DOCUMENT RESUME

ED 367 885 CE 065 997

AUTHOR Whichard, Judith A.; Cobb, R. Brian
TITLE Assessment Alternatives for Students in Vocational Education.
PUB DATE [93]
NOTE 46p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Cognitive Processes; Curriculum Based Assessment; *Evaluation Methods; Informal Assessment; Learning Processes; *Learning Theories; Literature Reviews; Portfolios (Background Materials); Postsecondary Education; Secondary Education; *Student Evaluation; *Thinking Skills; *Vocational Education
IDENTIFIERS Alternative Assessment

ABSTRACT

Emerging theories suggest that learning is contingent on cognitive development, which is, in turn, influenced by conative and ecological factors. These three domains--cognitive, conative, and ecological--all make important contributions to understanding student achievement. Their collective influences should be understood when formulating an opinion of student achievement. Thus, assessment becomes holistic, considering not only achievement outcomes, but also the cognitive processes through which the learning occurs, the learning context, and the learner's attitudinal preferences. A comprehensive assessment approach includes a number of assumptions for developing and interpreting measures. These assumptions suggest a number of directions for the focus of assessment. Four assessment alternatives--portfolio, exhibition, dynamic, and curriculum based--demonstrate characteristics that support the assumptions. Sensitive to the instructional needs of teachers, these alternatives also readily adapt to instructional settings, while addressing the three domains of learning. Moreover, they emphasize the measurement of student performances directly tied to instructional goals, so they could guide classroom curricula as well as measure short-term and long-term changes in students' knowledge. (Appendixes include 77 references and 2 tables that compare product and process assessment strategies.) (YLB)

Assessment Alternatives for Students in Vocational Education

Judith A. Whichard and R. Brian Cobb

School of Occupational and Educational Studies

Colorado State University

Running Head: ASSESSMENT ALTERNATIVES IN VOCATIONAL EDUCATION

Abstract

The research reported in this paper analyzed the literature to identify assessment strategies that would address the achievement of higher-order thinking skills. Initially, to understand the nexus between student assessment and learning, cognitive learning theory, conative structure development and ecological factors were reviewed. Next, a list of assumptions that addressed these cognitive, conative and ecological considerations was articulated, guiding the analysis of literature and identification of appropriate assessment strategies. Four assessment alternatives -- portfolio, exhibition, dynamic, and curriculum-based -- demonstrated characteristics supporting the assumptions. Sensitive to the instructional needs of teachers, these alternatives also readily adapted to instructional settings, while addressing the three domains of learning (i.e., cognitive, conative, and ecological). Moreover, they emphasized the measurement of student performances directly tied to instructional goals, thus they could guide classroom curricula, as well as measure short-term and long-term changes in students' knowledge.

Assessment Alternatives

for Students in Vocational Education

As America's share of domestic and international trade markets continues to erode, the need to combat educational deficiencies has intensified. Recognizing that inadequately prepared workers have significantly contributed to our diminishing economic competitiveness (Bailey, 1991), legislators have proposed policies for educational reformations of unprecedented comprehensiveness in an attempt to raise educational standards and thus forestall further economic deterioration. States have responded by increasing graduation requirements, and infusing more academic curricula into high school programs. Business and industry have been enlisted as educational allies to help articulate instructional competencies that are more responsive to workplace challenges (SCANS, 1991).

While the application of such policy changes has reconfigured the amount of academic content students receive, of greater consequence is its impact on the reconfiguration of educational literacy. The definition of literacy that sufficed for earlier generations and a different economy has been replaced by a host of different literacies such as computer, scientific, mathematic, civic and cultural. These literacies call for students to analyze, think critically, evaluate, synthesize information, communicate effectively, solve problems, know how to learn, and in general, learn far more effectively (Bailey, 1990; Brown, 1989).

As the emphasis to include these literacies in instruction has grown, so too has the need to reexamine the purposes and functions of educational institutions (Bailey, 1991; Resnick & Resnick, 1985). Expanded missions that are responsive to the structural changes in the nature of work appear to be paramount in balancing the needs of the economy with the talents of school graduates (Resnick & Resnick, 1985). Consequently, all educational fora have been given the task not only to analyze curriculum for its recency and germaneness, but also for its implied methods of delivery (Bailey, 1990; Darling-Hammond, 1990).

A partial corollary to the general education reformation efforts are the Carl D. Perkins Vocational and Applied Technology Education Act Amendments of 1990. Invoking major changes (cf., Rosenstock, 1991; Wirt, 1991), the Amendments have broadened both the purpose and the scope of vocational education, chiefly, by modifying its traditional mission. Under the new legislation, vocational education has been encouraged to de-emphasize its preparation of some students for full-time employment and to emphasize its role as an important instructional modality for all students for acquiring higher-order reasoning and problem-solving skills (Gray, 1991; Rosenstock, 1991).

Although this modified mission is not the only change, it has the greatest potential to contribute to the larger educational reformation effort by blurring the distinction between academic and vocational education. For the first time vocational education can play a pivotal role in planning curricula that engender higher student achievement through the integration of academic and vocational content (Gray, 1991; Rosenstock, 1991).

To assume this leadership role, however, requires an understanding of the factors that set and clarify intellectual standards. Educational researchers (Nickerson, 1989; Resnick & Resnick, 1985; Snow, 1989; Wiggins, 1989) cite two key considerations for establishing these standards--curriculum and student assessment. Albeit perceived as discrete components, in practice they are reflections of one another, continually defining and redefining the things that are taught, how they are taught, how they are measured, and how much learning students have achieved (Resnick & Resnick, 1985; Stiggins, 1991a; Wiggins, 1991a). In short, inextricably interwoven, curriculum and student assessment determine the directions of education and ultimately provide a measure of the effectiveness of the whole educational enterprise (Nickerson, 1989; Perrone, 1991).

Because of its acknowledged criticality to the understanding of instructional achievement, student assessment is the primary emphasis of this paper. The argument presented in this paper is developed in four parts. The first considers the modified direction of vocational education and its implications for student assessment. The second part, by presenting cognitive learning theory, conative

structure development and ecological considerations, develops the nexus between student assessment and learning. The third section reviews extant student assessment alternatives and their potential relevance for vocational educators. The paper concludes with a discussion of each assessment alternative's relative abilities to address critical learning domains.

The Direction of Vocational Education

Continuing to emphasize the importance of serving special populations, the new vocational education legislation also charts a new direction by altering its historical purpose. Responsive to the challenges of the changing workforce and workplace, the 1990 Amendments direct vocational education toward greater compliance with the larger school restructuring initiative by requiring the integration of academic and vocational education at both secondary and postsecondary levels (Rosenstock, 1991). Under its new definition, vocational education has become an educational delivery system -- not just a content area (Kolde, 1991; Vaughan, 1991). The focus of vocational education has shifted from the traditional job-skills orientation toward the broader purpose of using vocational education as a vehicle for "learning academic and other kinds of thinking skills and for linking thought with action" (Wirt, 1991, p. 426). Complying with this new focus requires changing the occupational content of vocational education to include more generic academic content, thereby facilitating the overall goal of student achievement of skills -- problem-solving, effective communication, productive application of knowledge, and creative and independent thinking -- believed essential to an increasingly technological and complex workplace (Bailey, 1990; Kolde, 1991; Vaughan, 1991; Wirt, 1991).

Legislative Implications for Student Assessment

Although the evidence of assessment as a key facilitator in maximizing students' educational success has a relatively firm empirical and theoretical foundation, vocational education has limited literature, and virtually no empirical research, confirming the effects of student assessments. The exception to these lacunae are the writings on performance-based processes suggested as tenable assessment methods for students with disabilities (cf. Albright & Cobb, 1988a; Peterson, 1985; Phelps

& Wentling, 1977; Sitlington, 1981; Stodden, Ianacone, Boone, & Bisconer, 1987). That the 1990 legislation creates the potential for dramatic vocational course content changes is apparent. Not so apparent, however, is the underlying assumption that the legislation places on the evaluation of the influences of these content changes on student achievement (Darling-Hammond, 1990). Essentially, by sanctioning course curricula that support the learning of the new workplace literacies, the Amendments appear consistent with the larger educational agenda of raising intellectual standards. As mentioned earlier, intellectual standards originate and are subsequently maintained and modified through curriculum; which, in turn, is continuously influenced by assessment (Resnick & Resnick, 1985; Snow, 1989; Wiggins, 1989). Thus, it becomes imperative to accurately and fully measure student achievement while attempting to effect planned curricular changes. Stated differently, to encourage curricular changes without emphasizing relevant measures of student achievement almost guarantees that the new literacies intended by the legislation will receive short shrift in vocational education classrooms (Lewis, 1990; Nickerson, 1989).

The Relationship of Learning and Assessment

When done well, assessment for student learning reflects learning and development theories, promotes additional student learning, and provides teachers with accurate indices of student achievement and progress (Nickerson, 1989; Snow, 1989; Stiggins, 1991b). The preceding goals notwithstanding, the recent emphasis on developing student higher-level thinking skills would appear to require increasingly sophisticated assessment. Currently, developing measures that validly assess learning demands understanding the psychological structures and processes related to learning (Heshusius, 1991: Snow, 1989).

