecting tests that might be used to evaluate the cognitive performance of military personnel and to make recommendations about the utility of specific tests; and finally, to identify areas in which research projects might improve our ability to evaluate the effects of nutrition on cognitive performance in military personnel. The papers presented in the first three sessions of the workshop and included in this volume provide an excellent background for discussion of these three issues. The final session of the workshop was devoted to discussions of the three aims of the workshop in light of this background material. Speakers and other participants in the workshop were formed into three groups, each with its own discussion leader. Individuals were assigned so that each group would include a mixture of investigators from basic and applied areas of cognitive testing, investigators from military laboratories, and members of the National Research Council Committee on Military Nutrition Research. Each discussion group was to give special attention to one of the three aims of the workshop, although, as the results of the working group discussions indicate, the three aims are so intricately connected that one cannot be addressed in isolation from the others.

Below are summaries of the discussions for each of the working groups. Following these are the general conclusions and recommendations developed in these working groups regarding the three aims of the workshop.

GROUP 1: DISCUSSION SUMMARY

This group was to give special attention to the first aim, the current state of the art in testing cognitive functions.

Discussion in the working group identified four major areas that need to be brought together for military nutrition research. The four areas are:

currently existing military tests and batteries;
military tasks that should be measured by the above-mentioned tests;
important nutritional variables that need to be investigated as independent variables; and
important modulating variables that should be either coindependent with nutrition variables or controlled variables, or have known values to minimize confounding.

It was agreed that the state of the art is such that performance relevant to important military tasks can be tested by using cognitive tests together with simulators and group-assessment methods. Following is a list of existing cognitive tests used by the military and a list of military tasks. These are reviewed in Session II.

Military tests: simulators, field studies, field exercises, Walter Reed performance assessment battery, Performance Evaluation Tests for Environmental Research (PETER) battery, Department of Defense Joint Working Group on Drugs Battery, Alluisi Battery, FDC Evaluation Methods, embedded performance, Strupp task.

Military tasks and behavior: flexibility, leader decision making, driven behaviors, self-paced behaviors, sensory-motor tasks, small-group performance, self-evaluation of performance, risk taking, visual scanning, affect.

Nutritional variables: caloric deprivation, water sufficiency, protein-carbohydrate ratio, electrolyte balance, vitamin sufficiency, performance modulators.

Modulating variables: mission demands, age, gender, race, heat, cold, altitude, noise, activation, repetition, sleep deprivation, circadian disruption, psychological stress, pollutants, drugs/alcohol, physical fitness, fatigue, nuclear, biological, chemical status.

It was also agreed that the ability exists to measure, manipulate, or control many of the modulating variables of importance. However, the sensitivity of existing tests to possibly subtle effects of nutritional variables is not known. The lists include nutritional variables that might affect cognitive performance and variables that might modulate or covary with the nutritional variables.

Other conclusions of the working group are as follows:

No single test battery for assessing cognitive function exists, and such a battery is inconceivable. The development of a single all-purpose, cognitive battery is not a reasonable research objective.

A finite set of military tasks with heavy cognitive loads need to be identified and submitted to psychometric analysis to identify component skills.

A number of batteries for assessing military performance have been developed, and the factorial composition of these tasks needs to be evaluated in the context of the preceding item.

Independent determination of the sensitivity of the tasks in the batteries needs to be done. Pharmacological and environmental variables should be used as positive controls.

Systematic studies with animals and humans should now proceed to isolate appropriate parameters for future field testing.

GROUP II: DISCUSSION SUMMARY

This group discussed the second aim, to recommend specific tests that may be used to assess the possible cognitive effects of nutritional deficits on military performance. The discussion centered on three issues: whether there is any basis for believing that nutrition deficits can influence cognitive measures, the need for a variety of tests that measure cognitive performance that are applicable to military performance, and the recommendation of specific cognitive processes that should be examined.

Although it was argued by some that the demands and the structure that are imposed on military personnel may preclude any subtle influence of a nutritional deficit, the majority of the discussants agreed that the effects of nutritional deficits may be equivalent to those produced by fatigue, sleep deprivation, or jet lag. Moreover, it was pointed out that poor nutrition may interact with fatigue, sleep deprivation, and jet lag and exaggerate their deleterious effects on military performance.

The discussion then shifted to the task of identifying specific tests of cognitive function that the military may use to study the possible influence of nutritional deficits. It was concluded that this task was impossible to complete at present because of the paucity of information on the cognitive processes that are important for optimal military performance. The discussion emphasized that research to identify these cognitive processes is critical to assess the degree to which nutritional deficits may impair military performance.

