DOCUMENT RESUME

| | | |
|---|---|---|
| ED 418 103 | | TM 028 197 |

| | |
|---|---|
| AUTHOR | O'Neil, Harold F., Jr.; Schacter, John |
| TITLE | Test Specifications for Problem-Solving Assessment. |
| INSTITUTION | National Center for Research on Evaluation, Standards, and Student Testing, Los Angeles, CA. |
| SPONS AGENCY | Office of Educational Research and Improvement (ED), Washington, DC. |
| REPORT NO | CSE-TR-463 |
| PUB DATE | 1997-12-00 |
| NOTE | 22p. |
| CONTRACT | R305B60002-97 |
| PUB TYPE | Numerical/Quantitative Data (110) -- Reports - Evaluative (142) |
| EDRS PRICE | MF01/PC01 Plus Postage. |
| DESCRIPTORS | *Computer Assisted Testing; Computer Software; Criteria; *Educational Assessment; *Measurement Techniques; *Problem Solving; *Scoring; Test Construction; Test Theory |
| IDENTIFIERS | High Stakes Tests; *Test Specifications |

ABSTRACT

This document reviews several theoretical frameworks of problem-solving, provides a definition of the construct, suggests ways of measuring the construct, focuses on issues for assessment, and provides specifications for the computer-based assessment of problem solving. As defined in the model of the Center for Research on Evaluation, Standards, and Student Testing (CRESST), problem solving is a cognitive process directed at achieving a goal when a solution method is not obvious to the problem solver. Issues that would drive the assessment specifications for a CRESST assessment would be: (1) selection/modification/generation of a conceptual framework; (2) what to measure; (3) assessment approaches; (4) criteria; (5) type of technology; (6) purpose of testing; (7) type of competency tested; (8) high or low stakes level of testing; (9) contexts; and (10) recommended testing time. With these points in mind, specifications for the assessment were developed, and are presented in table form. The feasibility study previously reported (J. Schacter et al., 1997), which used these specifications, illustrated that a Web-based, information-rich environment with these specifications can measure student problem solving. An appendix contains the scoring key for the developed assessment. (Contains 3 figures and 32 references.) (SLD)

CRESST

National Center for Research on Evaluation, Standards, and Student Testing



# Test Specifications for Problem-Solving Assessment

CSE Technical Report 463

Harold F. O'Neil, Jr.
University of Southern California/CRESST

John Schacter
CRESST/University of California, Los Angeles



- UCLA Center for the Study of Evaluation

in collaboration with:

- University of Colorado
- NORC, University of Chicago
- LRDC, University of Pittsburgh
- University of California, Santa Barbara
- University of Southern California
- The RAND Corporation

**Test Specifications for Problem-Solving Assessment**

CSE Technical Report 463

Harold F. O'Neil, Jr.
University of Southern California/CRESST

John Schacter
CRESST/University of California, Los Angeles

December 1997

Center for the Study of Evaluation
National Center for Research on Evaluation, Standards, and Student Testing
Graduate School of Education & Information Studies
University of California, Los Angeles
Los Angeles, CA 90095-1522
(310) 206-1532

The work reported herein was supported under the Educational Research and Development Center Program, PR/Award Number R305B60002, as administered by the Office of Educational Research and Improvement, U.S. Department of Education.

# TEST SPECIFICATIONS FOR PROBLEM SOLVING ASSESSMENT

**Harold F. O'Neil, Jr.**
**University of Southern California/CRESST**

**John Schacter**
**CRESST/University of California, Los Angeles**

In this document we review several theoretical frameworks of problem solving, provide a definition of the construct, suggest ways of measuring the construct, focus discussion on issues for assessment, and provide a specification for the computer-based assessment of problem solving.

In general, there is no widely accepted definition of problem solving that has led to a reliable and valid measure of it. The literature on problem solving is characterized by multiple theoretical frameworks—cognitive science or socio-cultural or information processing. Our theoretical frameworks are twofold: the CRESST model of learning and a specific model of problem solving. These frameworks are derived from the cognitive science literature.

The view of problem solving in the cognitive science literature not surprisingly when compared with the workplace literature (see O'Neil, 1997) tends to be less applied and more theoretical. What is useful, however, in the cognitive science literature is a set of distinctions concerning tasks, the role of domain knowledge, and a specification of the cognitive strategies/processes used in problem solving. These sets of distinctions are listed in Table 1 and are used later to focus our assessment of problem solving.

A synthesis of various theoretical literatures generated the CRESST model of learning. The CRESST model includes five families of cognitive learning, of which problem solving is one (Baker, 1995). Problem solving is a cognitive process directed at achieving a goal when a solution method is not obvious to the problem solver (Mayer & Wittrock, 1996, p. 47).