Cognitive Psychology and Educational Measurement

There are many new conceptions about the psychological and motivational structures and processes involved in learning (cf. Case & Bereiter, 1984; Kyllonen & Shute, 1989; Siegler, 1989;

Snow, 1989; Snow & Lohman, 1989). Although still in the developmental research stages, early findings indicate a wide range of cognitive and conative domains that assessment must address.

Cognitive structures. When mastering instructional content three overlapping yet sequential learning phases, or cognitive structures, are distinguishable: (a) the accretion of new information and its chunking, elaboration, and connection to existing knowledge; (b) informational restructuring, through which new knowledge organizations are formed, replacing and/or reformulating old conceptual relationships; and (c) the adaptation and application of knowledge structures in particular uses (Anderson, 1985; Snow, 1989). For example, in a vocational building trades classroom, a student might be presented with the task of framing an external wall to a house. Typically in new residential construction, external walls are built in modularized sections, including openings for doors and windows. Structural members are cut and assembled on the floor, or deck, of the new residence and the entire section is hoisted into place and affixed in its appropriate position at the periphery of the deck. As students are presented with the process of building an external wall, they first must learn the terminology and procedures for measuring and cutting wall members, and then for assembling and positioning the wall unit. New terminology and the discrete process of building an external wall are added to their existing knowledge of residential construction, and the physical/temporal act of external wall construction is integrated into their overall knowledge of house construction.

During the second learning phase, students would reconceptualize, or reformulate, their ideas about how external walls are built. Indeed, they would begin to compare and contrast the value of modularized construction, and relate these value judgments to those they have made about other forms of modularized assembly.

During the final learning phase, the students would begin to think about how this form of wall construction might work with internal walls, and with other forms of residential construction such as remodeling existing structures, frames for patios and decks, porches and garages, etc. This adaptation

and modification of knowledge structures is, of course, the target of educational reformers when they refer to critical thinking and higher order reasoning skills.

Researchers (Collins, 1988; Feuerstein et al., 1987; Glaser, 1988) have tied these three phases of cognition to classes or groupings of external measurable behaviors that appear to indicate students' relative levels of cognitive achievement. These behaviors then become the focal points upon which assessment strategies are built.

The first class of behaviors is that which reflects the coherence of knowledge. New knowledge about a given topic is fragmented and superficial. It is only as learning progresses and understanding occurs, that the learner begins to structure and organize concepts. Thus, assessment should measure the "connectedness" of concepts and the student's ability to access "interrelated chunks" of information (i.e., whether the student acts as a summarizer of information or whether she develops a point of view). In our building trades example, one way coherence of knowledge might be demonstrated would be that students would understand that both the top (plate) and bottom (shoe) horizontal members of the assembled external wall would be measured and marked from the same end, and in the same places, to assure that the vertical members (studs) are indeed vertical (plumb) after assembly.

Learning is demonstrated by principled problem solving. Advanced learners recognize underlying principles and patterns needed to solve a problem, generally ignoring the surface features of a task. Assessing for this component implies ferreting out the learner's ability to transfer appropriate algorithms across a variety of tasks. During external wall construction, students might be assessed for their ability to generalize the measuring, cutting, and locating of shortened studs around windows and doors from the way they performed these tasks on full-sized wall studs. Similarly, students might be asked to figure out how to frame the unique part of the external wall where interior walls (partitions) abut against the external wall.

The third class of behaviors are those associated with knowledge use, or the depth of conceptual understanding of the contextual conditions that facilitate the application of knowledge. Assessing for these behaviors means capturing the relevant information and skills as they are exercised in the context of the larger learning task (Wolf, Bixby, Glenn & Gardner, 1991). In the building trades example, students would be assessed not only for their skills in assembling an external wall, but in fitting the wall on the foundation deck. Clearly this implies a procedural issue around assessment -- that skills in external wall construction are best assessed on site, and during the actual process of performing the tasks.

Automatized skills are behaviors that reflect the degree to which the learner has inculcated routine, fundamental components necessary to apply the larger intellectual concept. Here, building trades students could be assessed relative to their abilities to independently apply their knowledge of modularized external wall construction in other sections of the residential construction (i.e., dormer walls, gable ends, internal wall partitions, etc.).

Metacognitive or self-regulatory skills include those behaviors suggesting learners' ability to monitor their own understanding, use strategies to make questions comprehensible, evaluate the relevance of accessible knowledge, and verify their own solutions. Assessments measuring such skills should include classroom practices through which students reflect on and critique their own work (Schwartz & Viator, 1990; Wiggins, 1989, Wolf et al., 1991). Here, building trades students would be asked to critique how structurally rigid their work was, or how well their walls conformed to requirements of external wall sheathing, internal dry wall application, or insulation installation.

Conative structures. Conative structures are those motivational (extrinsic and intrinsic stimulants or inducements) and volitional (intrinsic predilections) constructs such as confidence and interest that are rooted in the learner's ability or personality (Siegler, 1989; Snow, 1989). Although differentiated from cognitive constructs, conative constructs have direct implications in cognitive

processes, and thus must be understood and measured as they intercede in the learning context (Siegler, 1989; Snow, 1989).

Motivation for continued learning and achievement, interest in subject matter, a sense of confidence, and self-efficacy as a learner are thought to derive through task engagement and the subsequent personalization of task goals (Siegler, 1989; Snow, 1989). The relative degree of both learner task engagement and the personalization of goals, on the other hand, is the result of previous interaction with learning and the ensuing attitude toward achievement ability. These attitudinal orientations determine initial learner expectations about the instrumentality of learning for achieving personal goals, and the relative needs for achieving success and avoiding failure (Siegler, 1989). Moreover, students' attitudes toward their ability influences their motivational orientation toward learning, subsequently affecting knowledge acquisition and performance (Dweck & Leggett, 1988).

The implications for vocational assessment processes associated with the conative elements in learning theory are far less direct than for the cognitive structures noted earlier. An important construct of vocational assessment, in fact the most frequently cited assessment domain, is vocational interests. Our instrumentation and interpretive abilities to reliably measure enduring vocational interests of secondary vocational education students have lagged well behind our desire to measure those interests. Too, our vocational interest instrumentation and measurement technologies have focused almost exclusively on generalized images students may have of *occupations* rather than the training tasks, processes, and environments associated with *learning* about those occupations. Whether or not these student images truly reflect the complex mosaic of almost every occupation is highly suspect, especially for school-age youth.

In our building trades example, interest and motivation in this subject matter may well have to extend far beyond traditional assessments for interest in the occupations associated with residential construction. Rather if assessment processes are linked to this conative structure as a part of the

learning process, then measurement must be related to motivational issues around the tasks, processes, and environments in which the occupational skills are learned. While external wall construction, per se, may not be the focus of this form of assessment, building trades teachers might have to continually assess how students' interest levels fluctuate as significant new pieces of the occupational jigsaw puzzle of a residential carpenter are aggregated into their evolving perception of this occupation. Examples of these features include working out-of-doors in the rain, mud, and cold for sustained periods of time; working at significant and potentially precarious heights; working under conditions of heavy lifting; using dangerous and noisy portable power tools; and working with difficult non-wood products such as concrete, asbestos shingles, metal roofing, gypsum wallboard, fiberglass insulation, etc.

Ecological factors. There is a growing awareness that schools are social systems and that the social aspects of schooling are centrally related to student outcomes. Social processes such as interactions of students with teachers and peers shape the ways a person constructs and masters knowledge (Goodlad, 1984). Cultural values further contribute to knowledge mastery, defining and clarifying notions of personal meaning. Additionally, the specific nature of the curricular environment (i.e., the flexibility of the instructional delivery methods, settings and content in adapting to individual learning differences and preferences) strongly influences the ability of each student to learn that curriculum (Heshusius, 1991). Thus, to consider learning as separate from its context, incorrectly fragments the ensuing evaluation and ignores essential aspects of acquiring knowledge (Glaser, 1988; Heshusius, 1991).

Summary

Emerging theories suggest that learning is contingent on cognitive development, which is in turn, influenced by conative and ecological factors. Difficult to separate, these three domains -- cognitive, conative, and ecological -- all make important contributions to understanding student achievement. Providing an overview of learning constructs, Figure 1 illustrates the interrelationships of the three domains.

Insert Figure 1 about here

Referring to Figure 1, there are three phases defined under the cognitive domain. Research has revealed that these three phases of cognitive development -- acquiring data, restructuring data, and applying data -- are hierarchical. To measure the extent of development in each phase, assessment should target the characteristics of five manifestations of cognitive skill--coherence of knowledge, principled problem solving, knowledge use, automatized skills, and metacognitive, or self-regulatory, skills.

The conative domain has two dimensions--motivation and volition--that can be determined by evaluating the learner's self-confidence, self-efficacy, interest in subject, aptitude, ability, and learning style. The ecological setting, the third domain, refers to the influences of the learning environment such as instructional delivery (e.g., cooperative learning, inquiry, didactic, etc. methods), instructional setting (e.g., laboratory, regular classroom, community-based site, etc.), and curriculum format and content.

