DOCUMENT RESUME

ED 283 857 TM 870 380

AUTHOR	Lesgold, Alan; And Others
TITLE	Toward Intelligent Systems for Testing. Technical Report LSP-1.
INSTITUTION	Pittsburgh Univ., Pa. Learning Research and Development Center.
SPONS AGENCY	Office of Naval Research, Arlington, Va. Personnel and Training Research Programs Office.
PUB DATE	Mar 87
CONTRACT	N00014-85-K-0655
NOTE	32p.
PUB TYPE	Reports - Evaluative/Feasibility (142)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Adaptive Testing; *Artificial Intelligence; *Cognitive Measurement; Computer Assisted Instruction; *Computer Assisted Testing; Course Objectives; Criterion Referenced Tests; *Diagnostic Tests; *Individualized Instruction; Instructional Development; Test Construction

ABSTRACT

This report illustrates one way in which the technologies of testing might combine with cognitive science techniques to help steer instruction. Steering testing is brief diagnostic testing that steers, or individualizes, the course of instruction. Steering testing uses simple heuristics for reasoning about the level of a student's competence in a particular subskill and intelligently manufactures practice opportunities that will be especially revealing about the student's current competences. Theoretically, steering testing should permit a partly logical constraining of diagnosis and should be based on a representation of the knowledge needed to exercise the skill it purports to measure. Four types of knowledge, involved in dealing with a student, need clarification when designing computer systems for steering testing: (1) domain expertise; (2) curriculum knowledge; (3) planning knowledge; and (4) treatment knowledge. In addition, a student model, a knowledge structure specifying which subskills a student is thought to know or to not know, is embedded in the curricular goal structure of the system. When a diagnosis is needed, the student model is examined to identify areas of competence about which more information is needed. These areas represent constraints on the type of test item that will be informative. Once the constraints are posted, an intelligent item generator constructs test items that satisfy them. To illustrate these ideas, an intelligent computer-based tutor, with a problem solving mode, that teaches basic electrical principles is discussed. (BS)

University of Pittsburgh

LEARNING RESEARCH AND DEVELOPMENT CENTER

Toward Intelligent Systems for Testing

Alan Lesgold
Jeffrey Bonar
Joyce Ivill

March 1987

Technical Report LSP-1

This research was sponsored by the Personnel and Training Research Programs, Psychological Sciences Division, Office of Naval Research, under Contract No. N00014-85-K-0655. Contract Authority Identification Number, NR 677-430.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) UPITT/LRDC/ONR/LSP-1	5. MONITORING ORGANIZATION REPORT NUMBER(S) NR 4422539
6a. NAME OF PERFORMING ORGANIZATION Learning Research and Development Center, Univ. of Pittsburgh 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Personnel & Training Research Programs Office of Naval Research (Code 1142PT)
6c. ADDRESS (City, State, and ZIP Code)	7b. ADDRESS (City, State, and ZIP Code) 800 North Quincy Street Arlington, VA 22217-5000
8a. NAME OF FUNDING/SPONSORING ORGANIZATION 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-85-K-0655/P00004
8c. ADDRESS (City, State, and ZIP Code)	10. SOURCE OF FUNDING NUMBERS

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT ACCESSION NO.
61153N	RR04206	RR04206-00	NR442c524

11. TITLE (Include Security Classification)

Toward Intelligent Systems for Testing

12. PERSONAL AUTHOR(S)
Lesgold, Alan M.; Bonar, Jeffrey G.; Ivill, Joyce

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Technical	FROM ____ TO ____	March 1987	32

16. SUPPLEMENTARY NOTATION A version of this report will appear as a chapter in "Knowing and Learning" Issues for the Cognitive Psychology of Instruction, ed. by Lauren Resnick (Hillsdale NJ: Lawrence Erlbaum Associates).

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Artificial Intelligence; Cognitive Psychology; Cognitive Science; Instruction, Computer-based; Instruction, Intelligent computer-based; Testing; Testing, adaptive; Testing, diagnostic; Tutors, Tutors, Intelligent; Training.
05	09		

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

We describe an adaptive, cognitive approach to diagnostic testing for instruction: steering testing. Steering testing is brief diagnostic testing that steers, or individualizes, the course of instruction. When a diagnosis is needed, the student model (i.e., what is currently known about the student's competence) is examined, and areas of competence about which more information is needed are identified. These areas represent constraints on the type of test item that will be informative. Once the constraints are posted, an intelligent item generator constructs a test item that satisfies them. We discuss steering instruction in two intelligent computer-based tutors for adult skills: avionics troubleshooting and DC-circuit understanding.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT ☒ UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION
22a. NAME OF RESPONSIBLE INDIVIDUAL Dr. Susan E. Chipman	22b. TELEPHONE (Include Area Code) (202) 696-4318 22c. OFFICE SYMBOL ONR 1142Pt

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

Toward Intelligent Systems for Testing

Alan Lesgold
Jeffrey Bonar
Joyce Ivill

Learning Research and Development Center
University of Pittsburgh

One of Robert Glaser's special contributions to psychology and education is the concept of criterion-referenced testing (Glaser, 1963). While norm-referenced testing supports decisions that involve choosing among people or otherwise comparing them, criterion-referenced tests tell us something about what people know or what they can do. In introducing the concept, Glaser was beginning a long advocacy of adaptive education, of shaping education to each person's current competences rather than choosing to educate only the people who score highest on general tests.

While this was his goal, most work on criterion-referenced testing (cf. Hambleton, 1984) has focused on issues relating to certification, to setting of standards for educational outcomes, and to tracking, that is, on selection more than on adaptation. There are a number of reasons for this, but the situation can be summarized as follows. Adaptive education is a steering process. Norm-referenced tests are designed to indicate reliably who is out in front; criterion-referenced tests are designed to tell us exactly where each person is; but knowing where you are is not the same as knowing how to steer a course toward a planned destination.

The purpose of this chapter is to illustrate one way in which the technologies of testing might combine with certain cognitive science techniques to help steer instruction. We focus on steering an intelligent tutor, i.e., on student modeling. However, the approach can be generalized to other instructional forms, including reactive environments (exploratory microworlds) and perhaps even conventional classroom instruction. We are discussing diagnostic testing to be used often, in small amounts, to steer the course of instruction. Further, in contrast to relatively standard (e.g., pretest-treatment-posttest) designs for individualizing the teaching of children, we focus on individualizing the testing process to make it more efficient in steering instruction.

Problems of Diagnostic Testing

Any test, including a diagnostic test, consists of a number of items. The person being tested carries out some performance of each of the items, scores are assigned to those performances, and those scores are aggregated to arrive at an evaluation. To make steering tests, we need test items that are relevant to the specific steering decisions that must be made about a particular student in a particular context, and we need procedures for scoring performance on those items. Steering tests must be efficient to administer, since steering requires frequent, but not necessarily precise, feedback (given the inertia of teaching and learning, the steering error produced by believing an imprecise test will probably be canceled out by subsequent course corrections).

Standard psychometric methods are not designed for steering tests. They are designed to assure that different forms of a test are equivalent and that the scores on that test are reliable. The problem of steering tests is that they must be brief, so that testing does not take too much time from learning. This makes it difficult for them to be reliable, and steering requires at least some reliability of feedback to be successful.

There are two ways a test can be made more reliable. The first is to increase the extent to which performance on its items directly reflects the skills one wishes to assess. This can be done statistically or substantively. Statistical approaches such as item-response theory (Lord, 1980) help assure that different items are measuring the same thing, and thereby increase the reliability of scores, but not necessarily their validity. However, it is also possible to develop a microtheory of the competences one wishes to teach. Such a microtheory can help in specifying items that test particular subsets of the target skills.

The second way to make a test more reliable is to use knowledge about the student's performance on prior items to limit the information each new test item must provide. Adaptive testing algorithms have been developed for this purpose. They use a sequential strategy. After the student completes an item, an estimate of the student's performance based upon the items so far completed is used to select the most informative next item to administer, and then the score on that next item is used to update the estimate. The adaptive testing approach, which almost always requires a computer for the real-time estimates just mentioned, can be applied even when nothing more than the difficulty ordering of items is known. However, it is especially powerful if more detailed information about the items is available. Again, a theory that relates performance on various test items to underlying competences and their acquisition can be helpful, even if it is incomplete.

In at least one case, adaptive testing techniques were applied to diagnostic testing (Spineti & Hambleton, 1977). Spineti and Hambleton used learning hierarchies specified by rational task analysis (Gagne, 1965) to help constrain the estimation process. That is, they decided on items according to an analysis of the material being learned and to some theoretical predictions of the order of acquisition for parts of that material. Doing this, they were able to achieve a 50% reduction in the number of items required to achieve a given level of score reliability.

The approach we have taken to steering testing is somewhat different. It uses very simple heuristics for reasoning about the level of a student's competence in particular subskills. Its power derives primarily from its ability to intelligently manufacture practice opportunities (test items) for the student that will be especially revealing about his current competences. We believe, although it remains to be proven, that these practice opportunities are generally appropriate learning vehicles as well as test items. In that sense, we are pursuing steering as a unified system in which testing and learning are combined.

In our view, a cognitive theory of testing, and especially a theory of steering testing, should have two characteristics. First, it should permit a partly logical (in contrast to a purely statistical) constraining of diagnosis. Second, it should be based on a representation of the knowledge that is needed to exercise the skill it purports to measure. The logical approach is not at all foreign to our experience. When one is sick and goes to a physician, one is not satisfied with broad probabilistic statements. Rather, one expects a diagnosis constrained by the physician's knowledge of disease. More specifically, we expect the physician to be asking herself what diseases could produce the overall complex of symptoms and signs presenting themselves to her. Diagnosis in medicine, then, is the designing of a personal theory of a specific patient's pathology. This personal theory is rooted in theories of disease mechanisms and not just in unexplained statistical relationships.

The diagnosis process is dynamic. For example, based on the hypothesis that a patient has heart disease, the physician may probe for more explicit detail about certain symptoms or order a test that may confirm or refute her theory. A teacher does this too when prior knowledge about a student, combined with current observations, leads her to attribute grammatical errors in the student's paper either to inexperience with written language or to use of nonstandard dialect or to a mistaken sense of when formal conventions are needed.