Table 1

Cognitive Science Distinctions in Problem Solving

| Issue | Distinctions | Reference |
|---|---|---|
| Definition | Problem solving is cognitive processing directed at achieving a goal when no solution method is obvious to the problem solver. (p. 47) | Mayer & Wittrock, 1996 |
| Subprocesses | Representing, planning, executing | Mayer & Wittrock, 1996 |
| | Identifying the problem, forming an internal representation, encoding, planning, strategy selection, solution monitoring. | Sternberg & Davidson, 1992 |
| Problem-solving components | Domain-specific knowledge (content understanding)<br>Metacognition (planning, self-monitoring)<br>Domain-specific problem-solving strategies<br>Motivation (self-efficacy, effort) | O'Neil & Schacter (this document) |
| Problem-solving components | Cognitive resources, the body of facts and procedures capable of being brought to bear in a particular mathematical situation.<br>Heuristics, "rules of thumb" of effective problem solving, including such aids as drawing figures, introducing suitable notation, exploiting related problems' analysis, reformulating problems, working backwards, and testing and verification procedures.<br>Control, having to do with the efficiency, resource allocating, and metacognition with which individuals utilize the knowledge at their disposal.<br>Belief systems, one's perspectives regarding the nature of a discipline and how one goes about working in it, and level of effort. | Schoenfeld, 1989 |
| Problem-solving process | (1) Representing the problem and (2) searching for a means to solve the problem. | Hayes, 1981 |
| Problem-solving process | (1) Exploration of alternative ideas; (2) extraction of relevant material; (3) simplification by constraining the problem into parts; (4) organization, attending to externally provided feedback. | Schacter et al., 1997 |
| Math problem-solving definitions | A math problem for any student is a task (a) in which the student is interested and engaged, and for which he wishes to obtain a resolution, and (b) for which the student does not have a readily accessible mathematical means by which to achieve that resolution. | Schoenfeld, 1989 |

Table 1 (continued)

| Issue | Distinctions | Reference |
|---|---|---|
| Problem solving in electronics | Lesgold and Lajoie (1989) list the following components of problem solving in electronics: an analysis of electronics troubleshooting expertise, including (a) conceptual underpinnings, principles, and models; (b) methods, strategy, and tactics; and (c) social, personal and practical knowledge, battle experience, war stories, situational tuning. | Lesgold & Lajoie, 1989 |
| Content-process space | Content knowledge (rich, lean) vs. process skills (open, constrained) | Glaser & Baxter, 1997 |
| Application areas | Puzzles, math and science, troubleshooting | |
| European definitions of problem solving | There are two principal approaches to studying expertise in complex problem solving. One approach was largely initiated by Broadbent (1977) in England, the other by Dörner (1987) in Germany. The key difference between the two traditions is that in the former, one is able to specify a precise rule via a mathematical formula that would optimize problem solving, whereas in the latter, the problems are so complex that it is questionable whether we could even devise any mathematical or even computer simulation that would clearly optimize performance. What the two approaches have in common, however, is that one can give the problems to anyone of at least roughly average intelligence and get them to solve the problems. (p. 301) | Sternberg, 1995 |
| Mental simulation | A problem-solving strategy is use of mental simulation.<br><br>Four primary functions are served by mental simulations: generate a course of action, inspect and evaluate a course of action, explain a phenomenon, and discover and explore models of a phenomenon. (p. 337) | Klein & Crandall, 1995 |
| Problem-solving definition | Problem solving is defined broadly here, as higher level cognitive activity, either novel or routine, that requires previous learning of various types and that may result in new learning. . . . Key features of problem solving are (1) the task requires a solution or sets a performance goal, but the solution process may not be a defined procedure, and there may be a variety of correct solutions; (2) some degree of search takes place in the performer's thinking process; (3) the performer uses previously learned rules, verbal information, and cognitive strategies to reach a solution or achieve the goal; and (4) in the process of | Gagné & Medsker, 1996 |

Table 1 (continued)

| Issue | Distinctions | Reference |
|---|---|---|
| | solving the problem, the performer may learn a higher order rule or cognitive strategy that will help solve similar problems in the future. (pp. 124-125) | |
| Problem-solving definition (rich content knowledge) | In brief, competent students (a) provide coherent explanations based on underlying principles rather than descriptions of superficial features or single statements of fact; (b) generate a plan for solution that is guided by an adequate representation of the problem situation and possible procedures and outcomes; (c) implement solution strategies that reflect relevant goals and subgoals; and (d) monitor their actions and flexibly adjust their approach based on performance feedback. (pp. 2-3) | Baxter, Elder, & Glaser, 1996 |
| Domain-independent strategies | Mental simulations<br><br>Use of analogy<br>Use of multiple representations | Klein & Crandall, 1995;<br>Sparks, 1996;<br>Anderson, 1990 |
| General problem-solving strategies | Means-ends analysis; Working backward; Simplification; Generalization and specialization; Trial and error; Rules; Brainstorming; Contradiction; Restate the problem; Analogies and metaphors (pp. 163-167) | Crowl, Kaminsky, & Podell, 1997 |
| Domain-dependent strategies (troubleshooting)<br><br>Identifying the problem | The sequencing HYDRIVE differentiates between several forms of space splitting. There is power system elimination, which removes power system sources from the problem area (as in checking hydraulic pressure gauges or circuit breakers); there is active path splitting, which activates different combinations of components to achieve a particular system function (as in operating the rudders through the control stick and through the rudder pedals); and there is power path splitting, which either eliminates series of edges having the same power type or locates the failure to a particular power type (as in using electrical backup to replace mechanical function). (Adapted from p. 19) | Gitomer, Steinberg, & Mislevy, 1994 |
| Domain-dependent strategies (troubleshooting)<br><br>Fixing the problem | Removing and replacing a component and observing whether the change results in a fix to the system. A remove and replace strategy is expensive both in terms of time and equipment, and is recommended only when there is a high degree of certainty that the replaced component is faulty. A serial elimination strategy refers to actions that only provide information about one component at a time. (Adapted from p. 19) | Gitomer, Steinberg, & Mislevy, 1994 |

