MPR
Associates, Inc.

Assessing the Feasibility of a Test Item Bank and Assessment Clearinghouse

Strategies to Measure Technical Skill Attainment of Career and Technical Education Participants

Assessing the Feasibility of a Test Item Bank and Assessment Clearinghouse

Strategies to Measure Technical Skill Attainment of Career and Technical Education Participants

Seth Derner
Vivayic Corporation
610 J Street, Suite 210
Lincoln, Nebraska 68508

Steve Klein
MPR Associates, Inc.
205 SE Spokane Street, Suite 344
Portland, OR 97202

Don Hilber
DTI Associates, Inc
2920 South Glebe Road
Arlington, VA 22206

Contents

List of Tables

iv

Background

Career and technical education (CTE) coursework is offered in secondary and post-secondary institutions throughout the United States. Students completing a sequenced program of CTE coursework are expected to have mastered state- and/or industry-recognized academic knowledge and technical skills that prepare them to enroll in a community college, four-year college, or university; pursue advanced career training in a public or private proprietary institution; enter an apprenticeship program; obtain employment; or enlist in the military. In many instances, states have established statewide technical skill standards or specified guidelines to assist local education agencies in defining what students should know and be able to do.

As a condition for receipt of federal funding via the *Carl D. Perkins Career and Technical Education Act of 2006 (Perkins IV)*, states are required to assess students' attainment of challenging technical skill proficiencies that are aligned with post-secondary program or industry-recognized standards. Although a handful of states have established statewide secondary exams or contracted for third-party assessments, the majority lack standardized testing systems and the resources or political will to create them. At the postsecondary level, institutional tests tend to be occupationally focused and faculty driven, with standardized assessments, where they exist, drawn from existing national industry certification or state credentialing and licensing exams.

To support states in measuring students' technical skill attainment, the U.S. Department of Education, Office of Vocational and Adult Education (OVAE) is collaborating with the National Association of State Directors of Career Technical Education Consortium (NASDCTEc), the Association for Career and Technical Education (ACTE), and other technical education stakeholders to explore the feasibility of (1) establishing a test item bank containing questions submitted by various business, industry, and education sources, and (2) compiling an assessment clearinghouse containing information about industry-recognized national assessments that may be adopted or adapted for state use. It is anticipated that the adoption of either or both assessment systems will produce substantial economies of scale, granting educators access to a vast catalog of reliable industry-validated test items and exams for a fraction of what it would cost states to develop each system individually.

This report documents strategies that can be used to initiate development of a technical skill test item bank and/or assessment clearinghouse and quantifies the cost of creating and maintaining such a system. It is intended to inform state administrators on the potential uses and benefits of system participation, test developers on the needs and expectations of secondary and postsecondary educators, and federal policymakers on system capabilities and constraints. This report opens with a review of the rationale for creating a test item bank and assessment clearinghouse and includes a discussion of current state approaches to assessment and the drawbacks associated with existing systems. Section 2 summarizes findings from a survey of state administrators that was used to identify state testing needs and outlines the design criteria researchers will need to address to meet state needs. Sections 3 and 4 draw upon state-identified design components to propose an implementation model for the test item bank and assessment clearinghouse, respectively. Each section closes by identifying a recommended model and the cost estimates associated with its creation. The paper concludes with recommendations for initiating the nationwide development of a test item bank and assessment clearinghouse.

Section 1: Rationale for System Development

Accountability provisions contained within the Perkins IV legislation extend measurement expectations introduced in prior legislation. One significant departure from past practice, however, is OVAE's decision to issue nonregulatory guidance detailing preferred approaches for constructing state technical assessment measures. Specifically, states are encouraged only to report on CTE concentrators tested using a state-established technical skill assessment or with a third-party, industry-recognized certification or credentialing exam.[1] Concentrators in courses or programs for which such assessments do not exist may be excluded from the measure, though states must document the percentage of individuals tested and their plans for expanding the use of state-established or national industry-recognized assessments to additional program areas.

These new technical assessment guidelines are intended, in part, to remedy measurement deficiencies noted by the Office of Management and Budget (OMB) in its review of the federal Perkins Act, using its Program Assessment Rating Tool (PART).[2] Labeling Perkins as ineffective, primarily due to the lack of a common standard of validity and reliability in state performance measures, the OMB finding has been cited by some policymakers calling for the elimination of Perkins program funding. Changes in the measurement approach also are intended to provide students with a means of directly demonstrating their knowledge and skills, allowing for their improved transition from secondary to postsecondary programs of study and in conveying proof of skill mastery to potential employers.