As each domain interacts with the other, their collective influences should be understood when formulating an opinion of student achievement. Thus, assessment becomes holistic, considering not only achievement outcomes, but also the cognitive processes through which the learning occurs, the learning context, and the learner's attitudinal predilections (Feuerstein et al., 1987; Glaser, 1988; Heshusius, 1991; Siegler, 1989; Snow, 1989; Wiggins, 1989).

Implications of These Domains for Developing Assessment

A comprehensive assessment approach suggests a number of assumptions for both developing and interpreting measures (Heshusius, 1991). Compiled from the recent literature (Brown, 1989; Feuerstein et al., 1987; Glaser, 1988; Heshusius, 1991; Marston, 1989; Nickerson, 1989; Shepard, 1989; Siegler, 1989; Wiggins, 1989; Wolf et al., 1991), these assumptions suggest a number of redirections for the focus of assessment.

1. Valid indicators of learning are outcomes of authentic learning processes, which are defined as activities that are meaningful and purposeful to the learner. Meaning and context are central to all learning and, therefore, to any assessment activity.

2. Learning is directly attributable to self-organizing and self-regulating principles within the learners as they interact with the environment.

3. Legitimate assessment reflects the complexity of learning through multiple indicators and measures. Moreover, kinetic, intuitive, and artistic expressions are additional ways that learning can be expressed.

4. "Errors" are an intrinsic, natural, and valuable part of learning and can provide insights into students' cognitive and conative structures. Analyzing these "errors" to guide relearning is fundamental to assessment.

5. Learning does not occur in a stable, linear, progressive manner. Assessment activities need to allow for this natural variability in accomplishments. Additionally, both proximal and distal instructional goals should be measured; assessment targeting only immediate learning neglects the longer range objective of knowledge application and transfer.

Subscribing to the preceding assumptions, shifts the focus of assessments to direct measures of student performance. Direct assessments evaluate the cognitive skill of interest as it is expressed in the performance of some extended task. The bases of direct assessment are predominantly the processes and outcomes of actual learning activities as they occur in the actual settings (Heshusius, 1991; Shepard, 1989; Wolf et al., 1991). Ultimately, of course, focusing on direct forms of assessment places greater reliance on teacher-directed, classroom assessments (Shepard, 1989; Stiggins, 1991b; Wiggins, 1989).

Assessing Student Achievement

"An enlightened teacher is the best evaluator of students' growth in process learnings" (Costa, 1989, p. 2). Teachers can directly observe and collect evidence of student performances that demand

creativity, ingenuity, mastery and the application and transfer of knowledge. However, teachers as consumers and developers of assessments need strategies that can validly evaluate the complex learning objectives underlying the acquisition of higher cognitive processes (Costa, 1989; Snow, 1989; Stiggins, 1991b).

Validity in Assessment

Validity, typically defined as the ability of the test to measure what it is designed to measure, has always been a major concern of test developers and users (Nickerson, 1989). Generally, validity in testing has been perceived to be comprised of three discrete but interrelated facets: criterion-related validity, content validity, and construct validity. Recently, however, theorists have taken a broader view of validity, emphasizing not only the psychometric aspects of *test* validity, but also ethical, social, and organizational validities of *assessment processes* (Linn, Baker, & Dunbar, 1991; Frederiksen & Collins, 1989; Messick, 1989; Nickerson, 1989).

For example, Messick (1989) recently defined validity as "an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the *adequacy* and *appropriateness* of *inferences* and *actions* based on test scores or other modes of assessment " (p. 5). Through his definition, Messick encouraged test developers and users to realize and to consider the "social consequences" of decisions made on the basis of assessment results. Linn et al, (1991) have added to Messick's concept of consequences of assessment several additional validity criteria, including fairness, transfer and generalizeability, cognitive complexity, content quality and coverage, meaningfulness, and cost and efficiency. Thus, assessment validity would appear to circumscribe a much broader range of concerns than has traditionally been the case when considering assessment of student achievement in different subject matters.

Accepting the fact that validity encompasses the consideration of testing consequences raises the issue of systemic validity, or the causal role of assessments in driving instruction. Teachers,

influenced by district administrators, historically "teach to the test", particularly when the adequacy of schools is judged on the results of tests (Kirst, 1991; Shepard, 1991). Because this is such a pervasive practice, assessment should support the learning of classroom objectives that are guided by educational standards (Nickerson, 1989). This means designing systemically valid tests "that induce in the educational system curricular and instructional changes that foster the development of the cognitive skills that the test is designed to measure" (Frederiksen & Collins, 1989, p. 27). Teaching to such tests could then eventuate legitimate educational objectives as assessments would be reflections of both what students should be attempting to learn as well as standards of desired performance (Frederiksen & Collins, 1989; Nickerson, 1989).

Assessment Strategies

There is a growing consensus that many of the standardized instruments traditionally used to assess student's intellectual abilities or educational progress are not only ineffectual, but inaccurate measures of higher cognitive processes (Brown, 1989; Ennis, 1985; Frederiksen & Collins, 1989; Nickerson, 1989; Resnick, 1987). Such tests have been criticized on a number of dimensions. (For alternative perspectives on standardized testing, review Jones & Idol, 1991, and Linn, 1989).

1. Standardized tests assume an outcomes-only orientation, analyzing the students' current levels of performance and thus neglecting the processes that may have operated or failed to operate to bring about the performance (Campione & Brown, 1987; Feuerstein et al., 1987; Heshusius, 1991; Marston, 1989).

2. Standardized achievement and aptitude tests are poor predictors of student potential, limited in their abilities to measure the responsiveness of a student to instruction, and the interaction of learning and environment (Brown, 1989; Campione & Brown, 1987; Hiebert & Calfee, 1989; Heshusius, 1991; Worthen & Spandel, 1991).

3. The content validity of many standardized achievement tests has often been criticized as specious. Prevailing published tests frequently fail to accurately sample the curriculum taught to the

student, meaning that scores do not reflect "true" student skills level (Marston, 1987; Shepard, 1989). This is doubly regrettable as tests dictate or restrict (i.e., "teaching to the test") what is taught (Kirst, 1991; Livingston, Castle, & Nations, 1989; Shepard, 1989). Moreover, defining performance in global terms, standardized tests yield information that is not easily relatable to developing or to modifying instructional programs, particularly at the local school classroom level (Campione & Brown, 1987; Marston, 1989). Seldom are standardized data used to develop effective intervention strategies for ameliorating learning deficiencies. Additionally, they are often racially, culturally, and socially biased (Worthen & Spandel, 1991).

4. Standardized tests are frequently taken as a permanent characterization of the individual. Such implications have intrinsic and extrinsic ramifications, damaging both the student's self-perception about learning ability as well as the educational staff's expectations for student achievement (Campione & Brown, 1987).

5. Standardized achievement and aptitude tests generally measure only limited and superficial student knowledge and behaviors, ignoring the more important indices of higher-level cognitive mastery (Brown, 1989; Ennis, 1987; Frederiksen & Collins, 1989; Worthen & Spandel, 1991). Moreover, such assessments are made outside the learning context, consequently ignoring the intercessions of environmental factors and learning (Campione & Brown, 1987).

6. Standardized achievement tests do not promote student learning (Brown, 1989; Frederiksen & Collins, 1989; Wolf, 1989; Worthen & Spandel, 1991).

From the burgeoning dissatisfaction with norm-referenced, standardized tests, alternative forms of assessment have emerged (Archbald, 1991; Herman, 1992). Under the rubric of authentic assessments, they hold great promise for classroom and school-wide use. Sensitive to teacher needs, many of these alternatives can be readily implemented in instructional settings (Lesh & Lamon, 1992; Kulm, 1991; Kulm & Malcolm, 1991) and can address the three domains of learning (cognitive, conative and ecological). Furthermore, they provide data that can guide and support classroom

curriculum as they emphasize the measurement of student performances directly tied to instructional goals, as well as measure short-term and long-term changes in students' knowledge through multiple implementations. Additionally, they either closely approximate, or are direct by-products of, actual learning tasks as they occur within the learning environment. Finally, each facilitates personal meaning and additional instruction to the student along with immediate, detailed, and complex feedback useful to both students and teachers to gain insight into the processes of learning.

Initially, overviews of four authentic assessment strategies--portfolio, exhibition, dynamic, and curriculum-based--are presented, detailing purposes, assumptions, characteristics, and outcomes. The section concludes with a discussion of their relative merits in addressing the three learning domains described earlier.

Portfolio assessment. A portfolio is a purposeful collection of student work that exhibits the student's efforts, progress, and achievements in one or more areas. The collection must include student participation in selecting contents, the criteria for selection, the criteria for judging merit, and evidence of student self-reflection (Paulson, Paulson & Meyer, 1991, p. 60).

The purposes of portfolio assessment are twofold: (a) to provide a way of evaluating student learning that, while informing teachers and school systems, will also model personal responsibility in questioning and reflecting on one's own work; and (b) to capture growth over time so that students can become informed and thoughtful assessors of their own histories as learners (Paulson et al., 1991; Wolf, 1989; Wolf et al., 1989).