The CRESST model of problem solving (Figure 1) is adapted from the problem-solving models of Baxter, Elder, and Glaser (1996), Glaser, Raghavan, and Baxter (1992), and Mayer and Wittrock (1996), Sugrue (1995). It includes four elements: (a) content understanding, (b) problem-solving strategies (i.e., either domain-dependent or domain-independent), (c) metacognition, and (d) motivation. An example of a domain-dependent strategy would be a problem-solving information-seeking strategy to help a learner find information to better solve the problem. For instance, "use Boolean operators" in constructing a search argument or browsing using hypertext to survey information spaces. An example of a domain-independent strategy would be the use of analogies (e.g., "a tire pump is like a syringe" or "electrons around an atom is like the planets around the sun"). In general, to be a successful problem solver, one must know something (content knowledge), possess intellectual tricks (problem-solving strategies), be able to plan and monitor one's progress towards solving the problem (metacognition), and be motivated to perform.

Our R&D in problem solving using the above definition focused on the feasibility of using Web-based technology for the assessment of problem solving. Currently, the ideal assessment of problem solving is based on think-aloud

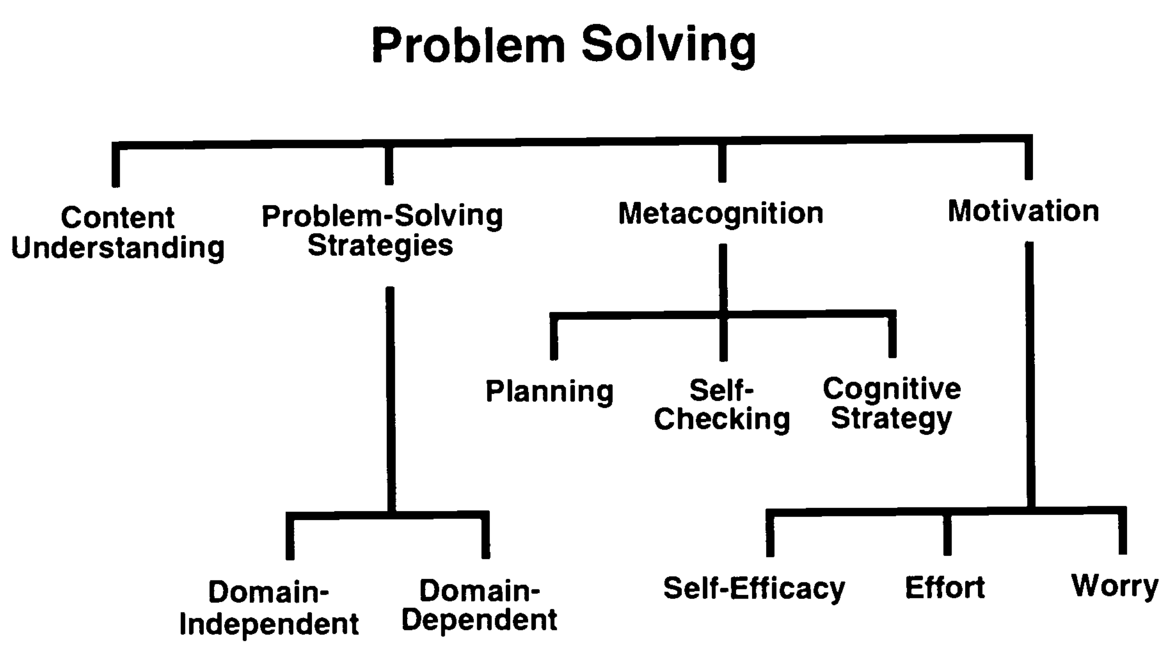


*Figure 1.* Model of problem solving.

protocols (Ericsson & Simon, 1993; Voss & Post, 1988; Voss, Tyler, & Yengo, 1983) or performance assessments that require extensive human rater scoring. However, such assessments are expensive and time consuming and result in delayed (up to months or years) reporting to students, teachers, and parents. In our work, we have taken a different approach to measuring these problem-solving processes. Instead of conducting interviews, naturalistic observation, or think-aloud protocols while students problem solve, we rely on computer (i.e., server) access logs to automatically record most of the process behaviors students engage in while on the computer. Our approach is to computerize the administration, scoring, and reporting of problem-solving measures, thus facilitating timely reporting and potentially increasing reliability and validity (Herl, O'Neil, Chung, & Dennis, 1997; Martinez, 1993; O'Neil, Chung, & Brown, 1997; Schacter et al., 1997).