The good teacher's diagnosis differs from that of a physician in one respect, though. We come to a physician to get a diagnosis when something is wrong -- she does not generally shape continuing decisions about how we should act (except perhaps in developing special regimens, e.g., diets for control of diabetes). A teacher, in contrast, is carrying out an active, goal-directed activity -- teaching -- which needs only small course corrections. Consequently, it [illegible] as reasonable to conduct the testing from the teacher's point of view, at least in part.

We would like to produce tests that capture some of the capabilities of the most perceptive and observant teachers. We want them to be driven mostly from the teacher's goal structures for teaching but also to respond to knowledge of the expertise the teacher is trying to convey, the treatments available to the teacher for effecting learning, and certain more global teacher concerns, such as adapting to general differences in aptitude and general characteristics of competence at different levels of learning.

In the next section, we discuss the different kinds of knowledge that are needed to adapt teaching to an individual student's course of learning. We take the viewpoint of intelligent tutoring system design, but the same concerns arise in all approaches to instruction. This is followed by sections in which a specific approach to the generation of diagnostically and educationally useful problems is discussed.

Components of Teaching and Testing Knowledge

Several different kinds of knowledge are required in our approach to steering testing. Especially when designing computer systems to teach or to test, it is important to clarify the knowledge, or competence, that is involved in dealing with a student. We have categorized that knowledge into four types.[1] These are domain *expertise*, *curriculum* knowledge, instructional steering knowledge, and *treatment* knowledge. Each type of knowledge has different structures, different generalized methods, and different purpose and applicability. Further, there are a variety of connections from one type of knowledge to another. Figure 1 shows these four categories with examples of the kinds of knowledge they contain for an electricity tutoring/testing system under development at the Learning Research and Development Center.

Domain Expertise

Domain expertise is always embodied in instructional decision making, either explicitly or implicitly. Deep diagnosis of student difficulties may require an explicit representation of the knowledge required for the performances that are the goals of instruction. For example, the ability of a computer-based tutor to diagnose bugs (systematic errors) in children's arithmetic performances requires having a model of the algorithms that experts use in executing those performances. Also, feedback on test performance and advice to the student may have to be couched in terms of procedures for acting rather than in terms of criteria for outcomes specified in the curriculum. One way or another, the performances that constitute the goals of a curriculum derive from information about the competences that constitute expertise.

Another aspect of domain expertise that is important in instruction and testing is knowledge of the target task environment. When we speak of what it is we want people to do, we are referring not only to the knowledge they need to perform successfully but also to the circumstances under which that knowledge must be employed. Again, knowledge of these circumstances might be the basis for curricular objectives, but those objectives rest upon domain expertise. If we have the objective that given situation X, the student can do Y, it rests upon knowledge of what kind of situation X is and how Y can be done in X. For example, a student might be able to solve a proportion problem at the time a lesson on proportion is presented but not be able to use that knowledge later in solving a word problem or even to solve the same problem as one of a set on mixed topics. When testing or teaching is done by a computer program, the underlying domain knowledge sometimes must be made explicit.

Curriculum Knowledge

Curriculum knowledge is the specification of the goal structure that guides the teaching of a body of expertise. Educational researchers and developers often treat the procedures that constitute expertise and the instructional goals that constitute curriculum as more or less the same. They assume that expertise can be split apart easily "at its joints" (to use Plato's phrase). The curriculum, then, is a natural hierarchy of goals and subgoals to teach the natural units of expertise. From this viewpoint, curriculum knowledge and domain expertise are the same thing. However, it appears that there are many different plans for splitting apart expertise, especially when expertise involves complex performances. For example, consider the curricular issues that arise in teaching simple electrical principles. There are some basic concepts -- voltage, current, and resistance -- and some basic laws -- Kirchhoff's Laws and Ohm's Law. In addition, there are different types of circuits -- series and parallel.

So, one legitimate decomposition of the subject might begin with voltage, teaching the behavior of voltage in series and parallel circuits, then teaching about resistance in the two types of circuits, and finally treating current. Another decomposition might, with equal legitimacy, build the entire curriculum on Kirchhoff's current laws. Yet another view might treat parallel circuits as being quite distinct from series circuits and redevelop the concepts of voltage, resistance and current separately for each. We need to capture these multiple viewpoints if they correspond to different curricular goals about which steering information may be needed. For this reason, the various subgoals of knowledge that the teacher or curriculum writer can have are best represented by multiple hierarchical goal structures; these goal structures overlap in the components of expert performance to which they refer.

Once we concede that instructional goals are not really a simple decomposition of the expertise being taught into discrete sets and subsets, we are in a position to understand why some testing that is part of a curriculum may not be as diagnostic as we would hope. Specifically, we can understand why a student might demonstrate clear competence on a curricular goal that is prerequisite to some other goal but still appear, from the standpoint of the teacher of that second goal, to not have mastered the first. For example, a student may demonstrate understanding of Kirchhoff's Current Law but fail to apply it in a circumstance for which it is relevant. Separating expertise from curriculum allows us to understand such situations better.

Suppose that we consider domain expertise to be represented by a surface. Expert knowledge is, after all, highly interconnected. Even if it is properly represented as some kind of network, it can be approximated by a continuous surface (specifically, a manifold of unspecified dimensionality). We start by assuming that each curricular subgoal corresponds to a region of the expertise continuum. The expertise subset corresponding to a curricular goal will likely be convex, in the sense that if two pieces of knowledge are part of the same curricular goal, then any strong relationship that directly ties them together should also be part of that goal. On the other hand, a curriculum goal's corresponding expertise is not a completely closed set, since concepts it subsumes may well have connections to other knowledge that goes beyond the goal. That is, the edges between the expertise subsets corresponding to different curricular subgoals are not necessarily clean edges with no connections to other knowledge.

The untargeted knowledge lying between the clusters of expertise directly addressed by the curriculum can be important in remediating lack of transfer from a curriculum goal's prerequisites to the final target capability.[2] Ordinarily, instruction is directed at the center of the expertise subset corresponding to a curricular goal (see Figure 2). This helps keep the new knowledge to be taught simple enough to be learned. However, this approach can sometimes backfire. For example, if two bundles of expertise are both curricular goals, their centers may be well taught but their peripheries ignored. For example, I may teach you how to compute the joint resistance of two resistors in series, and this may satisfy an instructional objective. Later, if you need to find the joint resistance of three resistors in order to solve a problem, you may be able to do that or you may not. In either case, simply reteaching the two-resistor algorithm will be insufficient.

If a higher-order curricular goal happens to depend upon the integration of the two lower-order subgoals, it is exactly the edges of their domain knowledge subsets on which it will likely depend. For decisions about what to teach when remediation seems necessary and also for decisions about how to interpret apparent inconsistencies in diagnosing whether a curricular subgoal has been achieved, domain expertise may be needed.

Planning Knowledge

In addition to specific curricular goals, there are some other higher-order curricular issues that need to be addressed in planning testing or teaching. Often, these are abstractions from, or specialized viewpoints on, the curricular goal structure. These may include learning skills, problem solving heuristics, rather general aptitudes, and even preferences. These concerns, e.g., the more general "inquiry" skill goals in a science course, overlap some of the higher-level goals in the curriculum. It could even be argued that these concerns really are part of the curriculum, but we retain the distinction since planning issues often color the exact form that goal-specific instruction might take.

For example, we would treat as a planning issue the complexity of arithmetic computation that is required to solve a word problem in a math course. The metagoal is for the student to be able to advance through the problem-solving part of the curriculum even if his arithmetic skills are developing more slowly than his problem solving skills. So, the arithmetic required in a word problem might be adjusted to keep it simple enough to let new problem solving skills develop. Later, when problem solving skills are strong, the situation might reverse, and increasingly tough arithmetic might be required whenever the student is predicted to find the problem solving tasks easy. Note that the issue of arithmetic skills getting in the way of problem solving could arise in curricula other than math, such as the electrical networks curriculum sketched in Figures 1 and 3. It is for this reason especially that we choose to treat the matter as a metacurricular planning issue. Sometimes capability on skills that are not the focus of instruction will require alteration of instructional and testing strategies for target skills. This is why instruction and testing systems need planning and metacurricular knowledge.

The planning of teaching must also take into account the long-term, higher-order aspects of education: metacognitive skills, mature and flexible preferences, and fundamental principles that apply in many domains. From the point of view of the steering test developer, though, these higher-order issues represent, for the most part, variables to be controlled. We can't really understand whether a student knows how to solve electrical network problems, for example, if his capability is hidden by slow arithmetic performance. So, we have to take account of metacurricular issues in selecting problems for instructional or measurement use. That is, problems can be selected to require domain-specific skills but to assure that the student answering a given problem will not be troubled by weakness on general basic skills that are not the current focus of measurement or instruction. For example, if a student is weak in arithmetic, a problem might be generated that required only small-integer arithmetic. If a different student finds it easier to receive information in graphical form, the information given for a problem might be presented via a diagram, graph, or even photographic image.

Treatment Knowledge

We turn now to the matter of educational treatments and test item development. Even when we know what to teach or what to measure, there remains a separate form of expertise involved in successfully generating a situation in which a piece of knowledge can be exercised. For example, several different types of problems can be created to test understanding of electrical network principles (or to provide opportunities for coached practice). Problems can be quantitative or qualitative. They can deal with unchanging situations or can focus on relative changes in different measurements of a circuit. Since electricity knowledge must be applied in slightly different ways for each type of problem, we could treat problem type as a curriculum issue. However, the knowledge an intelligent system needs about problem categories is different in form from knowledge about curricular goals. This is especially the case when we want to develop problems for practice or for steering tests that require integrated use of several different skill components that are separate curricular goals. The knowledge needed to develop such problems is specific to electricity and to the teaching of electricity.

Practice and testing that requires multiple skills to be combined is an important goal of our work. A contrasting approach is taken in some formal instructional development methodologies such as the Defense Department's ISD (Merrill & Tennyson, 1977) approach. As generally used, that approach consists of complete development and elaboration of the curriculum followed by the development of tests and treatments corresponding to each curricular goal. This seems entirely sensible, an extension of a management-by-objectives approach. However, if this method is applied superficially, difficulties can arise. We have already discussed the problem of too-narrow focusing on core concepts without adequate elaboration and qualification, but there are other, related problems as well. For example, a variety of apprenticeship situations involve simultaneous practice of a wide range of skill components, only some of which may be the current targets of instruction. When practice is provided on each skill component separately, without attention to when each should be used and how they tie together, fragmentary learning results. The instructor can show, on academic-style tests, that the student learned each subskill that was to be taught, but the subskills cannot be put together to solve real-world problems.