---

[1] Perkins IV defines a *CTE concentrator* as a secondary student who has earned three (3) or more credits in a single CTE program area (e.g., health care or business services) or two (2) credits in a single CTE program area, but only in those program areas where two credit sequences at the secondary level are recognized by the state and/or its local eligible recipients. A postsecondary/adult *CTE concentrator* is one who (1) completes at least 12 academic or CTE credits within a single program area sequence that is made up of 12 or more academic and technical credits and terminates in the award of an industry-recognized credential, a certificate, or a degree or who (2) completes a short-term CTE program sequence of less than 12 credit units that terminates in an industry-recognized credential, a certificate, or a degree.

[2] The PART rating can be accessed at http://www.whitehouse.gov/omb/expectmore/summary/10000212.2002.html.

3

State Approaches for Assessing Technical Skill Attainment

To date, states have attempted to develop performance measures for technical skill attainment that balance federal data needs with state capacity and resources to collect data. While some states have pioneered statewide assessment systems or have contracted for third-party exams that conform to high standards of validity and reliability, others have adopted indirect measures, such as program completion or grade point average (GPA), to proxy student attainment.

4

A review of state measurement approaches proposed in states' five-year Perkins plans suggests that federal nonregulatory guidance may be helping to reduce the variability in states' constructions of secondary measures, though postsecondary measures are less consistent. As illustrated in Table 1, 44 of 53 state secondary agencies (83 percent) are planning to use state-approved or national assessments to assess CTE concentrators' technical skill attainment, compared to 34 of 53 state postsecondary agencies (64 percent). Grades and cumulative GPAs, which may provide more subjective measures of student learning, are the second most prevalent measurement approach being used among states using alternative construction strategies.[3]

Table 1. State Perkins measurement approaches: Program year 2009–2013		
Measurement approach	Secondary	Postsecondary
National or state-approved assessment	44	34
Grades or GPA	4	10
Program completion	2	1
Assessment and program completion	0	1
Completion of 80 percent of standards	1	0
Mastered industry-validated standards	1	0
Course completion	1	1
Assessment and grades	0	5
Completed degree/certificate or transferred within four years	0	1

SOURCE: U.S. Department of Education, Office of Vocational and Adult Education.

This fragmented approach to state testing means that students across, and often within states, are assessed on differing technical content, using a range of assessment

---

[3] Care must be taken, however, when interpreting these findings. Past experience suggests that measurement approaches that appear similar in structure may produce substantially different results, because states may adopt differing criteria to define CTE concentrators, differing constructions to measure numerators and denominators, and differing methodologies to collect student data.

instruments, performance criteria, and rating systems. For example, measures based on state-approved, locally developed assessments may lack both validity and reliability, because instructors may assess students based on their own beliefs about workplace skill needs and/or rate students using subjective criteria or differing scoring rubrics to establish passing thresholds. Such widespread variability in standards and assessments undercuts federal efforts to assess CTE's contribution to preparing students for entry into higher education and careers and may lead to a misalignment between educational training and workforce needs.

5

States lacking standardized, statewide technical skill assessments often cite resource constraints as the limiting factor in developing their systems. Constructing statewide assessment systems requires the involvement of education, business, industry, and labor representatives in identifying and validating technical standards, test developers in designing assessment instruments, and psychometricians in assuring valid and reliable testing procedures. The high costs associated with convening experts and administering and updating assessments can make it cost prohibitive for individual states to undertake assessment development. Indeed, states that have pioneered statewide assessments, such as Arizona and Utah, have done so in response to state legislative policies mandating and funding system development.

Practical constraints will likely constrain test development even among states making good faith efforts to align their assessment systems with OVAE nonregulatory guidance. Due to technical, administrative, and resource constraints, not all states will be able to develop or use appropriate technical skill assessments that fully cover all of their CTE programs spanning secondary and postsecondary education. For most, the expectation is that there will be progress in the early years, but that gains will stall out for program areas at one or both levels, either because industry-recognized assessments are not available or there is resistance from educators at local institutions. Differential investments in assessment design will likely result in a patchwork of state systems, with some advanced and still more deficient in terms of assessment capacity.

The proposed technical skill test item bank and assessment clearinghouse can resolve many of the logistical, technical, and cost issues associated with performance measurement.[4] And because development of a single national assessment will not meet all states' needs, design flexibility engineered into the system would permit states to tai-

---

[4] While alternative testing systems may also prove feasible, the technical skill assessment task force convened for this study has focused on the test item bank and assessment clearinghouse as the most practical systems for nationwide use, for the reasons documented below. It is anticipated that the concepts posited in this report will undergo continued evolution and refinement based on the contribution of test developers and other stakeholders seeking involvement in this project.

lor assessments to conform to state-established standards, while still aligning the system with industry-recognized and validated skills. The following section lays out design criteria that may be used to structure a test item bank and assessment clearinghouse and summarizes states' interest in creating a national assessment system employing these criteria.