In our empirical work, content knowledge is assessed by concept maps. Domain-specific problem-solving strategies are measured through an information-seeking task that analyzes both search behavior and how newly found information is utilized. Domain-independent problem solving is assessed using metacognition and motivation survey instruments. The basic design involves having the student (a) create a concept map, (b) receive feedback on it, then, (c) using a simulated Web site, search for information to improve it, and (d) construct a final concept map. The final concept map serves as the outcome content understanding measure. Finally, metacognition and motivation are assessed by paper-and-pencil survey instruments.

We believe that content understanding and problem-solving strategies are best assessed domain-specifically whereas metacognition and motivation are best assessed as domain-independent constructs. We have created measures of trait metacognition and motivation that can be administered in 4 to 6 minutes and would be included in the assessment time for problem solving. The trait scales are found in the Appendix. Using constructs from state-trait anxiety theory (Spielberger, 1975) as an analogy, we have formulated a set of self-report, domain-independent trait and state measures of metacognition and motivation. We find the state versus trait distinction useful for both cognitive and affective measurement. Thus, we have generalized the key constructs from an affective domain (e.g., state and trait anxiety) to a cognitive domain (e.g., state and trait metacognition).

In our technical approach, we have created both process and outcome measures of problem-solving strategies. As was mentioned above, we assessed problem-solving search behavior by giving students the task of improving their existing representational maps (i.e., concept maps). Students were required to *search* for information that would enhance their existing concept maps and, at the same time, *justify* any changes they made to their maps. By monitoring students' online search, we can determine whether (a) students search well enough to access rich information sources within the "weak" concept areas in their maps (as determined from feedback on their initial concept maps); (b) students understand enough to define the problem through their search behavior; and (c) students can sufficiently justify the changes they make to their concept maps. This approach allows us to examine both the searching process (through analysis of student search strategies) and the final integrated product (the updated concept map).

## Specification Issues in Assessment of Problem Solving

The purpose of this section is to suggest various approaches for the assessment of problem-solving skills. These issues would drive the assessment specifications. In general, the issues in assessing these skills are conceptualized as (a) selection/modification/generation of a conceptual framework; (b) what to measure (e.g., cognitive processes, tasks, or characteristics of jobs and the setting or context); (c) assessment approaches (e.g., multiple-choice items, concept map, essay, or performances); (d) criteria (e.g., validity, fairness, cost); (e) type of technology (e.g., paper-and-pencil, computer); (f) purpose of testing (e.g., diagnostic, selection, accountability, credentialling); (g) type of competency tested (i.e., individual vs. team); (h) level of stakes (high vs. low); (i) contexts (e.g., school, work, home, community); (j) recommended testing time.

The instantiation of these general assessment issues for CRESST work can be seen in Table 2. For example, we use an approach to measure problem solving in which domain knowledge, problem-solving strategies, metacognition, and motivation are measured as facets of problem solving. Further, the assessment of problem solving will be considered to be for accountability purposes of a low-stakes nature for the individual involved and will assess only individual problem solving.

Table 2

Specification Issues for Assessment of Problem Solving

| Issue | Problem-solving approach |
|---|---|
| Modify a workforce competency framework with cognitive science definition of problem solving | SCANS (U.S. Department of Labor, 1992a, 1992b); Mayer and Wittrock (1996) |
| What to measure | Domain knowledge, problem-solving strategies, metacognition, motivation |
| Assessment approach | Task implemented in a problem-solving block |
| Assessment formats | Concept map, simulated Web space, survey instruments |
| Criteria | Reliability, validity, fairness, and transfer |
| Type of task(s) | Search |
| Type of technology | Computer-based |
| Purpose of testing | Program evaluation |
| Type of competency | Individual and not team |
| Level of stakes | Low for the individual |
| Contexts | Simulate problem-solving tasks at school |
| Testing time | 70 minutes |

## Assessment Specifications

The resulting detailed assessment specifications are shown in Figure 2 and Figure 3. These are the specifications that we used in our CAETI work.[1] They now serve as documentation of the existing assessment software. Figure 2 contains the specifications for the concept map. Figure 3 contains the specifications for the domain-specific, problem-solving search strategy.