This, of course, is a viewpoint that has been taken before. In the world of reading instruction, for example, we have just seen a long period in which holistic approaches have been taken. Similarly, case study approaches to the teaching of medicine and business are driven by the same motivation. There is, of course, some evidence against holistic approaches. For example, Chall (1967) surveyed a number of reading curricula and found that, on average, weaker students benefited from a phonics approach, in which recognition of each individual grapheme was the focus of separate instruction. In the professional world, it is regularly asked how we can be sure that a student who took a case study course really learned everything he should have. "What if I get a disease that was not one of the cases discussed?"

We can be a bit more formal about this problem if we view subskills as productions, actions to be performed under specific conditions. When subskills are taught in isolation, the conditions under which they should apply cannot be specified, since those conditions relate to the broader context of holistic performances. Also, there may be specific productions that are not represented as subgoals for instruction but that are the "glue" needed to combine the productions that were direct curricular targets.

An instructional synthesis of the holistic and componential approaches requires several things, including an understanding of the circumstances under which new subskills or concepts should be *introduced* in isolation even if they are later to be practiced more holistically. Of course, the missing productions, the "glue" that holds together the subskills we target in our curriculum, cannot be taught adequately in vitro; they require holistic instruction. The dilemma is that they also need to be assessed. We may need to help students attend to "gluing" their fragmentary knowledge together if they have trouble doing so on their own. Further, we may not always choose to introduce new pieces of knowledge formally and explicitly, hoping that they will be inferred through rich domain experience. If we take this approach, which may be very efficient, we need to be able to assess later whether there are any subgoals that were not well attained.

The basic approach we have taken is to generate test items (and instructional treatments, for that matter) in the course of testing. That is, at any given point in the course of testing, if a question arises about a specific curricular goal, a test item is generated for it by an intelligent subsystem of the tutoring program we (primarily the second author) are developing. The items can be shaped by metacurricular considerations. Further, if multiple skills are required for any realistic performance within the domain, sets of items can be developed over which particular subskill requirements are systematically varied.

So, our approach, given a family of cognitive analyses (of expertise, metacurricular issues, and problem environments in which the expertise can be manifested or practiced), is to intelligently generate the equivalent of a controlled experiment in which the need for various target pieces of knowledge is systematically varied. If the student fails to perform items requiring a piece of knowledge but does perform other items that do not require it, then we infer that work is needed on that knowledge. Further, we ask only about pieces of knowledge that are in the part of the curriculum through which we are steering. Finally, rather than make statistical decisions about whether a piece of knowledge is present or absent, we assume that knowledge can be present at various strength levels and use experience about the reliability with which a particular piece of knowledge manifests itself to specify the level of learning of that knowledge.

Summary. Perhaps the best way to illustrate the ideas just presented is to refer back to the example given above. Figure 3 elaborates the knowledge categories, in part, for our system to teach and test basic electricity principles. The curriculum knowledge includes three sets of goals: laws, concepts and architectures. Under each of these are subgoals. For example, the architectures being considered are series and parallel circuits (i.e., no bridge circuits). The planning knowledge includes two sets of planning concerns: the arithmetic difficulty of problems that are presented to the student and the circuit complexity. Both apply with respect to a variety of curricular subgoals. For example, circuit complexity may affect whether a student can handle parallel circuits, whether he can apply Kirchhoff's current law, etc. Arithmetic difficulty could also affect these subgoals, especially if quantitative problems are presented to the student. The treatment knowledge includes information on problem formats and feedback to the student. Finally, the domain expertise contains specific details of expertise in handling electrical networks that are referenced by the curriculum specification.

Generating Test Items from a Student Model

Having described the architecture of the knowledge in a steering testing system, we turn now to how one uses that knowledge to do assessment driven by a cognitive model of the target capabilities being taught. We offer as a first approximation an approach that has been tested in prototype form in an intelligent tutor. It assumes additional knowledge that we have not yet discussed: a student model, some sort of knowledge structure specifying which subskills the student is thought to know and which ones not.

We currently specify the student model by embedding it in the curricular goal structure of an intelligent tutor. For each curricular subgoal, there must be some sort of notation about the student's assumed competence relative to that subgoal. In one tutor the first author and his colleagues are building (Lesgold, Lajoie, et al., 1986), there are only four notations-*unlearned*, *perhaps* acquired, *probably* acquired, and reliably *strong*. These notations relate to an underlying cognitive model of learning derived from John Anderson's (1983) work. The rules currently used to change a subskill notation from one state to another are quite rough, but they are principled.

Movement to the *probably learned* state implies that a correct production, or set of productions, is assumed to have been developed by the student. The *perhaps* state indicates that the student has been observed to perform the target skill component, but that there is insufficient evidence to conclude that he knows the conditions as well as the actions for the subskill. The *perhaps* state is unstable. Either further correct performances will occur, prompting classification to the *probably* state, or we will assume that the single correct performance observed was accidental relative to the problem ecology for the curriculum, and the student will be moved back to the *unlearned* state. Recurrent reliable performance will move a student from *probably* to *strong*. One can imagine other approaches in which the notations might include indicators of misconceptions as well. The important point is that if we look in on a student who is in the midst of learning a skill, some of the subskills will be clearly demonstrated already, some will be manifesting obvious problems, some will be unlearned, and some will be in an unknown status.

If we consider how to diagnose student progress in a holistic practice environment given a current student model state, we see that a first issue to be addressed is what to test. In principle, the student could have learned anything since we last tested him or her. For that matter, any prior demonstration of competence might have been a fluke, so all positive entries in the student model are tentative. Nonetheless, it would make no use of the student model at all if we merely tested for every skill component at every opportunity. The student model enables testing for selected skill components efficiently and in realistic performance contexts. It is the equivalent for steering testing of the patient's chart for medical diagnosis.

We want to use the student model to generate constraints on the problems we pose to the student as test items. These constraints should have the property that they make the items maximally informative in tuning the student model to changes in the student's capabilities. What can guide our choices of curricular goals to test? There are several possibilities. We discuss them in terms of the four-level model of acquisition mentioned above (*Unlearned*, *Perhaps* learned, *Probably* learned, and *Strong*). The *Perhaps* stage may be the most volatile. Suppose a curricular goal to be the attainment of a specific production (carrying out a particular action when appropriate). When the action is initially performed and is successful, there is a considerable chance that the student may not notice the most important cues about the circumstance of the moment. So, he/she may be unable to demonstrate the production in other circumstances. For all practical purposes, it was never really learned at all. Till we have several demonstrations of the attainment of a curricular goal, we must assume that our assessment of the student is unstable. Once we see multiple successful performances, we will reclassify the student's competence to the *Probably* level. So, a first principle in selecting current curricular goals to test is to be sure to check up on goals in the *Perhaps* state.

A second issue has to do with prerequisite skills. If *Skill A* depends upon *Skill B*, then there is no point in regularly testing for *A* until *B* is demonstrated. Put another way, if there is ordering information about the curricular goals, we may want to concentrate testing on the region in the ordering between the goals in the *Strong* state and those in the *Unlearned* state, testing most often the *Perhaps* goals, checking for progress on the next few *Unlearned* goals, and checking occasionally to see if any goals have gone from *Probably* to *Strong* (operationally, we check to see if problems requiring this subgoal's skills are answered correctly for several consecutive occasions with varying requirements).

The next issue involves metacurricular concerns, especially those relating to extraneous sources of difficulty, such as requiring complicated arithmetic performance, presenting information in a medium known to be difficult for the student, etc. The basic rule of thumb we propose is to adapt these difficulty variables to the current student model level. For example, if the goal is to detect a movement from *Unlearned* to *Perhaps* for some curricular goal, then we want to set the metacurricular difficulty levels low, so that the initial weak acquisition of that subgoal's knowledge is not masked by too many other demands for processing capacity. For movement from *Perhaps* to *Probably*, an appropriate problem constraint is to have some situational changes from the problem in which the initial appearance of the relevant knowledge was first noted, since the theoretical motivation for the distinction is the possibility of the correct actions having been linked to imprecise conditions. For validating movement to *Strong* on some goal, there should be a demonstration of the relevant capability under more difficult circumstances, since the question is whether the relevant knowledge is robust enough to occur even under adverse conditions.

The Concept of Constraint Posting

The basic approach is to begin each cycle of diagnosis by sweeping through the curricular goal structure, noting which subskills are "ripe" for testing. When the sweep is completed, we try to build one or more problems that maximize our chances for accurately noting changes in the student's current knowledge state, using some of the rules of thumb just described. We then use performance on these made-to-order problems to decide how to update the student model -- we make a diagnosis.

Critical to the approach is the concept of *constraint posting* (Stefik, 1980). Rather than building test items as we sweep through the curricular goal structure, we instead simply add to a list of item constraints as we proceed. Each time we see an issue on which we would like more clarity, we post that concern as a constraint on the test item generation process. When the sweep through the curriculum is complete, we take the bundle of constraints and try to build items that satisfy them. Stefik (1980) has shown that in many complex problem solving tasks involving multiple sources of complexity and interactions between problem aspects (e.g., designing recombinant DNA experiments), this constraint posting approach is much more efficient than piecemeal search processes.

Constraint Posting Applied to Problem Generation

The item generation process, then, can work as follows. We first consider the student model. Some of the subskills may be marked as reliably strong. These represent beachheads in the conquest of ignorance. From these beachheads, as we venture out toward related subskills, we find some whose status is uncertain (subskills that may or may not have been acquired yet and acquired subskills that may or may not be reliable yet). We can make this search process more efficient if we know, for some subgoals, which other subgoals are prerequisite to them and which they are prerequisite for. A subgoal for which a just attained subgoal is prerequisite is likely to be a testing target, but we will also give some weight to all subgoals, using the rules of thumb discussed above. Since we are making steering decisions, we focus on the area of the curriculum that is currently the object of instruction. For each subgoal that is a current target of testing, at least one constraint is posted: a test problem must address that subgoal. For example, if we want to find out whether the student's capabilities in applying Ohm's Law to series circuits have improved, we post constraints that the problem must require Ohm's Law and must involve a series circuit.

We must also consider metacurricular planning issues. For example, a part of the system's planning component may address the question of whether or not a physics student has adequate math facility, or whether or not a student is able to learn information from graphical presentations. Constraints can be posted based on metacurricular aspects of the student model, too. We may, essentially, say to the test generator, "Since this student is poor in arithmetic, I can't find out if he has learned (moved from *unlearned* to *perhaps*) how to use Ohm's Law to compute the current in a circuit if the arithmetic comes out messy, so make the numbers come out simple."