These purposes are fostered by a number of assumptions about learning. First, the metacognitive or self-regulatory dimension of cognitive development is encouraged through self-analysis and self-critique of a learning task (Wolf, 1989; Wolf et al, 1991). Second, longitudinal assessments can provide insights into student cognitive growth and development (i.e., knowledge coherence, principled problem solving, knowledge use and automatized skills) as well as compensate for variability

in learning (Shepard, 1989; Wolf, 1989). Finally, students should be active managers of their own learning to engender personal relevance and meaning (Heshusius, 1991; Wolf, 1989).

The concept of portfolio assessment arose from the fields of art and literature where students would accumulate representative samples of their work. Recently adapted for use in a variety of instructional settings, portfolio assessments have several common characteristics. Students gather "biographies of works that represent a range of works and reflections" (Wolf, 1989, p. 37). Such biographies reveal the different developmental stages underlying the production of any major project. For a drafting student, biographies might include the notes, initial diagrams and sketches, and the final version of a remodeling project. The range of works is deliberately diverse to illustrate the depth and breadth of the project (Wolf, 1989; Wolf et al., 1991). Again, the student could collect photographs, published illustrations, hand-drawn sketches, computer-assisted diagrams as representative medium used in completing the project.

The reflections aspect of the portfolios are documents that come from the self-critiques wherein the student is expected to assume the stance of an expert (Wolf, 1989). During the critiquing process, students are required to note characteristics of their work, changes over time, and needed improvements (Wolf et al., 1991).

There are a number of beneficial educational outcomes attributed to portfolio assessments. Primarily, portfolios lend content validity to the assessment as the measured outcomes are the direct products of the classroom instruction (Wolf, 1989). Moreover, by involving students in the management and monitoring of the learning, the scope of learning is enlarged. For example, as the drafting student reviews her portfolio, noting the variety of products, she learns how her presentation and her choice of medium are crafted differently for varying audiences (Paulson et al, 1991; Wolf, 1989). Additionally, the portfolio facilitates the understanding of the learned process. As the student's work evolves, it provides evidence of how to perform the process. Knowing how to pursue the craft is as much a part of learning as is knowing the relative quality of the final product (Wolf et al., 1991).

Finally, portfolios can be used to articulate achievement levels as the student moves from course to course. Because portfolios provide direct evidence of performance, instructional planning that more realistically reflects individual ability is possible (Hiebert & Calfee, 1989).

In summary, portfolios weave together instruction and assessment in a way that facilitates student involvement in the learning; therefore, students are more likely to find relevance and meaning in the activities as well as higher levels of subject mastery. And, because the portfolios promote self-critiques of longitudinal evidence, students and teachers gain more valid information about cognitive developmental stages, facilitating more effective instructional planning.

Exhibition assessment.

> Exhibition activities challenge students to show off not merely their knowledge but their initiative; not merely their problem solving but their problem posing; not just their learning on cue, but their ability to judge and learn how to learn on an open-ended problem, often of their own design (Wiggins, 1989, p. 43).

Exhibition assessments are those that reflect intellectual ability through exhibitions of mastery. Such exhibitions should allow students to demonstrate initiative, knowledge, problem solving, problem posing, judgment, and the ability to learn. Additionally, the exhibitions should be representative of performances from the field (Wiggins, 1989).

Designing exhibitions implies a very different approach to assessment than is implied by criterion-referenced tests or outcome-based views of mastery. Exhibitions ideally embody and evoke desired outcomes in real-life contexts, requiring the student to have the judgment and skill to synthesize knowledge through a central challenge (Wiggins, 1989). The purposes behind the exhibition of mastery through the performance are to (a) design standard-setting tests that provide more direct evidence of a student's intellectual ability; (b) design tests that are able to stand by themselves as instructional objectives; (c) design intellectual challenges that reflect the heart of a discipline; and (d) design tests

that are more likely to engage students and motivate them to raise their own intellectual standards (Stiggins, 1991a; Wiggins, 1989).

Several assumptions guide exhibition assessment protocol. First, assessments are not after-the-fact tests for checking up on what students have learned; rather they are instructional, the central vehicle for clarifying and setting intellectual standards (Wiggins, 1989). Next, the responsibility of assessment is to capture the essential skills of inquiry and expression--a synthesis that requires questioning, problem posing and problem solving, independent research, creation of a product or performance, and a public demonstration of mastery calling for self-reflection and analysis of what a student has undergone and learned (Glaser, 1988; Shepard, 1989; Wiggins, 1989). Finally, a variety of assessment data yields a more vivid and reliable picture of student growth (Costa, 1989).

When implemented, exhibitions can assume several forms; however, each form has similar characteristics. The assessment process includes several direct-performance measures such as student portfolios and public presentations that give students the opportunity to persist in complex problem-solving situations and to pursue alternative problem-solving strategies (California Assessment Program, 1989). Every performance reflects the "ideal" standards of the school, and is designed to elicit depth in critical areas (e.g., problem solving, analyzing, synthesizing etc.) rather than breadth, or superficial knowledge in several areas (Wiggins, 1989). Assessment activities are also structured to direct students toward attaining more sophisticated levels of accomplishment. To this end, the assessment activities are both engaging and educational, encouraging student practice (Shepard, 1989; Wiggins, 1989). Furthermore, activities involve the student's own research or applications of knowledge, with self-assessment an integral part of the task. Additionally, the activities are judged on criteria that are understood by the students; criteria that are based on articulated performance standards (criterion-referenced) and that are judged by a multifaceted scoring system, which addresses each component of the activity (Wiggins, 1989). Finally, exhibition activities should be attempted by all students with the activity "scaled-up" rather than "scaled-down" to reflect student ability. Specifically, all students

shculd be held accountable for performing at some basic, acceptable level (Costa, 1989; Wiggins, 1989).

An example of an exhibition assessment might be an oral history project related to a vocational area. Students in landscape engineering, for instance, could be required to complete an oral history based on interviews and written sources, presenting the findings orally to the class. Subject matter choice could be left to the student, but such choices could include starting and/or running a small landscape company; a comparison of the artistic purposes of various landscape designs; landscapes and ecological factors; or water conservation and landscaping. As the topic is developed, the student could be asked to create three hypotheses based on preliminary investigations and questions to test each hypothesis. For a self-reflection component, a portfolio of evidence from the evolving history, including a written critique of the performance stages could be collected. The judgment criteria, clearly defined at the beginning of the assignment, would guide the project. A particularly effective way to develop meaningful criteria is through evaluation questions such as, "Did the student select appropriate sources for the interviews?" "Did the student use evidence to prove the ultimate best hypothesis?" "Did the student exhibit organization in writing and in presenting to the class?" Also, if the student learns that he has not satisfactorily completed aspects of the project, he could be given the opportunity to redo those aspects.

The outcomes of exhibitions support several of the standards suggested for student cognitive development. Students are directly and actively involved in the assessment process, both setting and achieving standards as well as self-critiquing; therefore students are encouraged to personalize tasks. Personalization facilitates meaning and relevance, which, in turn, sparks creativity, interest and the desire to achieve (Costa, 1989; Wiggins, 1989). Furthermore, assessment becomes a vehicle for additional learning, particularly when the judgment criteria are understood by the students (Wiggins, 1989).

Because the activities are contextualized, complex intellectual challenges, they reflect and promote higher-level cognitive development (i.e., problem solving, analyzing, and synthesizing). Ultimately, assessment and feedback become so central to student achievement that school schedules, structures, and policies are modified to support them (Wiggins, 1989).

In sum, exhibition assessments provide a forum for students to demonstrate control over the skills of inquiry and expression and control over an intellectual topic that approximates the expert's ability to use knowledge effectively and imaginatively.

Dynamic assessment.

> Dynamic assessment is an interaction between the examiner-as-intervener and a learner-as-active- participant, which seeks to estimate the degree of modifiability of the learner and the means by which positive changes in cognitive functions can be induced and maintained (Lidz, 1987, p. 4).

Dynamic assessment is focused on determining how a person learns rather than the outcomes or products of learning (Meyers, 1987). Both the examiner and the learner are active during the assessment; the examiner is an active intervener who continually modifies the interaction with the learner in order to induce successful learning (Jensen & Feuerstein, 1987; Lidz, 1987).

Although dynamic assessment has been used predominantly for assessing lower-functioning students, it has also been particularly successful in assessing culturally diverse students (Meyers, 1987). With increasing frequency, however, dynamic assessment models (cf. Feuerstein et al., 1987) have been recommended for use with all students (Hiebert & Calfee, 1989; Lidz, 1987).

Dynamic assessment has several purposes. Primarily, it attempts to evaluate as directly as possible the particular processes underlying successful performance (Campione & Brown, 1987). It also provides diagnoses of particular learning deficiencies within these processes in order to prescribe instructional interventions. Once a diagnosis is made, the effectiveness of the instructional intervention is monitored (Campione & Brown, 1987). Dynamic assessment also provides a testing modality that

accommodates the different cultural orientations of students, thus eliciting a more accurate estimate of learning potential (Budoff, 1987; Campione & Brown, 1987). Finally, it attempts to provide a testing forum wherein the learner can perform the task relatively self-sufficiently and independently depending on ability (Campione & Brown, 1987).