[1] The format but not the content of these specifications was adopted from Ruiz-Primo and Shavelson (1995).

| General specification attributes | Specific example |
|---|---|
| Task demands | Construct a concept map with a given set of concepts (terms) and links. |
| Task constraints | Student is provided with a fixed set of concepts and links.<br><br>Concepts are important ideas; range from 10-18. The final set of 18 concepts included ATMOSPHERE, BACTERIA, CLIMATE, CARBON DIOXIDE, DECOMPOSITION, EVAPORATION, FOOD CHAIN, OCEANS, PRODUCERS, CONSUMERS, RESPIRATION, SUNLIGHT, PLANTS, PHOTOSYNTHESIS, WASTE, WATER CYCLE, OXYGEN, and GREENHOUSE GASES.<br><br>Links are important relationships; range from 6-8. Links used were CAUSES, INFLUENCES, PART OF, PRODUCES, USED FOR, USES.<br><br>Duration: 20 minutes |
| Scoring | Focus is on the following components:<br>• propositions and agreement with criterion map(s)<br>• criterion maps: compare student's map with experts' map(s)<br>Criterion maps were obtained from 4 expert teachers. |
| Type of learning | Content understanding |
| Content domain | Environmental science |
| Outcome measures | Herl metric (Herl, Baker, & Niemi, 1996) comprised of:<br>Semantic content score, organizational structure score, number of terms used, number of links used, comparison to 4 expert maps |
| Prerequisite skills | None |
| Training | 8 minutes of instruction |
| Cognitive processes (domain-dependent) | Representation of content knowledge |
| Cognitive processes (domain-independent) | Metacognition |
| Affective processes (domain-independent) | Effort and self-efficacy |
| Student response | The student response will be on a computer. |
| Hardware and software components | Macintosh Power PC or Pentium PC; Mac OS 7.5.3 or Windows 95 or NT; 16 megabytes of RAM or more; 20 megabytes of hard disk space; Netscape 3.0 (Java and Javascript capable browser); Open Transport 1.1 or TCP/IP; Ethernet; programming language (Java, HyperCard); |
| Language | English |

*Figure 2.* Domain specifications for individually constructed concept map.

| General specification attributes | Specific example |
|---|---|
| Task demands | Use Netscape and a simulated Web space to find relevant information via searching to improve existing concept map (see Figure 2). Bookmark relevant information to justify new concepts and links. |
| Task constraints | Student are provided with a simulated Web environment:<br>• Our Web server software is AOL. Environment consists of 130 Web documents and over 400 images and diagrams about environmental science and other topic areas.<br>• The database is made up of both relevant and irrelevant information about environmental science that varies in degree of specificity, reading level, scope and comprehensiveness. Ninety percent of the information on the simulated Web was downloaded from the WWW using Web Whacker 1.0 software. The other 10% of the information was adapted from science textbooks and magazines.<br>• Four mechanisms to search a simulated Internet information space: (a) search engine (b) hierarchical subject listing (directory), (c) hypertext, and (d) a glossary of environmental science terms.<br>• A mechanism to bookmark Web pages encountered during information seeking and send those bookmarks to nodes in the map.<br>Duration: 35 minutes |
| Scoring | |
| Scoring problem-solving outcomes | Outcome of map score after searching; premap to postmap improvement |
| Relevance-rated database | Two experts rate all information in database on 4-point scale for its relevance to each concept in the task of constructing a concept map. |
| Extraction | Student bookmarks for each concept are compared to expert relevance ratings for each concept with scope and depth scores. |
| Simplification | Students searched on at least 20% of the concepts that they needed to work on. |
| Exploration | • Frequency counts of Boolean searching, synonym substitution, multiple terms in searching string<br>• Frequency counts of total search moves, browsing moves, scan and select moves, analytic queries made |
| Organization | Use of map feedback |

*Figure 3.* Specifications for domain-specific problem-solving search strategy.

| | |
|---|---|
| Type of learning | Problem-solving strategies |
| Content domain | Environmental science |
| Outcome measures | (a) Herl metric (Herl, Baker, & Niemi, 1996); (b) student bookmarks |
| Prerequisite skills | Computer concept mapping |
| Training | 12 minutes of instruction |
| Cognitive processes (domain-dependent) | Problem-solving processes (exploration, extraction, simplification, and organization) |
| Cognitive processes (domain-independent) | Metacognition |
| Motivational processes (domain-independent) | Effort, self-efficacy |
| Student response | The student response will be on a computer. |
| Hardware and software components | Macintosh Power PC or Pentium PC; Mac OS 7.5.3 or Windows 95 or NT; 16 megabytes of RAM or more; 20 megabytes of hard disk space; Netscape 3.0 (Java and Javascript capable browser; open transport 1.1 or TCP/IP; Ethernet; programming language (Java, HyperCard) |
| Language | English |

*Figure 3* (continued).

## Feasibility of a Web-Based Assessment of Problem Solving

In the feasibility study reported at the 1997 annual meeting of the American Educational Research Association (Schacter et al., 1997), which used the above specifications, we presented our first attempt at automated data collection and scoring of students' complex problem-solving processes and performance in Web-based, information-rich environments We are still developing the real-time scoring and reporting software. In general we found the computer-based problem-solving assessment to be feasible.

The feasibility study illustrated that a Web-based, information-rich environment can measure student problem solving. We argue that this finding is promising because it provides evidence that we can measure problem solving when working on realistic problems in realistic contexts that demand the activation of multiple cognitive processes. Such measurement is facilitated by the use of specifications.

## References


Anderson, J. R. (1990). *Cognitive psychology and its implications* (3rd ed.). New York: W. H. Freeman.

Baker, E. L. (1995, September). *Finding our way.* Presentation at the annual conference of the Center for Research on Evaluation, Standards, and Student Testing, University of California, Los Angeles.