Section 2: Design Criteria

When selecting components for a national assessment system, assessment task force members sought to accommodate the multiple educational goals that characterize CTE instruction across states and educational sectors within states. The following section provides an overview of the components deemed crucial to a national assessment system and summarizes states' attitudes toward these components based on a survey of state administrators. Drawing upon identified needs, the section closes by presenting a set of underlying assumptions used to guide test item bank and assessment clearinghouse design.

Components of a National System

In an increasingly global economy, employers are demanding workers who possess advanced knowledge and skills that are both verifiable and transferable. These skills are typically offered at the postsecondary level, where students enroll in occupationally specific coursework to obtain marketable skills that will enable them to find immediate employment. At the same time, states are working to develop secondary programs of study that introduce high school students to broadly defined CTE curricula spanning a career cluster and one or more pathways. Task force members also recognized that some states have already invested in building their own technical skill assessment systems, using assessments created by state agencies or third-party developers or employing national industry certifications and state-based licenses to verify students' technical skill attainment.

After considering available options, task force members opted to pursue dual assessment strategies that could, in some combination, address the multiple demands of a nationwide technical skill assessment system. The primary approach calls for developing a national test item bank, from which states could build their own assessments to measure student attainment of state-established, industry-recognized technical skill standards. To provide for flexibility across states, item bank elements could be harnessed to create national assessments at various skill levels, ranging from general work readiness, to career clusters and pathways, to occupationally specific skills. Initial plans call for furnishing the bank with test items aligned to the knowledge and skill statements identified by the States' Career Clusters Initiative (SCCI) in order to support states in evaluating outcomes of students who participate in CTE Programs of

Study. The item bank could also manage additional items aligned to other industry, postsecondary, or state standards.

A second focus calls for creating an assessment clearinghouse to catalog existing, standardized, national occupational certification, credentialing, and licensing exams that could be adopted (or adapted) for use at the secondary and postsecondary levels. This proposed approach would not provide states with direct access to exams or the ability to administer tests, but it could present users with information on their availability by pathway and specialty area, along with details on their purposes, length, cost, and administration dates.

The combination of the two assessment systems would provide states with the capacities shown in the Table 2 below.

Table 2.	Test item bank and assessment clearinghouse features		
Feature		Test item bank	Assessment clearinghouse
Assess knowledge and skill standards in SCCI identified Programs of Study		✓	
Assess skills identified in state-established Programs of Study		✓	
Assess skills identified in state-approved CTE programs or coursework		✓	
Provide access to valid and reliable assessments		✓*	✓
Improve consistency of state CTE assessments		✓	✓
Identify existing validated, industry-recognized assessments			✓
Compare existing assessments for a given field			✓

* Assembling test items into tests will require additional state-provided effort and resources.

Prior to embarking on a specific plan for design and deployment, task force members surveyed states' CTE administrators to determine their overall interest in an initiative to develop a national assessment, the features that they desired in such a system, and their willingness to finance assessment development.

Defining State Assessment Needs: State Survey Findings

In March 2008, task force members conducted a voluntary Web-based survey of state CTE directors. A total of 37 states responded to the survey, with participants including state secondary and postsecondary directors or their designees—typically a state-level Perkins accountability or assessment expert. It is not clear whether survey nonrespondents lacked interest in the project or simply declined to participate.

Among those providing information, responses were obtained from both secondary and postsecondary administrators in 27 states, and from either the secondary (5 states) or postsecondary (5 states) administrator in the remainder. While survey results can provide some insight into states' interest in the project, participating states were not representative of states nationwide, nor were individual respondents necessarily representative of official state policy. In some instances, states that previously had indicated an interest in the project also did not complete a survey. As such, care must be taken in interpreting findings.[5]

States were generally supportive of the national assessment approach, with 30 of the responding states reporting that they were "somewhat" or "very interested" in pursuing the creation of a test item bank. A majority of these respondents indicated that they would be "somewhat" or "very likely" to access the item bank assessments to create secondary-level tests that would assess the knowledge and skills identified in the SCCI standards (20 states) or to create customized exams to assess state-established standards (21 states). Interest was somewhat less pronounced at the postsecondary level, with roughly half of states indicating they were "likely" to use the item bank to develop assessments of SCCI or state-established standards (16 and 15 states, respectively). States expressed similar levels of interest in accessing the assessment clearinghouse to identify existing CTE exams for either secondary or postsecondary assessment purposes (26 states each).

States expressing interest in the test item bank were questioned further to identify important features of the proposed system. As illustrated in Table 3, state administrators placed high emphasis on having a Web-based system that would allow them to address Perkins accountability reporting requirements, while permitting them to create customized assessments to assess state-established secondary and postsecondary standards. State administrators placed less value on allowing individual school districts or postsecondary institutions to create their own local assessments.