Implemented most widely, the Feuerstein et al. (1987) dynamic assessment model has several assumptions about learning and assessment. First, deficiencies in cognitive functioning are not immutable, but rather modifiable. To determine such cognitive process deficiencies, assessment tasks must address higher mental processes. Next, students' problem-solving strategies and learning styles have a significant effect on learning (Meyers, 1987). Therefore, assessment cannot be a static process; the examiner must assume a participatory role, presenting a variety of instructional modalities throughout the assessment to maximize the student's learning potential. Last, one of the critical features of assessment is it instructional capability. Therefore, assessment must be based on actual learning tasks.

When implemented, dynamic assessments have a number of representative characteristics. All the tasks used during the assessment must address higher mental processes, and must exemplify optimal student performance levels (Feuerstein et al., 1987). After selection, each task is broken into its incremental components that are subsequently linked to specific cognitive processes (Feuerstein et al., 1987; Meyers, 1987). For the assessment, the student is asked to actively perform the task, with the examiner providing any assistance, noting the nature and frequency of the assistance, and hypothesizing about process deficiencies (Feuerstein et al., 1987; Meyers, 1987). During the assessment, the examiner introduces a variety of instructional modalities to determine which modality seems to best facilitate the learning. Finally, based on the assessment outcomes, instructional interventions are planned and implemented. Subsequent assessments are scheduled to determine the effectiveness of the interventions and the degree to which new knowledge is maintained (Feuerstein et al., 1987; Meyers, 1987).

An example of assessing for degree of math literacy will help illustrate the application of dynamic assessment. Initially, the student could be presented with a series of problems, progressing in degree of difficulty with those of greater difficulty subsuming the principles of the preceding problems. As the student works through the series, the teacher would give necessary assistance, noting the frequency and nature of the assistance. Should the student need the same type of help on each problem, it could indicate that he has not committed fundamental skills to memory. To test this preliminary hypothesis during the assessment, the teacher would instruct the student on such skills, noting any performance improvements. At the conclusion of the assessment, the student would be given supplementary instruction on the deficient skills. Subsequent assessments, usually conducted shortly after the initial assessment, would look for increased levels of autonomy in performance as well as new areas of process deficiencies.

Following a dynamic assessment model eventuates predictable outcomes. Both the amount and nature of interventions necessary to process changes are systematically determined. Learner strengths and weaknesses in performances are isolated and can be subsequently analyzed to determine the particular cognitive deficiencies. Hypotheses (e.g., Was the performance affected by different instructional modalities? Are the lower performances related to the level of task complexity?) about student performance can be generated and tested through follow-up student performances (Lidz, 1987; Meyers, 1987).

Additionally, etiologies of weaknesses and strengths can be evaluated for degree of severity and for degree of modifiability (Lidz, 1987). For example, if the failure is attributed to the learner's low degree of familiarity with content, additional materials can be assigned. On the other hand, if failure is attributed to one type of instructional modality (e.g., audio) and strength attached to another modality (e.g., visual), instruction can be conducted through the modality best facilitating learning.

In summary, dynamic assessment provides a flexible, interactive assessment context through which student learning deficiencies and potential can be discovered. Attentive to the learning process,

dynamic assessment assists in generating and testing hypotheses for modifying learning. Additionally, by providing follow-up assessments (i.e., test-treatment-retest), interventions can be evaluated for their effectiveness in meeting the learners' instructional needs.

Curriculum-based assessment.

> Curriculum-based assessment (CBA) is any set of measurement procedures that use direct observation and recording of a student's performance in local curriculum as a basis for gathering information to make instructional decisions (Deno, 1989, p. 1).

CBA, grounded in special education, emerged as an assessment technique for making screening, eligibility, and placement decisions for students with disabilities when it became apparent that standardized tests failed to accurately discriminate their learning deficiencies. Albeit not generally promoted as an assessment technique for regular education students, CBA is included in this discussion for two reasons. First, its strategies represent systematic data collection for identifying student deficiencies and proficiencies and can therefore be adapted across any instructional settings. Second, vocational education courses serve students with disabilities, and thus vocational teachers require assessment strategies sensitive to their needs. Moreover, an assessment model, curriculum-based vocational assessment (CBVA), based on CBA tenets and implemented in a number of vocational settings, appears to be a viable method for assessing mainstreamed students (cf., Albright & Cobb, 1988b; Stodden et al., 1987; Whichard-Morehouse, 1993).

The chief purpose of CBA is to provide data that are useful for guiding a number of instructional decisions such as identifying remediation needs, placing students in programs, and planning instruction. Additionally, CBA can monitor student progress as well as evaluate the instructional program (Marston, 1989).

A number of assumptions about assessment guide CBA. First, content validity is paramount to any assessment measure; therefore, all assessments must reflect the level of competence with the local-school curriculum (Shapiro, 1987). Moreover, standardized tests frequently fail to adequately measure

growth as they do not reflect the actual instruction received by the student (Marston, 1989). Next, indirect assessments (e.g., WISC-R and WRAT) provide no opportunity for task analysis of errors that could directly bear on instructional planning and placement. Furthermore, essential attributes of knowledge such as fluency and comprehension, are not always addressed in standardized tests. Consequently, observations of student performance are critical to adequately assessing ability (Marston, 1989).

Although there are several CBA models (e.g., fluency-based, accuracy-based, outcomes-based, and criterion-referenced), they share common characteristics (Fuchs & Deno, 1991; Jenkins, Deno & Mirkin, 1979; Marston, 1989). First, the measurement procedures are directly linked to the student's curriculum. All assessments use materials in which the students were instructed. Next, the assessment is of relatively short duration to facilitate frequent administration by educators. Frequent administration is crucial to the assessment of progress over time. The assessment can assume a variety of forms (e.g., verbal or written expression, manual performance, etc). However, each of these forms should be both time- and cost-effective, as well as easy to implement. Finally, the data from each assessment are graphed and charted, systematically monitoring student achievement.

Implementing CBA strategies produces characteristic outcomes. Because observations of performances are the basis of most CBA assessment, the observer (educator) can monitor the process the student used to derive the correct or incorrect answer. This latter piece is critical to error analysis and subsequent instruction. Additionally, performances allow students to display creative or novel solutions to problems, revealing unique learning styles (Marston, 1989). Because CBA is task analytic in nature, it provides data on the specific areas of deficiency or proficiency. Teachers can then respond with more effective instructional decisions (Deno, 1989; Marston, 1989).

As CBA techniques call for frequent administration of the assessment, the observer gains different aspects of student performance. Repeated, frequent measures allow the observer to view the student's performance across several days and at various stages in the problem solving. Such

information can provide important insights into student behaviors (e.g., attentiveness to task) as well as achievement. Also, if the assessment includes information recently acquired, the repeated assessment reinforces the learning (Marston, 1989). By taking repeated measures, the outcomes of instructional interventions can be monitored (Marston, 1989; Shapiro, 1987). Using systematic recording and graphing the assessment results, the teacher can determine the relative effectiveness of the interventions.

The following illustration demonstrates a CBA model used in a vocational setting for assessing student learning. To determine the learner's achievement level on a word-processing program, for example, the teacher might select the representative activity of typing and printing a business letter. Once selected, the teacher lists each procedural step comprising the task (i.e., task analyzes the process). During the actual assessment, the teacher would observe the student typing and printing the letter, making notes on the characteristics of the performance. When the task was completed, the final product would be evaluated for its conformity to course standards. Additionally, the typing and printing processes are reviewed, comparing the student's performance to the task analysis listing. Based on these reviews, necessary interventions would be assigned. If, for instance, the format of the letter was incorrect, interventions would support practice in correct formatting procedures. While noting the assigned intervention, the teacher would also chart the performance outcome. In this example, number of errors could be used to assess the letter's relative accuracy and to determine progress. Regularly scheduled repeat-assessments, would require the student to perform the same, or similar, task, with the teacher again monitoring the performance and noting both process and outcome changes. The charting and tracking of the interventions would continue throughout the assessments. The effectiveness of instructional interventions would be determined by the student's progress.

In summary, through systematic data collection, CBA provides a problem-solving format for determining student deficiencies, defining their underlying difficulties, considering alternative solutions, implementing alternative solutions, and determining when the deficiencies have been overcome.

Furthermore, it promotes performances linked to the instructional curricula, thus providing more accurate indices of student ability and growth.

Summary

To meet the increasingly complex demands placed on assessment, techniques that address the process of learning as well as the products of learning need to be developed. Moreover, they should be implemented at the classroom level, and should include interactive tasks that both reflect and encourage cognitive development.

Although improved assessment content is certainly essential, assessments also need to promote meaning and relevance to the students. Thus, assessments should, to the extent possible, incorporate projects that require the students to self-reflect, perform research, and develop their own standards of excellence. Moreover, assessments should provide immediate feedback that both the student and the teacher can use not only to measure progress, but also to set higher achievement goals.