Finally, the discussion centered on the identification of specific cognitive processes that were thought to be most vulnerable to nutritional influences. These processes include:

Attention. A battery of tests designed to measure attention should include tests of focused, selective, divided, and sustained attention. Again, it was emphasized that these tests of attention must be field-tested, and only those that measure the kinds of attentional processes required for optimal military performance should be used.

Memory. Many different kinds of tasks and tests of memory for different kinds of information relevant to the performance of military tasks should be included.

Self-Monitoring Behavior. These tasks should be directed towards evaluating the individuals' perception of their own ability to complete a particular military task.

Reading Ability. Tests of reading ability are important because they incorporate the assessment and integration of memory, reasoning, and verbal comprehension.

Tests of Affect and Morale. It was felt that food may play a very large, although indirect, role on military performance by affecting mood and morale.

Cognitive Strategies. Several discussants recommended that not only cognitive abilities should be assessed, but also cognitive strategies. It is possible to arrive at the same answer to a problem by using very different strategies. Knowing how nutritional deficits may affect the strategies that military personnel use may help predict the kinds of errors they would make in new situations.

Self-Initiated Cognitive Activities. Although it is currently difficult to measure, several discussants felt that this kind of cognitive measure may be particularly sensitive to the effects of poor nutrition. Its importance in military performance must be determined before it is included in any test battery.

Speed-Accuracy Trade-Off. Finally, it was felt that speed-accuracy trade-off functions are a cognitive dimension that may be sensitive to nutritional deficits and would be of considerable military importance.

GROUP III: DISCUSSION SUMMARY

This group was to give special attention to the third aim, identifying research areas that should be pursued to enhance our ability to test the cognitive effects of nutritional variables. The group began by considering how tests had been selected for the Meal Ready-to-Eat (MRE) study currently being conducted on the island of Hawaii. It was generally agreed that the testing procedures presented and discussed at the current workshop would be very helpful in planning future studies of the effects of nutritional variables on performance. The subsequent discussions of the group were concerned with five topics:

Are there enough cognitive testing procedures currently available to begin a program of research investigating the effects of nutritional variables on cognitive performance?

What should be the first step in a program of research investigating the cognitive effects of nutritional variables?

If a battery of cognitive tests were shown to be sensitive to the effects of nutritional variables, how would one then proceed to show that those tests were predictive of performance in real military situations?

What kinds of nutritional variables might be used in initial studies to identify sensitive cognitive tests?

What should be the role of preclinical animal studies in a program of research on the role of nutritional variables in the cognitive performance of military personnel?

Each of these topics was discussed at some length, and examples relevant to each topic were presented. It was recognized that the topics are not independent and that specific recommendations may be relevant to more than one topic. The general conclusions and recommendations of the group were as follows:

There are currently more than enough laboratory-based cognitive testing procedures available to begin a program of research on the effects of nutrition on cognitive performance.

Since most of the recently developed cognitive tests have not been used to evaluate the effects of nutritional variables, there is a critical need to determine the sensitivity of various tests to changes in nutritional status.

Tests thought most likely to be sensitive are those that involve sustained alertness and possibly storage of information in memory. Neurophysiologically, these kinds of tasks require integrated functioning of large areas of the neocortex and hippocampus.

Nutritional states that might have substantial effects on cognition and that might actually be encountered in military situations include acute caloric deprivation and submaximum water deprivation. These states might be used to determine which cognitive parameters are sensitive to nutritional status.

Once sensitive tests have been identified and the nutritional status most likely to impair performance on those tests has been determined, their predictive power would have to be measured. This would require testing the effects of identified nutritional variables in nonlaboratory situations, including simulators, field exercises, and measures of crew interaction.

Preclinical animal models would be useful for determining whether specific kinds of nutrient deficiencies might affect performance. These nutrient deficiencies could not be induced in humans.

The working group also agreed that several other issues need to be considered if a viable research program is to be developed, but had no specific conclusions or recommendations about how to address them. The issues were:

Nutritional research, if it is to be relevant to military needs, must be incorporated into existing programs.

Attention must be paid to the psychometric properties of the tests incorporated into any battery. Issues of reliability and sensitivity to change must be addressed.

For some purposes, normative test data obtained from relevant military populations may be useful.