Once the sweep through the curricular and planning structures is complete, the posted constraints must be analyzed before test items are generated. Are there too many to handle at once? If so, we might partition them into several clusters. Are the constraints inconsistent, in the sense that a problem embodying some of them cannot, in principle, embody the others? For example, if we constrain an electricity problem to be simple and we want to know both whether a student knows how to deal with two resistors in series and also whether he knows how to deal with two in parallel, this cannot all be done with one circuit problem. So, again, we might partition the constraints into bundles that can comfortably be handled.

Finally, one or more holistic problems that satisfy the constraints posted must be posed. From performance on a problem, either a diagnosis can be made immediately or a more focused problem can be specified for further testing. In essence, we are dealing with a qualitative process that has many of the properties of one of psychometrics' most important quantitative processes -- adaptive testing.

An Example from a Tutor for Basic Electricity Principles

To illustrate some of these ideas, we describe MHO, a tutor that teaches basic electrical principles (current, voltage, and resistance; Kirchhoff's Laws and Ohm's Law). MHO is designed to work in both a problem-posing and an exploration mode. In the exploratory mode, the student can make measurements on circuits and even build his own circuit. In the didactic mode, though, MHO must decide what problem to present to the student. Thus, it faces the same problem that a testing program would face: to examine the student model and determine which problem to pose to optimize the information value of the student's answer.

MHO's student model is a specialized form of checklist: a goal structure for teaching the specific knowledge it wants to teach. The checklist derives from the curriculum and planning issues shown in Figure 3 above. For each subgoal, the student is marked as being in one of the four states described above, as shown in Table 1. Quantitative scores could be entered as well. What is critical is that some student knowledge levels are considered to indicate potential for change while others are not. For example, a student who knows certain material is not likely to suddenly stop knowing it, but a student who has yet to learn some material is in a more changeable state.

From the subgoal scores and other knowledge, such as curricular sequencing and prerequisite relationships, it is possible to define a set of subgoals that are most unstable. These are the subgoals that may require more frequent measurement in order for instruction to be steered well. As discussed above, they represent the front along which instruction is progressing through the curriculum goal structure. The task of a test item generator, then, is to generate a test item that will be especially informative about this front. MHO does this by posting a set of constraints for the test problem. In the student model given above, the Series, Kirchhoff's Law, and Current subgoals are at this front. Each constraint helps adapt the steering feedback to the student's current state. To see how this is done, we need to consider MHO's architecture and the subject matter that it teaches and tests.

Architecture

At this time, MHO teaches and tests several levels of DC circuits. It poses problems such as the one shown in Figure 4. We call the architecture used in MHO the *Bite-Size Architecture*. It is an object-oriented architecture for intelligent tutoring systems.[3] An object is a semiautonomous piece of computer program that can be called upon to achieve particular goals. It includes both data structures and procedural capabilities. Object-oriented programming involves designing sets of objects that can efficiently interact to solve problems. Each curriculum subgoal (and also each metacurricular planning issue and each problem format) is represented by an object called a "bite." Within the computer program, a bite contains a record of the student's performance on a subgoal and the knowledge needed to post a constraint for that subgoal.

Voltage, for example, is represented by a bite in MHO. That bite has rules for teaching about voltage. It contains information pertinent to developing an understanding of what voltage represents, including the constraints it should post to create relevant problems. Also, it can update the student model information by noting how the student does on problems relevant to its subgoal. One byproduct of this architecture and the curricular model on which it is based is that a tutoring program's knowledge is modular and can easily be expanded by adding additional curricular objects along with their pointers to the other knowledge components (which may involve additions to those components as well). For example, MHO's designers are now expanding it to include curricular goals involving simple alternating current circuits.

Problem Generation

MHO poses problems by presenting a circuit diagram and asking a question about it. The machinery used in problem generation chooses most of the circuit components randomly, but it is constrained by both general and specific curricular subgoals (bites) which the student has not yet mastered. Some of the choices represented by these constraints are the following:

a. A problem can be posed in qualitative, quantitative or relative form.

b. The problem can vary in the complexity of the arithmetic it requires and the complexity of the circuit diagram to which it refers. This is determined by a global assessment of how much of the curriculum the student has mastered.

c. The problem can require knowledge of Ohm's Law or either of Kirchhoff's Laws.

d. The problem can focus on voltage, current or resistance.

e. The problem can focus on series or parallel circuit topologies. (MHO also worries about where the meters are placed in circuit diagrams, since there are some placements that students have particular difficulty handling, but we ignore that matter to make presentation of the basic approach more straightforward).

The product of constraint posting is stored as a list structure (see Footnote 3) to be used as the basis for problem generation and problem solving. This list structure contains information that specifies how to create a circuit and a problem based on that circuit, what the circuit should look like, and what electronic concepts are relevant. An example of such a list, derived from the student model shown in Table 1) is:

[1] *((((Rel Simple) ($ Kirchhoff)) ($ I=Series)) (UninterruptedS)) Series).*

This list represents the constraints that have been posted in sweeping the model shown in Table 1 and is the starting point for automatic generation of a problem. *Rel* stands for a Relative problem that will pose a simple question asking if two areas of the circuit will have the same measurement (in this case, current). *Simple* specifies the student's level of general understanding and will cause the circuit to be very simple in structure. *Kirchhoff* is the law this problem centers around. *I=Series* is a specialization of Kirchhoff's law, that *current* is equal at all points in a series array. *UninterruptedS* informs the problem generator that one meter should appear next to another with no other components between them (this is the simplest form for a problem looking at Kirchhoff's Law). Constrained by this information, the problem generator can develop many different circuits and pose many different problems about them, so it is quite plausible to do as much steering testing as any student requires and also to give students sets of appropriate problems as homework.

At the next, more elaborated, level of representation the circuit is designated as a network of resistors, a combination of series and parallel subnets with a power source. A more detailed list breaks this circuit into four nodes, each of which represents a side of a rectangular circuit. The nodes are created separately and then put together to make up a circuit. One at a time, the nodes are passed into a recursive function called **MakeCircuitString** to be elaborated further. **MakeCircuitString** makes decisions such as how many resistors are placed on a node, and whether these resistors should appear in a parallel or series net. These decisions are based on the information from the first list.

Simple instructs **MakeCircuitString** to limit the number of resistors that appear and to otherwise make the circuit conform to the specifications of a simple circuit. The *Simple* specifications keep the components that will be drawn to a minimum. *Simple* also informs **MakeCircuitString** that depending on what net we are working with all nodes should be of this kind. *I=Series* specifies the net to be used: all sides are series arrays. If this were a *Difficult* problem, some sides might have parallel subnets and others series. An example of a simple circuit, [1], that has passed through **MakeCircuitString** is

[2] *((VoltageSource) (Series (Resistor) (Resistor)) (Parallel (Resistor) (Resistor)) (Wire)).*

Figure 5 below shows the circuit designated by [2].

The final specifications development step is determining what problem should be posed about the circuit, where meters should be placed and what question should be asked about them. This step requires some information from the first list, e.g. [1]. *I=Series* reveals whether current or voltage is the target concept, while *UninterruptedS* holds information pertaining to how many problems and where problems should appear. Several recursive functions tear apart the second list and insert problem information (mainly meters) where it is best suited. Using the above example and placing several meters into the list, one example of the next stage is

[3] *((Problem Rel current after on (VoltageSource)) (Series (Resistor) (Problem Rel current before off (Resistor)) (Parallel (Problem Rel current after on (Resistor) (Resistor)) (Wire)).*

This list is then passed to an intelligent problem developer, which composes and draws the circuit. Figure 6 below shows a display corresponding to [3]. The question posed to the student will end up being, *"Is the current at Meter A higher, lower, or the same as the current at Meter B?"*

The Simulator assigns values to the components, i.e. resistance and voltage, and then finds the dependent values, i.e. current, voltage drops over resistors, etc. It can, for simple problems, ensure that all the values for current and voltage will be integral, and also can determine whether or not resistors and voltage sources should be displayed. If the circuit were more complex, an iterative propagation would occur next. Resistance for a subnet of a complex circuit, for example, would be calculated by asking each subnet component its resistance and then adding them together. Parallel structures are handled recursively as well, using the appropriate formulae.

The Softness of Student Classifications

We conclude by reconsidering more broadly the issue of diagnostic assessment of cognitive skills to steer instruction. Fundamentally, cognitive skill, like physical skill, often requires substantial practice of its basic components in the contexts in which they are to be applied. Actions can be learned without learning the exact conditions for which they are appropriate. Newly learned, and consequently weak, knowledge can fail to be used because stronger but incorrect knowledge is overgeneralized from related situations. Processing capacity demands due to one subskill may be so great as to make the execution of another, newly formed subskill impossible. This means that for most of the course of learning, a fundamental principle is true:

One cannot be sure a subskill has not been learned just because it was not demonstrated on an occasion where it should have been.

On the other hand, cognitive skill, like physical skill, is partly redundant. Weak methods can sometimes overcome the lack of appropriate domain knowledge. Sometimes, a problem that in theory should require a particular subskill is solved correctly by accident. The correct action may be taken with incorrect knowledge of the conditions under which it is appropriate, or an incorrect action may turn out to be "safe" this time only. This leads to a second fundamental principle.

One cannot be sure a subskill is completely learned just because it has been demonstrated.

These two principles suggest that the steering approach to diagnostic testing, in which local microtesting is embedded in the curriculum to steer instruction, is a more valid approach than the broader diagnostic testing that has become part of many current monitoring programs in our schools. By asking broad, generic questions (e.g., "What can I diagnose knowing nothing about the student in advance and giving only a general test?") we can get only broad, generic answers. That is, we can know how well, in general, learning is proceeding, but we can't steer specific children's education with such broad indicators, any more than we could steer a ship if all we had was an hourly account of how close to the correct path we were.

Empirical experience and cognitive theory tell us that an inherent property of cognitive performance is that it is unreliable unless substantial practice has occurred and that success can come for multiple reasons. These factors have to be taken into account in diagnosis. Ironically, perhaps, the less reliable steering testing approach provides better steering capability than the highly refined approaches used in current psychometric efforts at diagnosis. But this is no different than the irony that continuously knowing approximately where you are affords better steering capability than occasionally knowing how well you are steering, in general.