It should be evident from the foregoing discussion that portfolios and exhibitions fit chronologically at the end of the learning sequence, while curriculum based and dynamic assessments function more readily during the process of learning. Emphasizing direct observations of behavior, portfolio and exhibition assessments that include longitudinal collections of student work; long-term, comprehensive research projects; writing samples; and public performances; encourage students to synthesize and analyze information from a variety of sources. Summarized in Table 1, these "product assessments" primarily focus on the outcomes of instruction, reflecting student mastery of pre-established goals and objectives through a variety of activities. Such activities also appear to promote problem solving,

---

Insert Table 1 about here

---

critical thinking, and self-monitoring. Moreover, they require students to produce and arrange information in ways that others will comprehend.

On the other hand, the systematic problem-solving paradigm engendered by both CBA and dynamic assessment supports informed teacher instructional interventions. Labeled in Table 2 as "process assessments", they encourage the students and teachers to systematically examine the details of how the instruction was perceived and assimilated, thereby becoming diagnostic tools.

Insert Table 2 about here

Through regular observations, teachers are able to determine areas of cognitive deficiencies, assign additional instruction, and subsequently monitor the effectiveness of such instruction. Students are therefore able to attain task mastery with greater ease and efficiency.

Portfolio, exhibition, dynamic, and curriculum-based assessments offer viable options for teachers evaluating student achievement. They emphasize critical components of the new literacies and generate information that supports teacher decisions. Moreover, when used concomitantly, they can capture instructional goals in differing ways, providing more valid representations of students' performances.

Adequacy of Alternative Strategies for Assessing Student Achievement

Performance-based assessment measures such as the four preceding strategies have been increasingly acclaimed by educators as more relevant, and thus more accurate, measures of student ability. Moreover, they appear to address the broader definitions of validity incorporating activities that integrate cognitive, conative, and ecological systems; and stressing the benefits of testing/assessment for improving student performance. For instance, deep understanding and higher order skills are demonstrated (and encouraged) during performances in which students have to generate explanations and assemble skilled, comprehensive presentations (Snow, 1989; Wiggins, 1991b). These

performances provide insight into the cognitive constructs of knowledge use, coherence of knowledge, principled problem solving, automatized skills and self-regulatory skills (Snow, 1989). Additionally, observing the final and preparatory performances offers the teacher and the student opportunities to monitor growth and progress across cognitive and conative domains. As students persist through problem solving and problem forming, they continually acquire and master new knowledge, which in turn, fosters the development of new cognitive skills. Furthermore, the development of short-term, regular assessment such as performance observations and test-instruction-retest strategies, can help isolate cognitive and conative weaknesses, thus facilitating timely, effective interventions (Snow, 1989).

Admittedly, there are little empirical data that demonstrate the manner in which performance-based strategies can validly and reliably test for student achievements. Indeed there is much disagreement about what constitutes legitimate student achievement particularly when testing for critical thinking ability. (For a discussion on the epistemological implications of testing for critical thinking, see Stephen P. Norris, 1989). However, other parts of the world have been assessing student performance directly for a number of years. The Canadians, British, Australians, French, Italians, Germans, Russians, and Dutch (to name a few) all routinely demand of their students the production of high-quality documents and oral performances--even in mathematics (Wiggins, 1991b). Encouraged by the achievements of students in these countries, there are currently several American efforts to promote performance assessments. Such efforts include: (a) statewide writing assessments in over two dozen states; (b) hands-on assessment in vocational programs in most states; (c) National Assessment of Educational Progress (NAEP) performance assessments in science, writing, history, and reading (Wiggins, 1991b). The technical soundness of performance tests, has been addressed (albeit limitedly) through the College Outcome Measures Project (COMP). COMP, a battery of assessments, have been used for ten years to evaluate the results of a liberal arts education, by requiring oral presentations, written position papers, as well as other performances. An evaluation of COMP conducted by American College Testing concluded that performance tests were both valid and reliable, measuring

certain process abilities such as problem solving that are not tested by the multiple-choice format (Wiggins, 1991b).

Additionally, several studies (cf., Deno, Mirkin & Chiang, 1982; Fuchs, Fuchs & Maxwell, 1988; Marston, Fuchs & Deno, 1986; Shinn, Tindall & Stein, 1988) have indicated that CBA processes are reliable and valid for measuring reading and math abilities of special education students. Also, a number of studies illustrated the efficacy of one particular dynamic assessment model for measuring the intelligence and learning potential of lower-functioning students in several educational settings (cf., Budoff, 1987). If learning processes can be generalized, and there is reason to believe they can (cf., Case & Bereiter, 1984; Norris, 1989; Snow, 1989), then theoretically such assessment practices could be successfully incorporated into any instructional setting and used with all students.

Discussion

While there is much disagreement about the possible ways to construct and to conduct student assessment, there is general agreement that it is crucial to student achievement. The quality of instruction and student learning is directly related to the quality of assessment both inside and outside the classroom.

Furthermore, it is becoming increasingly clear that understanding learning has more to do with acknowledging, respecting, and trying to make explicit individual interests, values and motives. Judging students' learning by standardized tests is limited at best as these tests ignore the conative and ecological factors impinging on the performance, generally measuring knowledge fragments out of meaningful contexts. Thus, more comprehensive, performance-based measures appear to be better for obtaining the kinds of information that promote cognitive development and learning. Most importantly, if higher-order cognitive functioning is a major goal of education, making use of continuous and informed teacher judgments supported by relevant, accurate assessments are essential to achieving this goal.

Placing teachers in such critical roles, however, raises the issue of teacher training. Unfortunately, teachers observe, comment on, and grade students' actions, homework, and projects generally without the benefit of professional training in assessment (Stiggins, 1991b). Yet, teachers are the keys to effective assessment, not only in developing valid assessments, but also in understanding the relationship between assessment and instruction and how this relationship fosters educational standards. Clearly, teacher training is a paramount concern and should be given national priority.

The debate over the technical adequacy of performance-based assessments will undoubtedly continue until they become more widely practiced. However, it should be recognized that the disagreement over the "true" meaning of test validity fuels the debate. Should validity only be concerned with the traditional scientific dimensions of criterion, content and construct? Or, should it address the broader implications of testing such as its causal role in driving instruction and its use in classifying and labeling students? If the ethics of assessment are to be considered, then validity might be better served by assessments that measure what students should be attempting to learn, assisting students in attaining standards of desired performance. Regardless of the ultimate definition, some agreement should be attempted prior to judging the technical adequacy of any assessment.

References

Albright, L., & Cobb, R.B. (1988a). Curriculum-based vocational assessment: A concept whose time has come. Journal for Vocational Special Needs Education, 10(12), 13-16.

Albright, L., & Cobb, R.B. (1988b). Assessment of students with handicaps in vocational education: A curriculum-based approach. Alexandria, VA: American Vocational Association.

Anderson, J.R. (1985). Cognitive psychology and its implications (2nd ed.). San Francisco: W.H. Freeman.

Archbald, D.A. (1991, December). Authentic assessment: What it means and how it can help schools. Madison: University of Wisconsin, National Center for Effective Schools Research and Development.

Bailey, R. (1991). Jobs of the future and the education they will require: Evidence from occupational forecasts. Educational Researcher, 20(2), 11-20.

Bailey, R., (1990). Changes in the nature and structure of work: Implications for skill requirements and skill formation. Berkeley, CA: University of California, National Center for Research in Vocational Education.

Brown, R. (1989). Testing and thoughtfulness. Educational Leadership, 46(7), 31-33.

Budoff, M. (1987). The validity of learning potential assessment. In C. S. Lidz (Ed.), Dynamic assessment (pp. 52-81). New York: Guilford Press.

California Assessment Program. (1989). Authentic assessment in California. Educational Leadership, 46(7), 6.

Campione, J.C., & Brown, A.L. (1987). Linking dynamic assessment with school achievement. In C. S. Lidz (Ed.). Dynamic assessment (pp. 82-115). New York: Guilford Press.

Case, R., & Bereiter, C. (1984). From behaviorism to cognitive behaviorism to cognitive development: Steps in the evolution of instructional design. Instructional Science, 13, 141-158.

Collins, A. (1988). Reformulating testing to measure learning and thinking. In N. Frederiksen, R. Glaser, A. Lesgold, & M. Shafto (Eds.), Diagnostic monitoring of skills and knowledge acquisition (pp. 75-87). Hillsdale, NJ: Erlbaum.

Costa, A.L. (1989). Re-assessing assessment. Educational Leadership, 46(7), 2.

Darling-Hammond, L. (1990). Achieving our goals: Superficial or structural reforms? Phi Delta Kappan, 72, 286-295.

Deno, S.L. (1989). Curriculum-based measurement and special education services: A fundamental and direct relationship. In M. R. Shinn (Ed.), Curriculum-based measurement (pp. 1-17). New York: Guilford Press.

Deno, S.L., Mirkin, P.K., & Chiang, B. (1982). Identifying valid measures of reading. Exceptional Children, 49, 36-45.