AUTHORS

Louis E. Banderet, Ph.D.
Health and Performance Division
U.S. Army Research Institute of Environmental Medicine
Natick, MA 01760

Stanley F. Bolin, Ph.D.
Chief, Special Projects
Training and Simulation Technical Area
Training Research Laboratory
Army Research Institute (PERI-IIA)
5001 Eisenhowner Avenue
Alexandria, VA 22333-5600

Kenneth L. Davis, M.D.
Professor of Psychiatry and Pharmacology
Mount Sinai School of Medicine
Psychiatry Service (116A)
Veterans Administration Medical Center
Bronx, NY 10468

David A. Eckerman, Ph.D.
Psychology, Davie Hall 013A
University of North Carolina
Chapel Hill, NC 27514

Earl Hunt, Ph.D.
Psychology Department NI-25
University of Washington
Seattle, WA 98195

Jared B. Jobe, Ph.D.
Training Technology Agency
P.O. Box 313
Ft. Monroe
Hampton, VA 23651

Kent A. Kimball, Ph.D.
Box 77
U.S. Army Aeronautical Research Lab
Ft. Rucker, AL 36362

Youngjai Kim, Ph.D.
Long Island Jewish-Hillside Medical Center
7559 23rd Street
Glen Oaks, NY 11004

David Levitsky, Ph.D.
Division of Natural Sciences
118 Savage Hall
Cornell University
Ithaca, NY 14853

Terry M. Rauch, Ph.D.
Captain, U.S. Army
Director, Health and Performance Division
U.S. Army Research Institute of Environmental Medicine
Natick, MA 01760

Andrew P. Smith, Ph.D.
Medical Research Council Perceptual and Cognitive Performance Unit
Laboratory of Experimental Psychology
University of Sussex
Brighton, Sussex BN1 9Q6
United Kingdom

Larry R. Squire, Ph.D.
Veterans Administration Medical Center, V-116
San Diego, CA 92161

Barbara Strupp, Ph.D.
Division of Nutritional Sciences
Cornell University
Ithaca, NY 14850

Myron F. Weiner, M.D.
Professor of Psychiatry
University of Texas
South Western Medical School
Dallas, TX 75202

Christopher D. Wickens, Ph.D.
Department of Psychology
University of Illinois
Champaign, IL 61820

PARTICIPANTS

Richard C. Mohs, Ph.D., Chairman
Psychiatry Service (116A)
Veterans Administration Medical Center
Bronx, NY 10436

Elsworth R. Buskirk, Ph.D.
119 Noll Laboratory
The Pennsylvania State University
University Park, PA 16802

Carl E. Englund, Ph.D.
Research Psychologist
Naval Health Research Center
P.O. Box 85122
San Diego, CA 92138

Murray Glanzer, Ph.D.
Department of Psychology
New York University
6 Washington Place
New York, NY 10003

R. Curtis Graeber, Ph.D.
MAJ(P), MSC
Aeronautical Human Factors Research Office
MS:239-21
National Aeronautics and Space Administration
Ames Research Center
Moffett Field, CA 94035

Frederick W. Hegge, Ph.D.
Chairman, JWGD3 Military Performance
SGRD-VWI-C
Walter Reed Army Institute of Research
Washington, DC 20307

Ed Hirsch, Ph.D.
U.S. Army Natick Research and Development Center
STRNC-YBH
Natick, MA 01760

Harris R. Lieberman, Ph.D.
Department of Psychology E10-130d
Massachusetts Institute of Technology
Cambridge, MA 02139

Maria G. Milles, Ph.D.
Division of Nutrition, HFF-265
Food and Drug Administration
Washington, DC 20204

Richard D. Popper, Ph.D.
U.S. Army Natick Research and Development Command
Natick, MA 01760

Paul Rambaut, Ph.D.
Code EBM
National Aeronautics and Space Administration Headquarters
Washington, DC 20546

Roger Remington, Ph.D.
National Aeronautics and Space Administration
MS 239-3
Ames Research Center
Moffett Field, CA 94035

David Schnakenberg, Ph.D.
Lieutenant Colonel, U.S. Army
U.S. Army Research Institute of Environmental Medicine
Natick, MA 01760

Thomas J. Sobotpa, Ph.D.
Leader, Neurobehavioral Toxicology Team
Division of Toxicology HFF-162
Center for Food and Safety and Applied Nutrition
Food and Drug Administration
Washington, DC 20204

Bonnie Spring, Ph.D.
Department of Psychology and Social Relations
Harvard University
Cambridge, MA 02138

John M. Talbot, Ph.D.
Life Sciences Research Offices
Federation of American Societies for Experimental Biology
9650 Rockville Pike
Bethesda, MD 20814

National Research Council Cognitive Testing Workshop Staff

Samuel B. McKee
Committee on Military Nutrition Research

John Wagner
Board of Army Science and Technology

Alexander Wigdor
Committee on Performance of Military Personnel