The field of testing has worked to try to become efficient at making precise estimates from inherently unreliable data, and it has done very well at this. Approaches such as item-response theory and adaptive testing have allowed the broad and vague measures that tests provide to be made ever more efficiently. Further progress, and especially progress in steering testing (as opposed to certification and selection testing) will depend on better use of information we already have, or can readily get, about the cognitive requirements of the performances and student competences relative to those performances that interest us. Like the physician, we will, in steering the course of a child's education, be better guided by sketchy data tied to specific theoretical analysis than by precise, but general, indicators.

Our approach can be contrasted to the steering forms used in the curricula that grew from Bob Glaser's work on individualized instruction. There, the steering idea was also used. However, the technology of the time did not permit more than a short, uniform mastery test after each lesson. This allowed adequate teaching of the higher-aptitude student but did not handle the remediation problem discussed above. That is, it suffered from having to treat each curricular goal and its corresponding student capability as separable from every other, and it could not handle the problem of core learning without fringe transfer. There was much discussion during the period of that curriculum development about having remediation that was more than just doing the same thing again. The present approach to steering testing, which permits adaptation grounded in cognitive analysis of the instructional domain, rests on the goal structure for educational research established during the period of work by Bob Glaser and his colleagues on individually-prescribed instruction.

References

Anderson, J. R. (1983). *The architecture of cognition.* Cambridge, MA: Harvard University Press.

Bonar, J. G., Cunningham, R., & Schultz J. (September 1986). An object-oriented architecture for intelligent tutoring systems. *Proceedings of the first annual conference on object oriented programming, systems, languages, and applications.*

Chall, J. (1967). *Learning to read: The great debate.* New York: McGraw-Hill.

Gagne, R. M. (1965). *The conditions of learning.* New York: Holt Rinehart and Winston.

Glaser, R. (1963). Instructional technology and the measurement of learning outcomes. *American Psychologist, 18,* 510-522.

Glaser, R., Lesgold, A. M., Lajoie, S. P. (In press). Toward a cognitive theory for the measurement of achievement. In Glover, J. (Ed.), *The influence of cognitive psychology on testing and measurement.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Hambleton, R. K. (1984). Criterion-referenced measurement. In T. Husen & T. N. Postlethwaite (eds.), *International encyclopedia of education.* New York: Pergamon Press.

Lesgold, A. M. (in press). Toward a Theory of Curriculum for Use in Designing Intelligent Instructional Systems. In H. Mandl & A. Lesgold (Eds.), *Learning Issues for Intelligent Tutoring Systems.* New York: Springer-Verlag.

Lesgold, A. M., Lajoie, S. P., et al. (April, 1986). *Cognitive task analysis to enhance technical skills training and assessment.* Technical report. Learning Research and Development Center, University of Pittsburgh, Pittsburgh, PA.

Lord, F. M. (1980). *Applications of item response theory to practical testing problems.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Merrill, M. D. & Tennyson, R. D. (1977). *Teaching concepts: An instructional design guide.* Englewood Cliffs, NJ: Educational Technology Publications.

Spineti, J. P. & Hambleton, R. K. (1977). A computer simulation study of tailored testing strategies for objective-based instructional programs. *Educational and Psychological Measurement,* 37, 139-158.

Stefik, M. J. (1980). Planning with constraints. *Artificial Intelligence, 16,* 111-139.

Footnotes

This research was supported by a contract from the Office of Naval Research, Personnel and Training Branch, for which the first author is Principal Investigator. The methodology derives from work done under a subcontract from Universal Energy Systems, Inc., for the Air Force Human Resources Laboratory. The contents of this chapter have not been reviewed by either organization and no endorsement by them should be inferred. Arlene Weiner, Ronald Hambleton, and Lauren Resnick provided many helpful comments on an earlier draft.

[1]In Lesgold (in press) a three-category model was presented. Since then, we have become convinced that the curriculum and treatment categories should be separated.

[2]This issue is addressed more completely in Lesgold (in press).

[3]See Bonar, Cunningham and Schultz, 1986, for a description of *An Object-Oriented Architecture for Intelligent Tutoring Systems.* MHO is implemented in Loops, Xerox's proprietary object-oriented specialization of the standard artificial intelligence language Lisp. The graphics and student interface are handled via an interface package called *Chips*. Chips is a program developed at the Learning Research and Development Center, primarily by John D. Corbett and Robert E. Cunningham, with some contribution by Andrew D. Bowen. The *Chips* tools allow circuit displays to be designed so the student can click the mouse (a mouse is a pointing device that causes a marker to move on the screen as the device is moved on a table top; it often contains buttons as well, so that the computer user can point to an object on the screen by moving the marker over that object and then pressing a button) on any of the components and thereby cause a menu of query options to appear. Each object can behave differently: when a student clicks on a meter, a question is asked; when he/she clicks on a resistor a special menu of options is presented.

Figure 1. Types of Knowledge Needed in Teaching and Testing.

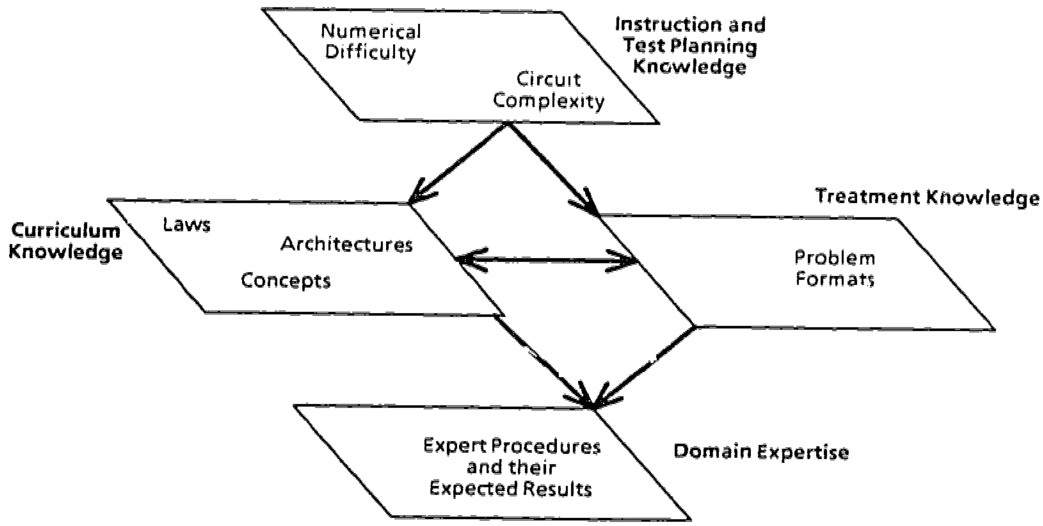

Figure 2. Remedial Knowledge May Not Be Core Knowledge.

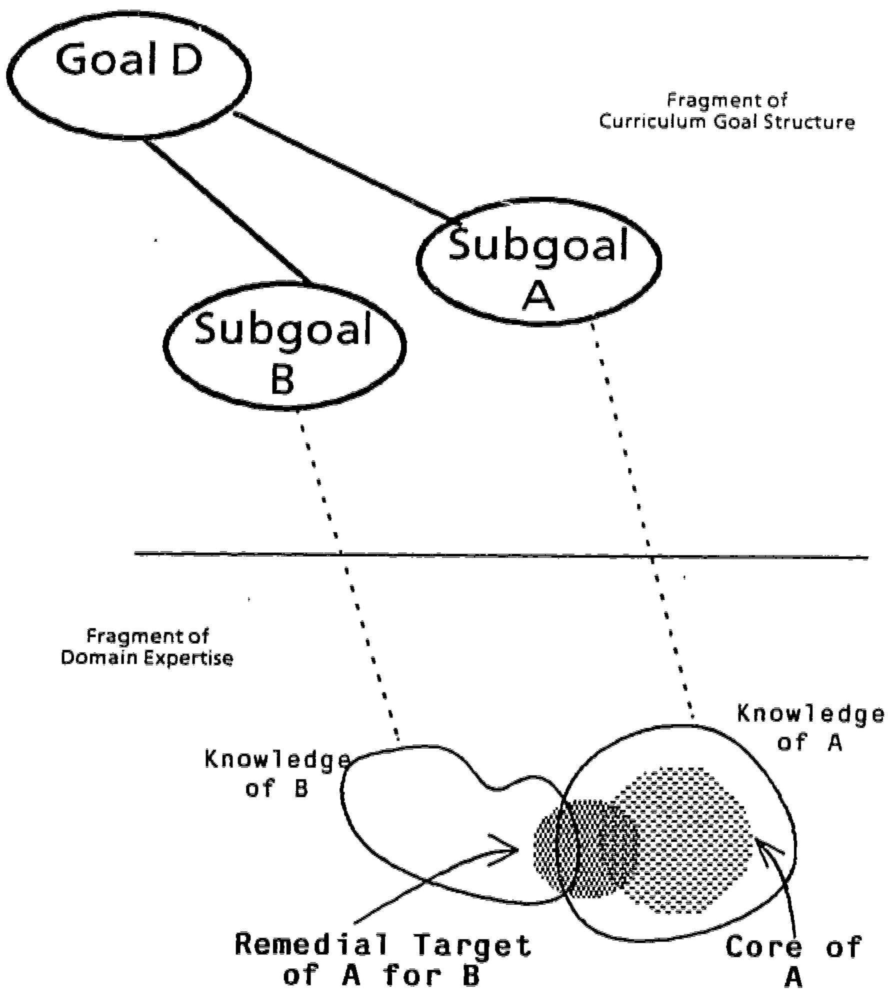

Figure 3. Examples of Different Knowledges Needed for Steering Testing.