Baker, E. L., & Schacter, J. (1996). Expert benchmarks for student academic performance: The case for gifted children. *Gifted Child Quarterly, 40,* 61-65.

Baxter, G. P., Elder, A. D., & Glaser, R. (1996). *Assessment and instruction in the science classroom* (CSE Tech. Rep. No. 418). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

Crowl, T. K., Kaminsky, S., & Podell, D. M. (1997). *Educational psychology. Windows on teaching* (Annotated instructor's edition). Madison, WI: Brown & Benchmark Publishers.

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (rev. ed.). Cambridge, MA: MIT Press.

Gagné, R. M., & Medsker, K. L. (1996). *The conditions of learning. Training applications.* Fort Worth, TX: Harcourt Brace College Publishers.

Gitomer, D. H., Steinberg, L. S., & Mislevy, R. J. (1994, April). *Diagnostic assessment of troubleshooting skill in an intelligent tutoring system* (RR-9421-ONR). Princeton, NJ: Educational Testing Service.

Glaser, R., & Baxter, G. (1997, February). *Improving the theory and practice of performance-based assessment.* Paper presented to the Board of Testing and Assessment, National Research Council/National Academy of Sciences Science Education Standards Conference: "The Assessment of Science Meets the Science of Assessments," Orlando, FL.

Glaser, R., Raghavan, K., & Baxter, G. (1992). *Design characteristics of science performance assessments* (CSE Tech. Rep. No. 349). Los Angeles: University of California, Center for Research on Evaluation, Standards, and Student Testing.

Hayes, J. R. (1981). *The complete problem solver.* Philadelphia, PA: The Franklin Institute Press.

Herl, H. E., Baker, E. L., & Niemi, D. (1996). Construct validation of an approach to modeling cognitive structure of U.S. history knowledge. *Journal of Educational Research, 89,* 206-218.

Herl, H. E., O'Neil, H. F., Jr., Chung, G. K. W. K., & Dennis, R. A. (1997, March). Feasibility of an on-line concept mapping construction and scoring system. In H. F. O'Neil, Jr. (Chair), *An integrated simulation approach to assessment.* Symposium presented at the annual meeting of the American Educational Research Association, Chicago.

Klein, G., & Crandall, B. W. (1995). The role of mental simulation in problem solving and decision making. In P. Hancock, J. Flach, J. Caird, & K. J. Vicente (Eds.), *Local applications of the ecological approach to human machine systems* (pp. 324-358). Hillsdale, NJ: Lawrence Erlbaum Associates.

Lesgold, A., & Lajoie, S. (1989). *Complex problem solving in electronics.* Manuscript in preparation, University of Pittsburgh, Learning Research and Development Center.

Martinez, M. E. (1993). Problem-solving correlates of new assessment forms in architecture. *Applied Measurement in Education, 6,* 167-180.

Mayer, R. E., & Wittrock, M. C. (1996). Problem-solving transfer. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 47-62). New York: Simon & Schuster Macmillan.

O'Neil, H. F., Jr. (Ed.). (1997). *Workforce readiness: Competencies and assessment.* Mahwah, NJ: Lawrence Erlbaum Associates.

O'Neil, H. F., Jr., Chung, G., & Brown, R. (1997). Use of networked simulations as a context to measure team competencies. In H. F. O'Neil, Jr. (Ed.), *Workforce readiness: Competencies and assessment* (pp. 411-452). Mahwah, NJ: Lawrence Erlbaum Associates.

Ruiz-Primo, M. A., & Shavelson, R. J. (1995, April) *Concept maps as potential alternative assessments in science.* Paper presented at the annual meeting of the American Educational Research Association, San Francisco.

Schacter, J., Herl, H. E., Chung, G. K. W. K., O'Neil, H. F., Jr., Dennis, R. A., & Lee, J. L. (1997, March). Feasibility of a Web-based assessment of problem solving. In H. F. O'Neil, Jr. (Chair), *An integrated simulation approach to assessment.* Symposium presented at the annual meeting of the American Educational Research Association, Chicago.

Schoenfeld, A. H. (1989). Teaching mathematical thinking and problem solving. In L. B. Resnick & L. E. Klopfer (Eds.), *Toward the thinking curriculum:*

*Current cognitive research* (pp. 83-103). Alexandria, VA: Association for Supervision and Curriculum Development.

Sparks, P. (1996). *Building mental models for problem solving with illustrations and analogies.* Unpublished doctoral dissertation, University of Southern California, Los Angeles.

Spielberger, C. D. (1975). Anxiety: State-trait process. In C. D. Spielberger & I. G. Sarason (Eds.), *Stress and anxiety* (Vol. 1, pp. 115-143). Washington, DC: Hemisphere.

Sternberg, R. J. (1995). Expertise in complex problem solving: A comparison of alternative conceptions. In P. A. Frensch & J. Funke (Eds.), *Complex problem solving. The European perspective* (pp. 295-321). Hillsdale, NJ: Lawrence Erlbaum Associates

Sternberg, R. J., & Davidson, J. E. (1992). Problem solving. In M. Alkin (Ed.), *Encyclopedia of educational research* (6th ed., pp. 1037-1045). New York: Macmillan.