Dweck, C.S., & Leggett, E.L. (1988). A social-cognitive approach to motivation and personality. Psychological Review, 95, 256-273.

Ennis, R.H. (1985). Critical thinking and the curriculum. National Forum, 65(1), 28-31.

Feuerstein, R., Rand, Y, Jensen, M.R., Kaniel, S., & Tzuriel, D. (1987). Prerequisites for assessment of learning potential: The LAPD model. In C.S. Lidz (ed.), _ Dynamic assessment (pp. 35-51). New York: Guilford Press.

Frederiksen, J.R., & Collins, A. (1989). A systems approach to educational testing. Educational Researcher, 18(9), 27-32.

Fuchs, L.S., & Deno, S.L. (1991). Paradigmatic distinctions between instructionally relevant measurement models. Exceptional Children, 57(6) ,488-500.

Fuchs, L.S., Fuchs, D., & Maxwell, S. (1988). The validity of informal reading comprehension measures. Remedial and Special Education, 9(2), 20-28.

Glaser, R (1988). Cognitive and environmental perspectives on assessing achievement. In Assessment in the service of learning: Proceedings of the 1987 ETS Invitational Conference. Princeton, NJ: Educational Testing Service.

Goodlad, J.I. (1984). A place called school. New York: McGraw-Hill.

Gray, K. (1991). Vocational education in high school: A modern phoenix. Phi Delta Kappan, 72, 437-442.

Herman, J. (1992, June). What's happening with educational assessment? Los Angeles: University of California, Center for the Study of Evaluation.

Heshusius, L. (1991). Curriculum-based assessment and direct instruction: Critical reflections on fundamental assumptions. Exceptional Children, 57(6), 315-328.

Hiebert, E.H. & Calfee, R.C. (1989). Advancing academic literacy through teachers' assessments. Educational Leadership, 46(7), 50-54.

Jenkins, J.R., Deno, S.L., & Mirkin, P.K. (1979). Measuring pupil progress toward the least restrictive environment. Learning Disability Quarterly, 2, 81-92.

Jensen, M.R., & Feuerstein, R. (1987). The learning potential assessment device: From philosophy to practice. In C.S. Lidz (Ed.), Dynamic assessment (pp. 379-402). New York: Guilford Press.

Jones, B.F. I Idol, L. (Eds.), Dimensions of thinking: Implications for testing. Hillsdale, NJ: Erlbaum.

Kirst, M.W. (1991). Interview on assessment with Lorrie Shepard. Educational Researcher, 20(2), 21-23.

Kolde, R.F. (1991). Integrated learning for a competitive work force. Phi Delta Kappan, 72, 453-455.

Kulm, G. (Ed.). (1991). Assessing higher order thinking in mathematics. Washington, DC: American Association for the Advancement of Science.

Kulm, G., & Malcom, S.M. (Eds.) (1991). Science assessment in the service of reform. Washington, DC: American Association for the Advancement of Science.

Kyllonen, P.C., & Shute, V.J. (1989). A taxonomy of learning skills. In P.L. Ackerman, R.J. Sternberg, & R. Glaser (Eds.), Learning and individual differences (pp. 117-162) New York: Freeman.

Lesh, R., & Lamon, S. J. (Eds.). (1992). Assessment of authentic performance in school mathematics. Washington, DC: American Association for the Advancement of Science.

Lewis, A.C. (1990). The murky waters of monitoring achievement of the national goals. Phi Delta Kappan, 72, 260-261.

Lidz, C.S. (1987). Historical perspectives. In C.S. Lidz. (Ed.), Dynamic assessment (pp.3-34). New York: Guilford Press.

Linn, R.L., Baker, E.L., & Dunbar, S.B. (1991). Complex, performance-based assessment: Expectations and validation criteria. Educational Researcher, 20(8), 15-21.

Linn, R.L. (1989). (Ed.), Educational measurement (3rd. ed.). New York: Macmillan.

Livingston, C., Castle, S., & Nations, J. (1989). Testing and curriculum reform: One school's experience. Educational Leadership, 46(7), 23-25.

Marston, D.B. (1989). A curriculum-based measurement approach to assessing academic performance: What it is and why do it. In M.R. Shinn (Ed.), Curriculum-based measurement (pp. 18-78). New York: Guilford Press.

Marston, D., Fuchs, L.S., & Deno, S.L. (1986). Measuring pupil progress: A comparison of standardized achievement tests and curriculum-related measures. Diagnostique, 11, 77-90.

Messick, S. (1989). Meaning and values of test validation: The science and ethics of assessment. Educational Researcher, 18(2), 5-11.

Meyers, J. (1987). The training of dynamic assessors. In C.S. Lidz (Ed.), Dynamic assessment (pp. 403-425). New York: Guilford Press.

Nickerson, R.S. (1989). New directions in educational assessment. Educational Researcher, 18(9), 3-7.

Norris, S.P. (1989). Can we test validly for critical thinking? Educational Researcher, 18(9), 21-26.

Paulson, F.L., Paulson, P.R., & Meyer, C.A. (1991). What makes a portfolio a portfolio? Educational Leadership, 48(5), 60-63.

Perrone, V. (1991). Toward more powerful assessment. In V. Perrone (Ed.), Expanding student assessment. Alexandria, VA: Association for Supervision and Curriculum Development.

Peterson, M. (1985). Models of vocational assessment of handicapped individuals. Career Development for Exceptional Individuals, 8(2), 110-114.

Phelps, L.A., & Wentling, T.L. (1977). A proposed system for the identification, assessment and evaluation of special needs learners. Journal of Industrial Teacher Education, 14(3), 19-35.

Resnick, L.A. (1987). Education and learning to think. Washington, DC: National Academy Press.

Resnick, D.P., & Resnick, L.B. (1985). Standards, curriculum, and performance: A historical and comparative perspective. Educational Researcher, 14(4), 5-20.

Rosenstock, L. (1991). The walls come down: The overdue reunification of vocational and academic education. Phi Delta Kappan, 72, 434-436.

Schwartz, J., & Viator, K. (Eds.), (1990). The prices of secrecy: The social, intellectual, and psychological costs of current assessment practices. Cambridge, MA: Educational Technology Center.

Secretary's Commission on Achieving Necessary Skills (SCANS). (1991). What work requires of schools: A SCANS report for America 2000. Washington, DC: U.S. Department of Labor.

Shapiro, E.S. (1987). Behavioral assessment in school psychology. Hillsdale, NJ: Lawrence Erlbaum Associates.

Shepard, L.A. (1991). Psychometricians' beliefs about learning. Educational Researcher, 20(7), 2-16.

Shepard, L.A. (1989). Why we need better assessments. Educational Leadership, 46(7), 4-9.

Shinn, M.R., Tindall, G., & Stein, S. (1988). Curriculum-based assessment and the identification of mildly handicapped students: A research review. Professional School Psychology, 3, 69-85.

Siegler, R.S. (1989). Strategy diversity and cognitive assessment. Educational Researcher, 18(9), 15-20.

Sitlington, P.L. (1981). Vocational assessment and the individualized vocational plan. Journal for Vocational Special Needs Education, Spring Issue, 27-29.

Snow, R.E. (1989). Toward assessment of cognitive and conative structures in learning. Educational Researcher, 18(9), 8-14.

Snow, R.E. & Lohman, D.F. (1989). Implications of cognitive psychology for educational measurement. In R.L. Linn (Ed.,), Educational Measurement (3rd ed., pp. 263-331). New York: Macmillan.

Stiggins, R. (1991a). Assessment literacy. Phi Delta Kappan, 72, 534-539.

Stiggins, R. (1991b). Relevant classroom assessment training for teachers. Educational Measurement: Issues and Practice, 7-12.

Stodden, R.A., Ianacone, R.N., Boone, R.M., & Bisconer, S.W. (1987). Curriculum-based vocational assessment: A guide for addressing youth with special needs. Honolulu, HI: Centre Publications.

Vaughan, R.J. (1991). The new limits to growth: Economic transformation and vocational education. Phi Delta Kappan, 72, 446-449.

Whichard-Morehouse, J.A. (1993). Efficacy of a utility technology approach for selecting a vocational assessment process (Doctoral dissertation. Colorado State University, 1992). Dissertation Abstracts International, 53A(10), 3506.

Wiggins, G. (1989). Teaching to the (authentic) test. Educational Leadership, 46(7), 41-47.

Wiggins, G. (1991a). Standards, not standardization: Evoking quality student work. Educational Leadership, 48(5), 18-25.

Wiggins, G. (1991b). A response to Cizek. Phi Delta Kappan, 72, 700-703.

Wirt, J.G. (1991). A new federal law on vocational education: Will reform follow? Phi Delta Kappan, 72, 425-433.

Wolf, D.P. (1989). Portfolio assessment: Sampling student work. Educational Leadership, 46(7), 35-39.

Wolf, D., Bixby, J., Glenn, J., and Gardner, H. (1991). To use their minds well: Investigating new forms of student assessment. Review of Research in Education, 17, 31-74.