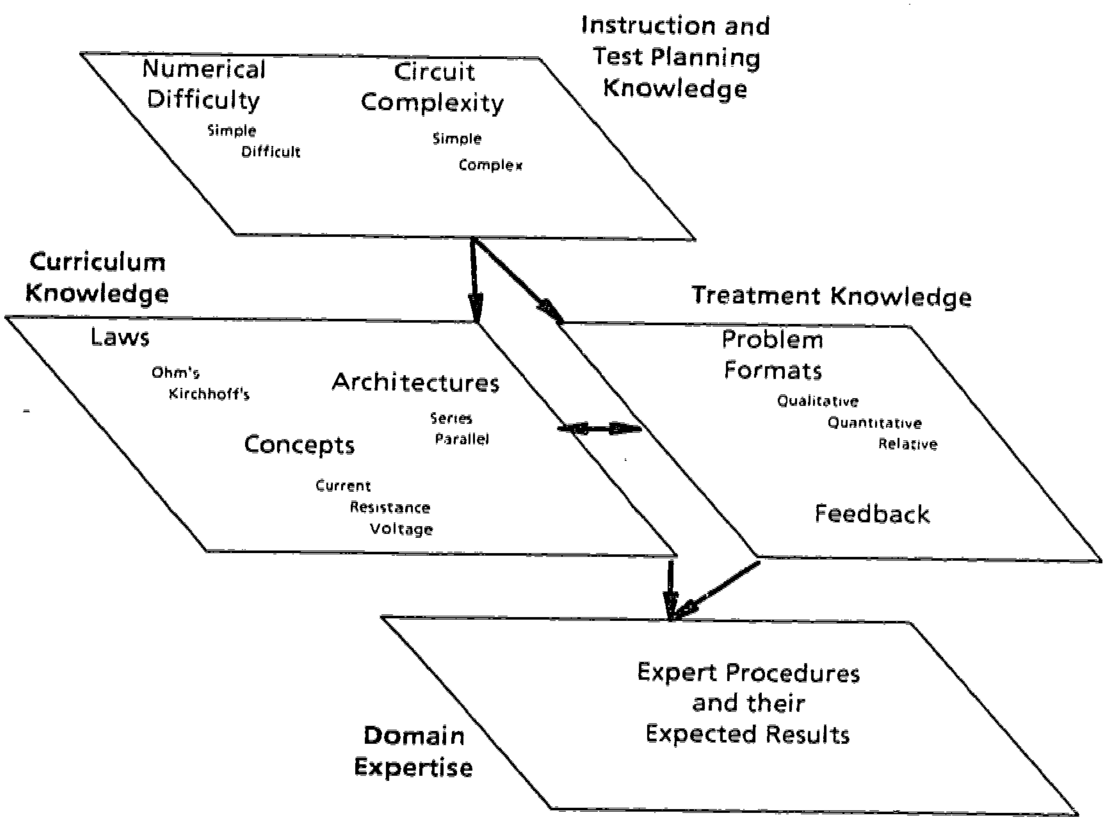

Figure 4. Example Problem from MHO Test Generator.

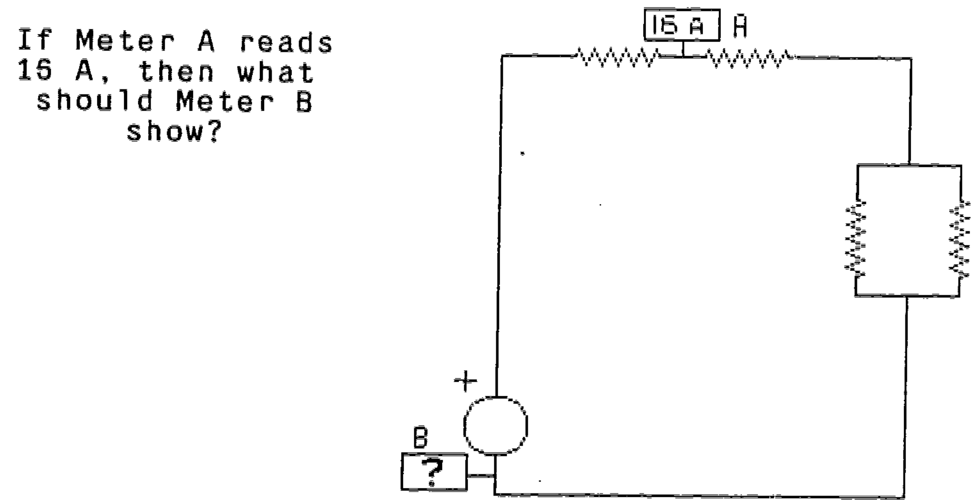

Figure 5. Circuit described by Eq. 2.

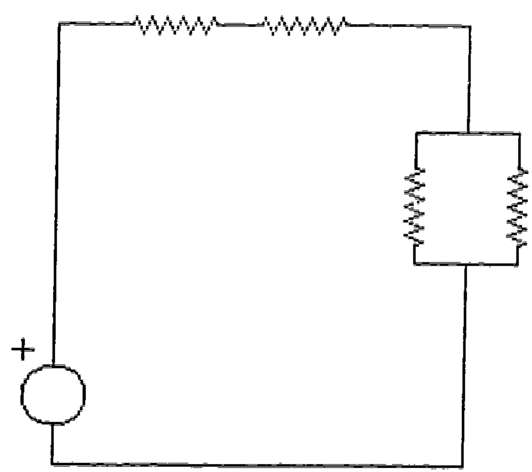

Figure 6. Circuit described by Eq. 3.

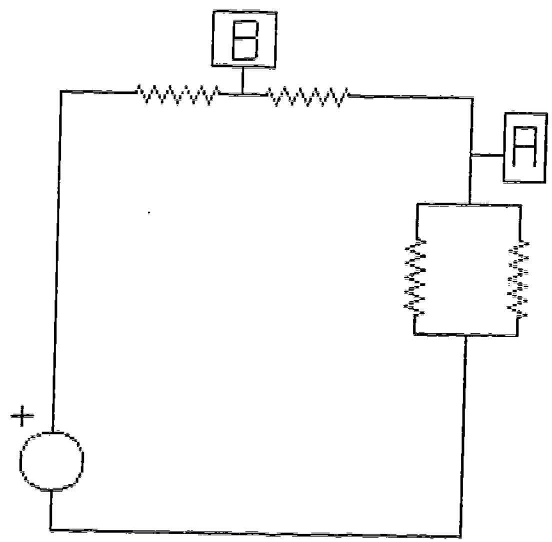

Table 1. Example Student Model.

Metacurricular Issues	
Numerical Difficulty	Simple vs. Difficult
Circuit Complexity	Simple vs. complex
Curricular Subgoals	**Current Student State**
Laws	
Ohm's	Unlearned
Kirchhoff's	Perhaps
Architecture	
Series	Perhaps
Parallel	Unlearned
Concepts	
Current	Perhaps
Resistance	Unlearned
Voltage	Unlearned
Treatment Issues	
Problem Formats	
Qualitative	
Quantitative	
Relative	

1987/03/09

Distribution List [Pittsburgh/Lesgold] NR 4422539

Dr. Phillip L. Ackerman
University of Minnesota
Department of Psychology
Minneapolis, MN 55455

Dr. Beth Adelson
Department of Computer Science
Tufts University
Medford, MA 02155

Air Force Human Resources Lab
AFHRL/MPD
Brooks AFB, TX 78235

AFOSR,
Life Sciences Directorate
Bolling Air Force Base
Washington, DC 20332

Technical Director,
Army Human Engineering Lab
Aberdeen Proving Ground
MD 21005

Dr. Robert Ahlers
Code N711
Human Factors Laboratory
Naval Training Systems Center
Orlando, FL 32813

Dr. Ed Aiken
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. John Allen
Department of Psychology
George Mason University
4400 University Drive
Fairfax, VA 22030

Dr. James Anderson
Brown University
Center for Neural Science
Providence, RI 02912

Dr. John R. Anderson
Department of Psychology
Carnegie-Mellon University
Pittsburgh, PA 15213

Dr. Nancy S. Anderson
Department of Psychology
University of Maryland
College Park, MD 20742

Technical Director, ARI
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Alan Baddeley
Medical Research Council
Applied Psychology Unit
15 Chaucer Road
Cambridge CB2 2EF
ENGLAND

Dr. Patricia Baggett
University of Colorado
Department of Psychology
Box 345
Boulder, CO 80309

Dr. Eva L. Baker
UCLA Center for the Study
of Evaluation
145 Moore Hall
University of California
Los Angeles, CA 90024

Dr. Meryl S. Baker
Navy Personnel R&D Center
San Diego, CA 92152-6800

prof. dott. Bruno G. Bara
Unita di ricerca di
intelligenza artificiale
Universita di Milano
20122 Milano - via F. Sforza 23
ITALY

Dr. William M. Bart
University of Minnesota
Dept. of Educ. Psychology
330 Burton Hall
178 Pillsbury Dr., S.E.
Minneapolis, MN 55455

Dr. Jackson Beatty
Department of Psychology
University of California
Los Angeles, CA 90024

1987/03/09

Distribution List [Pittsburgh/Lesgold] NR 4422539

Dr. John Black
Teachers College
Columbia University
525 West 121st Street
New York, NY 10027

Dr. Jeff Bonar
Learning R&D Center
University of Pittsburgh
Pittsburgh, PA 15260

Dr. Gordon H. Bower
Department of Psychology
Stanford University
Stanford, CA 94306

Dr. Robert Breaux
Code N-095R
Naval Training Systems Center
Orlando, FL 32813

Commanding Officer
CAPT Lorin W. Brown
NROTC Unit
Illinois Institute of Technology
3300 S. Federal Street
Chicago, IL 60616-3793

Dr. John S. Brown
XEROX Palo Alto Research Center
3333 Coyote Road
Palo Alto, CA 94304

Maj. Hugh Burns
AFHRL/IDE
Lowry AFB, CO 80230-5000

Dr. Jaime Carbonell
Carnegie-Mellon University
Department of Psychology
Pittsburgh, PA 15213

Dr. Pat Carpenter
Carnegie-Mellon University
Department of Psychology
Pittsburgh, PA 15213

LCDR Robert Carter
Office of the Chief of Naval Operations
OP-01B
Pentagon
Washington, DC 20350-2000

Chair, Department of Psychology
College of Arts and Sciences
Catholic University of America
Washington, DC 20064

Dr. Michelene Chi
Learning R & D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15213

Dr. L. J. Chmura
Computer Science and Systems
Code: 7590
Information Technology Division
Naval Research Laboratory
Washington, DC 20375

Mr. Raymond E. Christal
AFHRL/MOE
Brooks AFB, TX 78235

Assistant Chief of Staff for Research, Development, Test, and Evaluation
Naval Education and Training Command (N-5)
NAS Pensacola, FL 32508

Dr. Allan M. Collins
Bolt Beranek & Newman, Inc.
50 Moulton Street
Cambridge, MA 02138

Dr. Stanley Collyer
Office of Naval Technology
Code 222
800 N. Quincy Street
Arlington, VA 22217-5000