Sugrue, B. (1995). A theory-based framework for assessing domain-specific problem-solving ability. *Educational Measurement: Issues and Practice, 14*(3), 29-36.

U.S. Department of Labor. (1992a). *Learning a living: A blueprint for high performance.* Washington, DC: U.S. Department of Labor, The Secretary's Commission on Achieving Necessary Skills (SCANS).

U.S. Department of Labor. (1992b). *Skills and tasks for jobs: A SCANS report for America 2000.* Washington, DC: U.S. Department of Labor, The Secretary's Commission on Achieving Necessary Skills (SCANS).

Voss, J. F., & Post, T. A. (1988). On the solving of ill-structured problems. In M. T. Chi, R. Glaser, & M. J. Farr (Eds.), *The nature of expertise* (pp. 261-287). Hillsdale, NJ: Lawrence Erlbaum Associates.

Voss, J. F., Tyler, S. W., & Yengo, L. A. (1983). Individual differences in the solving of social science problems. In R. F. Dillon & R. R. Schmeck (Eds.), *Individual differences in cognition* (pp. 204-232). New York: Academic Press.

## APPENDIX

## CAETI SCORING KEY (25 August 1996)

## TRAIT THINKING QUESTIONNAIRE

| Scales | Items |
|---|---|
| Planning | 1, 8, 11, 19, 25, 28, 33, 36 |
| Cognitive strategy | 2, 7, 12, 17, 21, 29, 34, 38 |
| Self-checking | 3, 6, 13, 16, 22, 30, 35, 39 |
| Effort | 4, 9, 14, 18, 23, 26, 31, 40 |
| Self-efficacy | 5, 10, 15, 20, 24, 27, 32, 37 |

PLANNING

1. I determine how to solve a task before I begin.
8. I carefully plan my course of action.
11. I try to understand tasks before I attempt to solve them.
19. I try to understand the goal of a task before I attempt to answer.
25. I figure out my goals and what I need to do to accomplish them.
28. I imagine the parts of a task I have to complete.
33. I make sure I understand just what has to be done and how to do it.
36. I try to determine what the task requires.

COGNITIVE STRATEGY

2. To understand a task, I draw a graph if at all possible.
7. I think through the steps of a plan in my mind.
12. While solving a task, I try more than one way to do it.
17. I think through the meaning of tasks before I begin to answer them.
21. I select and organize relevant information to solve a task.
29. I spend more time trying to understand difficult tasks.
34. I attempt to discover the main ideas in a task.
38. I ask myself how this task relates to what I already know.

SELF-CHECKING

3. I check how well I am doing when I solve a task.
6. I ask myself questions to stay on track as I do a task.
13. I check my work while I am doing it.
16. I almost always know how much of a task I have to complete.
22. I judge the correctness of my work.
30. I correct my errors.
35. I check my accuracy as I progress through a task.
39. I ask myself, how well am I doing, as I proceed through tasks.

EFFORT

4. I work hard to do well even if I don't like a task.
9. I put forth my best effort on tasks.
14. I work as hard as possible on tasks.
18. I am willing to do extra work on tasks to improve my knowledge.
23. I concentrate as hard as I can when doing a task.
26. I work hard on a task even if it does not count.
31. A task is useful to check my knowledge.
40. Practice makes perfect.

SELF-EFFICACY

5. I believe I will receive an excellent grade in this course.
10. I'm certain I can understand the most difficult material presented in the readings for this course.
15. I'm confident I can understand the basic concepts taught in this course.
20. I'm confident I can understand the most complex material presented by the teacher in this course.
24. I'm confident I can do an excellent job on the assignments and tests in this course.
27. I expect to do well in this course.
32. I'm certain I can master the skills being taught in this course.
37. Considering the difficulty of this course, the teacher, and my skills, I think I will do well in this course.

**Trait Thinking Questionnaire**

Name (please print): ______________________________

Teacher: ____________________ Date: ____________________

Directions: A number of statements which people have used to describe themselves are given below. Read each statement and indicate how you generally think or feel on learning tasks by marking your answer sheet. There are no right or wrong answers. Do not spend too much time on any one statement. Remember, give the answer that seems to describe how you generally think or feel.