Worthen, B.R., & Spandel, V. (1991). Putting the standardized test debate in perspective. Educational Leadership, 48(5), 65-69.

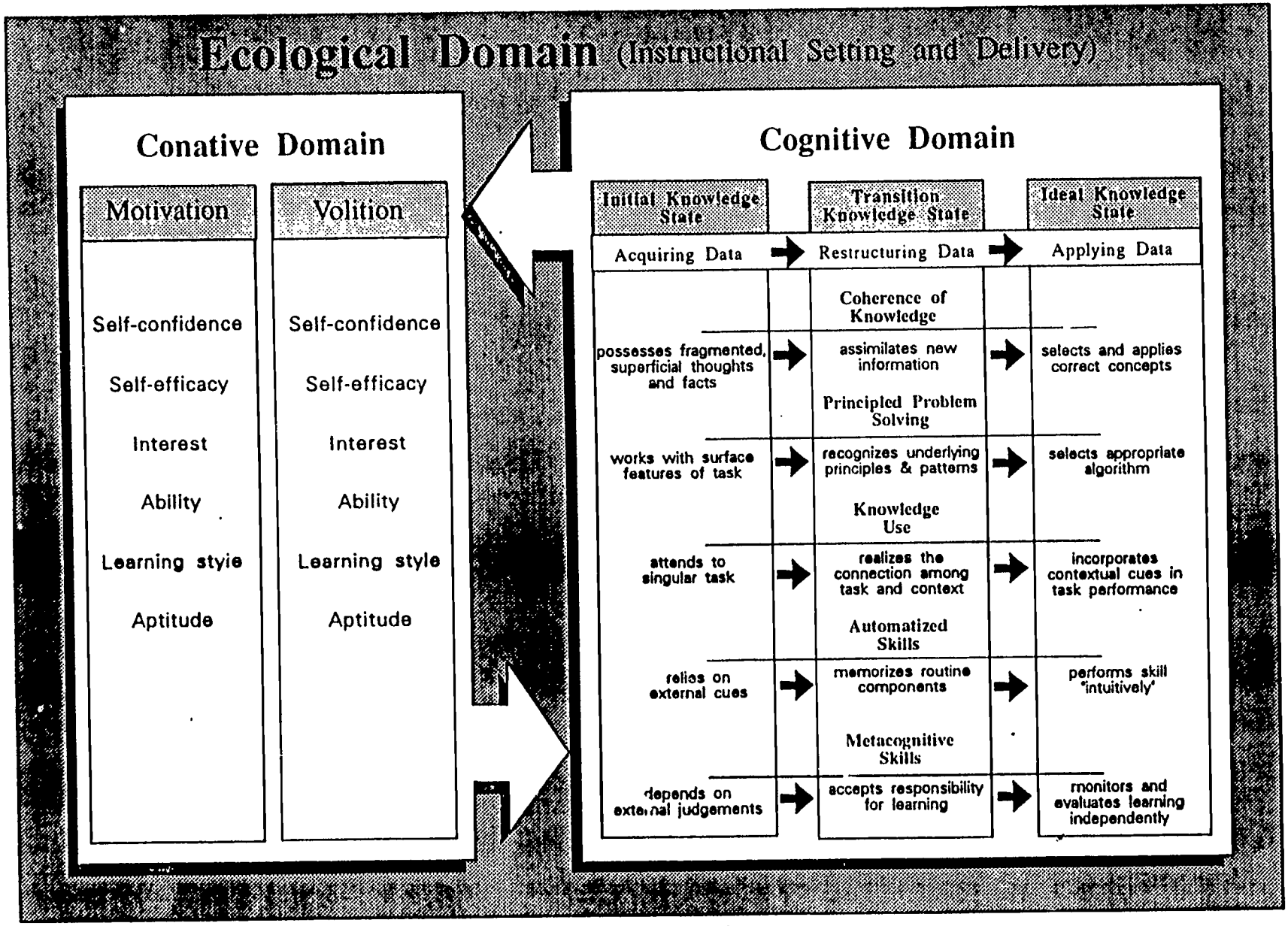

Figure 1. Influences of Conative, Cognitive and Ecological Factors on Learning

Table 1

Comparison of Product Assessment Strategies

	Portfolio	Exhibition
Definition	• Purposeful collection of student work that exhibits students' efforts, progress, and achievements. The collection must include student participation in selecting contents, selection criteria, criteria for judging merit, and evidence of student self-reflection.	• Assessments that reflect intellectual ability through exhibitions of mastery. Such exhibitions should allow students to demonstrate initiative, knowledge, problem solving, problem posing, judgment, and the ability to learn. Also, exhibitions should represent performances from the field.
Purposes	• Provide a way of evaluating student learning that while informing teachers and school systems, also will encourage student responsibility for learning. • Capture growth over time so students can become informed and thoughtful assessors of their learning histories.	• Provide direct evidence of a student's intellectual ability. • Are able to stand alone as instructional objectives. • Reflect the heart of a discipline. • Engage and motivate students to raise their intellectual standards.
Assumptions	• Metacognitive skills are facilitated through self-analysis and self-critique of learning tasks. • Longitudinal assessments provide insights into student cognitive growth and development as well as account for variability in learning. • Actively managing their own learning engenders personal relevance and meaning of learning to students.	• Assessments should be instructional. • Assessments should set and clarify intellectual standards. • Assessments should capture the skills of inquiry and expression. • Assessments should yield a variety of data for a more vivid and reliable picture of student growth.
Characteristics	• Students gather biographies of works that represent a range of works and reflections. • Biographies emphasize the different developmental stages underlying the production of a major project. • Range of collections is purposely diverse to illustrate the depth and breadth of learning. • Reflections aspect includes self-critiques of work characteristics, changes over time, and needed improvements.	• Assessment process includes several direct-performance measures. • Assessments give students the opportunity to persist in complex problem-solving situations. • Assessments are student performances that reflect the "ideal" standards of the school. • Assessments are designed to elicit depth in critical-thinking areas rather than breadth or superficial knowledge in several areas. • Assessment activities incorporate student's own research. • Assessment criteria are understood by student at the outset of the project.
Outcomes	• Give content validity to assessment. • Enlarge the scope of learning. • Facilitate understanding of how the process is learned. • Articulate achievement levels as student moves from course to course. • Support the instructional objectives.	• Students are directly and actively involved in the assessment process. • Students learn to critique own work. • Students are encouraged to personalize learning, which, in turn, sparks the desire to achieve. • Assessment reflects and promotes higher-level cognitive development. • Encourages students to use knowledge effectively and imaginatively.

Table 2

Comparison of Process Assessment Strategies

	Dynamic	Curriculum-based
Definition	• An interaction between the teacher-as-intervener and the learner-as-active-participant, which seeks to estimate the learner's ability to learn and how best to induce and maintain learning. During the assessment process the teacher focuses on how the student learns as well as the outcomes and products of learning.	• Any set of measurement procedures that use direct observation and recording of a student's performance in local curriculum as a basis for gathering information to make instructional decisions.
Purposes	• Evaluate the underlying processes of successful student performances. • Diagnose learning deficiencies to prescribe interventions. • Monitor the effectiveness of an intervention. • Assess learning potential of student. • Provide testing modality that accommodates different student learning orientations.	• Provide data that are useful for guiding instructional decisions such as identifying remediation needs, placing students in programs, and planning instruction.
Assumptions	• Deficiencies in cognitive functioning are modifiable. • To determine cognitive deficiencies, assessment must tap higher mental processes. • Learner's problem-solving strategies and learning styles have significant effects on learning. • Assessment cannot be a static process; teacher must actively participate to determine how student learns best. • Assessment also should be instructional.	• Content validity is paramount in assessment. • Assessment must measure students on what they have learned in the classroom. • "Errors" provide important information for analyzing student abilities. • Observations of student performances are essential for assessing fluency and comprehension.
Characteristics	• Assessment tasks address higher-order mental processes and must exemplify optimal student performance levels. • Student activity performs assessment task while the teacher observes and provides needed assistance. • Teacher notes the nature and frequency of the assistance. • Teacher introduces a variety of instructional modalities during assessment to determine which modality best facilitates learning. • Teacher determines necessary interventions. • Subsequent assessments are scheduled regularly to determine the effectiveness of the intervention and the degree to which the new knowledge is maintained.	• Assessments are directly linked to student's curriculum. • Assessments are relatively short. • Assessments are conducted frequently. • Assessments can assume a variety of forms but are always actively performed by students. • Assessments are easy to implement in the classroom. • Assessment data are systematically graphed and charted. • Instructional interventions are monitored regularly.
Outcomes	• Necessary interventions are systematically determined and monitored. • Learner strengths and weaknesses are isolated and subsequently analyzed to determine cognitive deficiencies. • Learner's learning potential is assessed. • Instructional interventions are continually monitored. • Student progress is systematically monitored.	• Students are given the opportunity to display creative or novel solutions to problems. • Teachers are given data that isolate areas of student deficiencies. • Student progress is monitored longitudinally. • Student progress is monitored systematically and regularly. • Assessment reinforces learning.