Brian Dallman
3400 TTW/TTGXS
Lowry AFB, CO 80230-5000

Distribution List [Pittsburgh/Lesgold] NR 4422539

Dr. Diane Damos
Arizona State University
Department of Psychology
Tempe, AZ 85287

Dr. Denise Dellarosa
Department of Psychology
Yale University
Box 11A, Yale Station
New Haven, CT 06520

Dr. R. K. Dismukes
Associate Director for Life Sciences
AFOSR
Bolling AFB
Washington, DC 20332

Dr. Stephanie Doan
Code 6021
Naval Air Development Center
Warminster, PA 18974-5000

Dr. Emanuel Donchin
University of Illinois
Department of Psychology
Champaign, IL 61820

Defense Technical
Information Center
Cameron Station, Bldg 5
Alexandria, VA 22314
Attn: TC
(12 Copies)

Dr. Susan Embretson
University of Kansas
Psychology Department
426 Fraser
Lawrence, KS 66045

Dr. Randy Engle
Department of Psychology
University of South Carolina
Columbia, SC 29208

Dr. William Epstein
University of Wisconsin
W. J. Brogden Psychology Bldg.
1202 W. Johnson Street
Madison, WI 53706

ERIC Facility-Acquisitions
4833 Rugby Avenue
Bethesda, MD 20014

Dr. K. Anders Ericsson
University of Colorado
Department of Psychology
Boulder, CO 80309

Dr. Martha Farah
Department of Psychology
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

Dr. Beatrice J. Farr
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Marshall J. Farr
Farr-Sight Co.
2520 North Vernon Street
Arlington, VA 22207

Dr. Paul Feltovich
Southern Illinois University
School of Medicine
Medical Education Department
P.O. Box 3926
Springfield, IL 62703

Dr. Craig I. Fields
ARPA
1400 Wilson Blvd.
Arlington, VA 22209

J. D. Fletcher
9931 Corsica Street
Vienna VA 22180

Dr. Kenneth D. Forbus
University of Illinois
Department of Computer Science
1304 West Springfield Avenue
Urbana, IL 61801

Dr. John R. Frederiksen
Bolt Beranek & Newman
50 Moulton Street
Cambridge, MA 02138

1987/03/09

Distribution List [Pittsburgh/Lesgold] NR 4422539

Dr. Alfred R. Fregly
AFOSR/NL
Bolling AFB, DC 20332

Dr. Michael Friendly
Psychology Department
York University
Toronto ONT
CANADA M3J 1P3

Julie A. Gadsden
Information Technology
Applications Division
Admiralty Research Establishment
Portsdown, Portsmouth PO6 4AA
UNITED KINGDOM

Dr. Michael Genesereth
Stanford University
Computer Science Department
Stanford, CA 94305

Dr. Dedre Gentner
University of Illinois
Department of Psychology
603 E. Daniel St.
Champaign, IL 61820

Dr. Lee Giles
AFOSR
Bolling AFB
Washington, DC 20332

Dr. Robert Glaser
Learning Research
& Development Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

Dr. Marvin D. Glock
13 Stone Hall
Cornell University
Ithaca, NY 14853

Dr. Sam Glucksberg
Department of Psychology
Princeton University
Princeton, NJ 08540

Dr. Daniel Gopher
Industrial Engineering
& Management
TECHNION
Haifa 32000
ISRAEL

Dr. Sherrie Gott
AFHRL/MODJ
Brooks AFB, TX 78235

Dr. T. Govindaraj
Georgia Institute of Technology
School of Industrial & Systems
Engineering
Atlanta, GA 30332

Dr. Richard H. Granger
Department of Computer Science
University of California, Irvine
Irvine, CA 92717

Dr. James G. Greeno
University of California
Berkeley, CA 94720

Dr. Henry M. Halff
Halff Resources, Inc.
4918 33rd Road, North
Arlington, VA 22207

Dr. Bruce Hamill
The Johns Hopkins University
Applied Physics Laboratory
Laurel, MD 20707

Dr. John M. Hammer
Center for Man-Machine
Systems Research
Georgia Institute of Technology
Atlanta, GA 30332

Dr. Ray Hannapel
Scientific and Engineering
Personnel and Education
National Science Foundation
Washington, DC 20550

Dr. Harold Hawkins
Office of Naval Research
Code 1142CS
800 N. Quincy Street
Arlington, VA 22217-5000

1987/03/09

Distribution List [Pittsburgh/Lesgold] NR 4422539

Dr. Barbara Hayes-Roth
Department of Computer Science
Stanford University
Stanford, CA 95305

Dr. Frederick Hayes-Roth
Teknowledge
525 University Ave.
Palo Alto, CA 94301

Dr. Joan I. Heller
505 Haddon Road
Oakland, CA 94606

Dr. Geoffrey Hinton
Carnegie-Mellon University
Computer Science Department
Pittsburgh, PA 15213

Dr. James D. Hollan
MCC,
Human Interface Program
3500 West Balcones Center Dr.
Austin, TX 78759

Dr. John Holland
University of Michigan
2313 East Engineering
Ann Arbor, MI 48109

Dr. Melissa Holland
Army Research Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Robert W. Holt
Department of Psychology
George Mason University
4400 University Drive
Fairfax, VA 22030

Ms. Julia S. Hough
Lawrence Erlbaum Associates
6012 Greene Street
Philadelphia, PA 19144

Dr. James Howard
Dept. of Psychology
Human Performance Laboratory
Catholic University of
America
Washington, DC 20064

Dr. Earl Hunt
Department of Psychology
University of Washington
Seattle, WA 98105

Dr. Ed Hutchins
Intelligent Systems Group
Institute for
Cognitive Science (C-015)
UCSD
La Jolla, CA 92093

Dr. Janet Jackson
Rijksuniversiteit Groningen
Biologisch Centrum, Vleugel D
Kerklaan 30, 9751 NN Haren (Gn.)
NETHERLANDS

Dr. R. J. K. Jacob
Computer Science and Systems
Code: 7590
Information Technology Division
Naval Research Laboratory
Washington, DC 20375

Dr. Zachary Jacobson
Bureau of Management Consulting
365 Laurier Avenue West
Ottawa, Ontario K1A 0S5
CANADA

Pharm.-Chim. en Chef Jean Jacq
Division de Psychologie
Centre de Recherches du
Service de Sante des Armees
108 Boulevard Pinel
69272 Lyon Cedex 03, FRANCE

Dr. Robert Jannarone
Department of Psychology
University of South Carolina
Columbia, SC 29208

Dr. Claude Janvier
Directeur, CIRADE
Universite' du Quebec a Montreal
P.O. Box 8888, St. "A"
Montreal, Quebec H3C 3P8
CANADA

1987/03/09

Distribution List [Pittsburgh/Lesgold] NR 4422539

COL Dennis W. Jarvi
Commander
AFHRL
Brooks AFB, TX 78235-5601

Dr. Robin Jeffries
Hewlett-Packard Laboratories
P.O. Box 10490
Palo Alto, CA 94303-0971

Dr. Douglas H. Jones
Thatcher Jones Associates
P.O. Box 6640
10 Trafalgar Court
Lawrenceville, NJ 08648

Dr. Marcel Just
Carnegie-Mellon University
Department of Psychology
Schenley Park
Pittsburgh, PA 15213

Dr. Daniel Kahneman
Department of Psychology
University of California
Berkeley, CA 94720

Dr. Milton S. Katz
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Steven W. Keele
Department of Psychology
University of Oregon
Eugene, OR 97403

Dr. Wendy Kellogg
IBM T. J. Watson Research Ctr.
P.O. Box 218
Yorktown Heights, NY 10598

Dr. David Kieras
University of Michigan
Technical Communication
College of Engineering
1223 E. Engineering Building
Ann Arbor, MI 48109

Dr. Walter Kintsch
Department of Psychology
University of Colorado
Campus Box 345
Boulder, CO 80302

Dr. David Klahr
Carnegie-Mellon University
Department of Psychology
Schenley Park
Pittsburgh, PA 15213

Mr. Al Kleider
Army Research Office
P.O. Box 12211
Research Triangle Park
North Carolina 27709-2211

Dr. Ronald Knoll
Bell Laboratories
Murray Hill, NJ 07974

Dr. Stephen Kosslyn
Harvard University
1236 William James Hall
33 Kirkland St.
Cambridge, MA 02138

Dr. Kenneth Kotovsky
Department of Psychology
Community College of
Allegheny County
800 Allegheny Avenue
Pittsburgh, PA 15233

Dr. David H. Krantz
2 Washington Square Village
Apt. # 15J
New York, NY 10012

Dr. Patrick Kyllonen
325 Aderhold
Department of Educational
Psychology
University of Georgia
Athens, GA 30602

Dr. David R. Lambert
Naval Ocean Systems Center
Code 441T
271 Catalina Boulevard
San Diego, CA 92152-6800

1987/03/09

Distribution List [Pittsburgh/Lesgold] NR 4422539

Dr. Jill Larkin
Carnegie-Mellon University
Department of Psychology
Pittsburgh, PA 15213

Dr. R. W. Lawler
ARI 6 S 10
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

Dr. Alan M. Lesgold
Learning Research and
Development Center
University of Pittsburgh
Pittsburgh, PA 15260

Dr. Alan Leshner
Deputy Division Director
Behavioral and Neural Sciences
National Science Foundation
1800 G Street
Washington, DC 20550

Dr. Jim Levin
Department of
Educational Psychology
210 Education Building
1310 South Sixth Street
Champaign, IL 61820-6990

Dr. John Levine
Learning R&D Center
University of Pittsburgh
Pittsburgh, PA 15260

Dr. Clayton Lewis
University of Colorado
Department of Computer Science
Campus Box 430
Boulder, CO 80309

Matt Lewis
Department of Psychology
Carnegie-Mellon University
Pittsburgh, PA 15213

Library,
Naval War College
Newport, RI 02940

Library,
Naval Training Systems
Center
Orlando, FL 32813

Science and Technology Division,
Library of Congress
Washington, DC 20540

Dr. Jane Malin
Mail Code SR 111
NASA Johnson Space Center
Houston, TX 77058

Dr. Sandra P. Marshall
Dept. of Psychology
San Diego State University
San Diego, CA 92182

Dr. Humberto Maturana
University of Chile
Santiago
CHILE

Dr. Richard E. Mayer
Department of Psychology
University of California
Santa Barbara, CA 93106