| | | Almost never | Sometimes | Often | Almost always |
|---|---|---|---|---|---|
| 1. | I determine how to solve a task before I begin. | 1 | 2 | 3 | 4 |
| 2. | To understand a task, I draw a graph if at all possible. | 1 | 2 | 3 | 4 |
| 3. | I check how well I am doing when I solve a task. | 1 | 2 | 3 | 4 |
| 4. | I work hard to do well even if I don't like a task. | 1 | 2 | 3 | 4 |
| 5. | I believe I will receive an excellent grade in this course. | 1 | 2 | 3 | 4 |
| 6. | I ask myself questions to stay on track as I do a task. | 1 | 2 | 3 | 4 |
| 7. | I think through the steps of a plan in my mind. | 1 | 2 | 3 | 4 |
| 8. | I carefully plan my course of action. | 1 | 2 | 3 | 4 |
| 9. | I put forth my best effort on tasks. | 1 | 2 | 3 | 4 |
| 10. | I'm certain I can understand the most difficult material presented in the readings for this course. | 1 | 2 | 3 | 4 |
| 11. | I try to understand tasks before I attempt to solve them. | 1 | 2 | 3 | 4 |
| 12. | While solving a task, I try more than one way to do it. | 1 | 2 | 3 | 4 |
| 13. | I check my work while I am doing it. | 1 | 2 | 3 | 4 |
| 14. | I work as hard as possible on tasks | 1 | 2 | 3 | 4 |
| 15. | I'm confident I can understand the basic concepts taught in this course. | 1 | 2 | 3 | 4 |
| 16. | I almost always know how much of a task I have to complete. | 1 | 2 | 3 | 4 |
| 17. | I think through the meaning of tasks before I begin to answer them. | 1 | 2 | 3 | 4 |

| | | Almost never | Sometimes | Often | Almost always |
|---|---|---|---|---|---|
| 18. | I am willing to do extra work on tasks to improve my knowledge | 1 | 2 | 3 | 4 |
| 19. | I try to understand the goal of a task before I attempt to answer. | 1 | 2 | 3 | 4 |
| 20. | I'm confident I can understand the most complex material presented by the teacher in this course. | 1 | 2 | 3 | 4 |
| 21. | I select and organize relevant information to solve a task. | 1 | 2 | 3 | 4 |
| 22. | I judge the correctness of my work. | 1 | 2 | 3 | 4 |
| 23. | I concentrate as hard as I can when doing a task. | 1 | 2 | 3 | 4 |
| 24. | I'm confident I can do an excellent job on the assignments and tests in this course. | 1 | 2 | 3 | 4 |
| 25. | I figure out my goals and what I need to do to accomplish them. | 1 | 2 | 3 | 4 |
| 26. | I work hard on a task even if it does not count. | 1 | 2 | 3 | 4 |
| 27. | I expect to do well in this course. | 1 | 2 | 3 | 4 |
| 28. | I imagine the parts of a task I have to complete. | 1 | 2 | 3 | 4 |
| 29. | I spend more time trying to understand difficult tasks. | 1 | 2 | 3 | 4 |
| 30. | I correct my errors. | 1 | 2 | 3 | 4 |
| 31. | A task is useful to check my knowledge. | 1 | 2 | 3 | 4 |
| 32. | I'm certain I can master the skills being taught in this course. | 1 | 2 | 3 | 4 |
| 33. | I make sure I understand just what has to be done and how to do it. | 1 | 2 | 3 | 4 |
| 34. | I attempt to discover the main ideas in a task. | 1 | 2 | 3 | 4 |
| 35. | I check my accuracy as I progress through a task. | 1 | 2 | 3 | 4 |
| 36. | I try to determine what the task requires. | 1 | 2 | 3 | 4 |
| 37. | Considering the difficulty of this course, the teacher, and my skills, I think I will do well in this course. | 1 | 2 | 3 | 4 |
| 38. | I ask myself how this task relates to what I already know. | 1 | 2 | 3 | 4 |
| 39. | I ask myself, how well am I doing, as I proceed through tasks. | 1 | 2 | 3 | 4 |
| 40. | Practice makes perfect. | 1 | 2 | 3 | 4 |

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
Educational Resources Information Center (ERIC)

ERIC®

# REPRODUCTION RELEASE

(Specific Document)

## I. DOCUMENT IDENTIFICATION:

| | |
|---|---|
| **Title:** Test Specifications for Problem-Solving Assessment | |
| **Author(s):** Harold F. O'Neil, Jr. and John Schacter | |
| **Corporate Source:** UCLA Center for the Study of Evaluation | **Publication Date:** December 1997 |

## II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following two options and sign at the bottom of the page.

☒ Check here
**For Level 1 Release:**
Permitting reproduction in microfiche (4" x 6" film) or other ERIC archival media (e.g., electronic or optical) *and* paper copy.

The sample sticker shown below will be affixed to all Level 1 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

**Level 1**

The sample sticker shown below will be affixed to all Level 2 documents

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

**Level 2**

☐ Check here
**For Level 2 Release:**
Permitting reproduction in microfiche (4" x 6" film) or other ERIC archival media (e.g., electronic or optical), but *not* in paper copy.

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

*"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."*

**Sign here→ please**

| | | |
|---|---|---|
| Signature: Kim Hurst | Printed Name/Position/Title: Kim Hurst/Admin Assistant | |
| Organization/Address: | Telephone: 310-206-1532 | FAX: 310-825-3883 |
| CRESST The National Center for Research on Evaluation, Standards, and Student Testing & The UCLA Center for the Study of Evaluation 301 GSE&IS/Mailbox 951522 Los Angeles, California 90095-1522 | E-Mail Address: kim@cse.ucla.edu | Date: 2/25/98 |