Dr. James McBride
Psychological Corporation
c/o Harcourt, Brace,
Javanovich Inc.
1250 West 6th Street
San Diego, CA 92101

Dr. James L. McGaugh
Center for the Neurobiology
of Learning and Memory
University of California, Irvine
Irvine, CA 92717

Dr. Gail McKoon
CAS/Psychology
Northwestern University
1859 Sheridan Road
Kresge #230
Evanston, IL 60201

Dr. Joe McLachlan
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. James S. McMichael
Navy Personnel Research
and Development Center
Code 05
San Diego, CA 92152

Dr. Barbara Means
Human Resources
Research Organization
1100 South Washington
Alexandria, VA 22314

Dr. Douglas L. Medin
Department of Psychology
University of Illinois
603 E. Daniel Street
Champaign, IL 61820

Dr. George A. Miller
Department of Psychology
Green Hall
Princeton University
Princeton, NJ 08540

Dr. Andrew R. Molnar
Scientific and Engineering
Personnel and Education
National Science Foundation
Washington, DC 20550

Dr. William Montague
NPRDC Code 13
San Diego, CA 92152-6800

Dr. Nancy Morris
Search Technology, Inc.
5550-A Peachtree Parkway
Technology Park/Summit
Norcross, GA 30092

Dr. Randy Mumaw
Program Manager
Training Research Division
HumRRO
1100 S. Washington
Alexandria, VA 22314

Dr. Allen Munro
Behavioral Technology
Laboratories - USC
1845 S. Elena Ave., 4th Floor
Redondo Beach, CA 90277

Chair, Department of
Computer Science
U.S. Naval Academy
Annapolis, MD 21402

Chair, Department of
Systems Engineering
U.S. Naval Academy
Annapolis, MD 21402

Technical Director,
Navy Health Research Center
P.O. Box 85122
San Diego, CA 92138

Dr. Allen Newell
Department of Psychology
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

Dr. Mary Jo Nissen
University of Minnesota
N218 Elliott Hall
Minneapolis, MN 55455

Dr. A. F. Norcio
Computer Science and Systems
Code: 7590
Information Technology Division
Naval Research Laboratory
Washington, DC 20375

Dr. Donald A. Norman
Institute for Cognitive
Science C-015
University of California, San Diego
La Jolla, California 92093

Deputy Technical Director,
NPRDC Code 01A
San Diego, CA 92152-6800

Director, Training Laboratory,
NPRDC (Code 05)
San Diego, CA 92152-6800

Director, Manpower and Personnel
Laboratory,
NPRDC (Code 06)
San Diego, CA 92152-6800

Director, Human Factors
& Organizational Systems Lab,
NPRDC (Code 07)
San Diego, CA 92152-6800

Fleet Support Office,
NPRDC (Code 301)
San Diego, CA 92152-6800

Library, NPRDC
Code P201L
San Diego, CA 92152-6800

Technical Director,
Navy Personnel R&D Center
San Diego, CA 92152-6800

Commanding Officer,
Naval Research Laboratory
Code 2627
Washington, DC 20390

Dr. Harold F. O'Neil, Jr.
School of Education - WPH 801
Department of Educational
Psychology & Technology
University of Southern California
Los Angeles, CA 90089-0031

Dr. Michael Oberlin
Naval Training Systems Center
Code 711
Orlando, FL 32813-7100

Dr. Stellan Ohlsson
Learning R & D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15213

Office of Naval Research,
Code 1142BI
800 N. Quincy Street
Arlington, VA 22217-5000

Office of Naval Research,
Code 1142
800 N. Quincy St.
Arlington, VA 22217-5000

Office of Naval Research,
Code 1142PS
800 N. Quincy Street
Arlington, VA 22217-5000

Office of Naval Research,
Code 1142CS
800 N. Quincy Street
Arlington, VA 22217-5000
(6 Copies)

Psychologist,
Office of Naval Research
Branch Office, London
Box 39
FPO New York, NY 09510

Special Assistant for Marine
Corps Matters,
ONR Code 00MC
800 N. Quincy St.
Arlington, VA 22217-5000

Psychologist,
Office of Naval Research
Liaison Office, Far East
APO San Francisco, CA 96503

Dr. Judith Orasanu
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Douglas Pearse
DCIEM
Box 2000
Downsview, Ontario
CANADA

Dr. James W. Pellegrino
University of California,
Santa Barbara
Department of Psychology
Santa Barbara, CA 93106

Dr. Virginia E. Pendergrass
Code 711
Naval Training Systems Center
Orlando, FL 32813-7100

Distribution List [Pittsburgh/Lesgold] NR 4422539

Dr. Nancy Pennington
University of Chicago
Graduate School of Business
1101 E. 58th St.
Chicago, IL 60637

Military Assistant for Training and
Personnel Technology,
OUSD (R & E)
Room 3D129, The Pentagon
Washington, DC 20301-3080

Dr. Steven Pinker
Department of Psychology
E10-018
M.I.T.
Cambridge, MA 02139

Dr. Martha Polson
Department of Psychology
Campus Box 346
University of Colorado
Boulder, CO 80309

Dr. Peter Polson
University of Colorado
Department of Psychology
Boulder, CO 80309

Dr. Michael I. Posner
Department of Neurology
Washington University
Medical School
St. Louis, MO 63110

Dr. Mary C. Potter
Department of Psychology
MIT (E-10-032)
Cambridge, MA 02139

Dr. Paul S. Rau
Code U-32
Naval Surface Weapons Center
White Oak Laboratory
Silver Spring, MD 20903

Dr. Lynne Reder
Department of Psychology
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

Dr. James A. Reggia
University of Maryland
School of Medicine
Department of Neurology
22 South Greene Street
Baltimore, MD 21201

Dr. Wesley Regian
AFHRL/MOD
Brooks AFB, TX 78235

Dr. Fred Reif
Physics Department
University of California
Berkeley, CA 94720

Dr. Gil Ricard
Mail Stop C04-14
Grumman Aerospace Corp.
Bethpage, NY 11714

Dr. Linda G. Roberts
Science, Education, and
Transportation Program
Office of Technology Assessment
Congress of the United States
Washington, DC 20510

Dr. Paul R. Rosenbaum
Educational Testing Service
Princeton, NJ 08541

Dr. William B. Rouse
Search Technology, Inc.
5550-A Peachtree Parkway
Technology Park/Summit
Norcross, GA 30092

Dr. David Rumelhart
Center for Human
Information Processing
Univ. of California
La Jolla, CA 92093

Dr. Walter Schneider
Learning R&D Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, PA 15260

1987/03/09

Distribution List [Pittsburgh/Lesgold] NR 4422539

Dr. Miriam Schustack
Code 51
Navy Personnel R & D Center
San Diego, CA 92152-6800

Dr. Marc Sebrechts
Department of Psychology
Wesleyan University
Middletown, CT 06475

Dr. Colleen M. Seifert
Intelligent Systems Group
Institute for
Cognitive Science (C-015)
UCSD
La Jolla, CA 92093

Dr. Ben Shneiderman
Dept. of Computer Science
University of Maryland
College Park, MD 20742

Dr. Robert S. Siegler
Carnegie-Mellon University
Department of Psychology
Schenley Park
Pittsburgh, PA 15213

Dr. Herbert A. Simon
Department of Psychology
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

LTCOL Robert Simpson
Defense Advanced Research
Projects Administration
1400 Wilson Blvd.
Arlington, VA 22209

Dr. H. Wallace Sinaiko
Manpower Research
and Advisory Services
Smithsonian Institution
801 North Pitt Street
Alexandria, VA 22314

Dr. Richard E. Snow
Department of Psychology
Stanford University
Stanford, CA 94306

Dr. Richard Sorensen
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Kathryn T. Spoehr
Brown University
Department of Psychology
Providence, RI 02912

Dr. James J. Staszewski
Research Associate
Carnegie-Mellon University
Department of Psychology
Schenley Park
Pittsburgh, PA 15213

Dr. Robert Sternberg
Department of Psychology
Yale University
Box 11A, Yale Station
New Haven, CT 06520

Dr. Kurt Steuck
AFHRL/MOD
Brooks AFB
San Antonio TX 78235

Dr. Paul J. Sticha
Senior Staff Scientist
Training Research Division
HumRRO
1100 S. Washington
Alexandria, VA 22314

Dr. John Tangney
AFOSR/NL
Bolling AFB, DC 20332

Dr. Kikumi Tatsuoka
CERL
252 Engineering Research
Laboratory
Urbana, IL 61801

Dr. Perry W. Thorndyke
FMC Corporation
Central Engineering Labs
1185 Coleman Avenue, Box 580
Santa Clara, CA 95052

Distribution List [Pittsburgh/Lesgold] NR 4422539

Dr. Sharon Tkacz
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Douglas Towne
Behavioral Technology Labs
1845 S. Elena Ave.
Redondo Beach, CA 90277

Headquarters, U. S. Marine Corps
Code MPI-20
Washington, DC 20380

Dr. William Uttal
NOSC, Hawaii Lab
Box 997
Kailua, HI 96734

Dr. Kurt Van Lehn
Department of Psychology
Carnegie-Mellon University
Schenley Park
Pittsburgh, PA 15213

Dr. Beth Warren
Bolt Beranek & Newman, Inc.
50 Moulton Street
Cambridge, MA 02138

Dr. Keith T. Wescourt
FMC Corporation
Central Engineering Labs
1185 Coleman Ave., Box 580
Santa Clara, CA 95052

Dr. Douglas Wetzel
Code 12
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Barbara White
Bolt Beranek & Newman, Inc.
10 Moulton Street
Cambridge, MA 02238

Dr. Christopher Wickens
Department of Psychology
University of Illinois
Champaign, IL 61820

Dr. Heather Wild
Naval Air Development
Center
Code 6021
Warminster, PA 18974-5000

Dr. Robert A. Wisher
U.S. Army Institute for the
Behavioral and Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333

Dr. Martin F. Wiskoff
Navy Personnel R & D Center
San Diego, CA 92152-6800

Dr. Dan Wolz
AFHRL/MOE
Brooks AFB, TX 78235

Dr. Wallace Wulfeck, III
Navy Personnel R&D Center
San Diego, CA 92152-6800

Dr. Joe Yasatuke
AFHRL/LRT
Lowry AFB, CO 80230

Dr. Joseph L. Young
Memory & Cognitive
Processes
National Science Foundation
Washington, DC 20550