---

[5] See Appendix for a description of the survey methods and cross-tabulations of survey results.

Table 3. States rating test item bank features as extremely or very important	
Feature	Number
Collect and aggregate data for Perkins reporting purposes	23
Provide on-line test administration	22
Create customized statewide assessments for Secondary	22
Allow access to states to choose items, create own assessments	19
Create customized statewide assessments for Postsecondary	15
Allow access to school districts to create own assessments	7
Allow access to faculty / teachers to create own assessments	6

States expressed some interest in supporting the creation of item banks, with 17 states indicating they would be likely to contribute questions for expert consideration. States indicated that they would be less willing to commit resources to the project, with 10 states reporting that they would provide start-up funding to seed item bank development and 15 states reporting that they would supply annual funding to gain access to item bank questions. States' unwillingness to offer start-up resources may be due, in part, to vagaries associated with project status and components; indeed, in follow-up conversations with a subset of respondents, administrators reported that they would be hesitant to promise resources until they had a clearer understanding of how the test item bank would function.

Guiding Assumptions

Form must follow function, and as such, task force members have crafted a set of assumptions to guide system development. These assumptions, which build upon information gathered though electronic surveys of state CTE administrators, formal and informal feedback collected at national presentations, and meetings of stakeholder groups conducted by task force members, provide a basis for subsequent feasibility models contained in this report.[6]

Assumption 1: Federal Role in System Development

It is anticipated that OVAE administrators will play a vital role in the design, acceptance, and promotion of the test item bank and assessment clearinghouse, but will neither provide major funding nor retain ownership or management rights over system development and deployment. OVAE will support efforts to legitimize the col-

---

[6] Interim study findings were presented at two national meetings: the spring meeting of the National Association of State Directors of Career and Technical Education Consortium, held in Washington, DC in April 2008, and the annual conference of the National Association for Career and Technical Education Information, held in Boise, ID in May 2008.

laborative item bank, and may encourage its use, but it is expected that a third-party public or private organization will own the rights and responsibilities to its operation. It is further assumed that OVAE will be actively involved to ensure that data generated from the test item bank and its derived tests, as well as assessment items identified in the clearinghouse, will be acceptable for Perkins reporting requirements.

Assumption 2: Test Item Bank Design and Test Delivery Mode

It is assumed that the test item bank and the system for designing and delivering tests will be Web-based and built and managed in an electronic database format that meets current industry standards for use, portability, and accessibility. Likewise, it is assumed that test delivery, test specification design, user management, and reporting functions will use electronic applications and database structures requiring Web access. There will not be a system option to access test items, tests, or reports in a hardcopy format or in a stand-alone computer application. However, these options may be available, using a state's own computing capabilities, once test items are retrieved from the item bank. While it may be possible to create a unifying structure, or at minimum a portal, that makes access and use of the item bank and assessment clearinghouse tools easier for state CTE leaders, the two may not necessarily be housed on the same platform.

Assumption 3: Role of Derived Assessments and Accountability Reporting

Seldom is a test or the items for a test developed for the sole purpose of reporting information. It is assumed that the primary reason for developing a collaborative item bank is to provide a means to improve student attainment of knowledge and skills identified in the SCCI Career Clusters and Pathways knowledge and skill charts. While it is anticipated that the results from customized statewide tests derived from the test item bank will satisfy state reporting requirements for Perkins, its use likely will depend on the final nature, purpose, specifications, and reliability of the item bank and the procedures used to derive and deliver tests.

Assumption 4: Client Base

Item bank access will be limited to state officials with authority over CTE programs. Local or regional officials, instructors/teachers, and other state education leaders will not contract for, nor have direct access to, system resources, although these groups

may be involved in test specification design, test deployment, test administration, data review, and training. The decision over which items will be drawn from the item bank and how these items will be integrated into a test delivery system also will reside with state officials or their designee. These individuals will have sole responsibility for using the items in a way that is consistent with agreed-upon state assessment purposes and parameters.

Any effort to deploy a collaborative item bank will require that states be involved on a voluntary basis. Therefore, an additional assumption is that states choosing to participate will be afforded some opportunity to shape decisions for test purpose, item type, item population strategy, access controls, costs, item bank security, test administration, and reporting. It also is assumed states not participating in the development and deployment effort will have limited input into those decisions; however, they will have the opportunity to access the national test item bank when ready for use although they will be charged a fee or greater fee for access than those that initially participate and agree to the terms and conditions for test item use, as set forth by the governing entity of the test item bank.

Assumption 5: Standards for Item and Test Development

Regardless of the final purpose of the tests derived from the collaborative test item bank (which is a critical factor in many decisions related to item and test development), an effort of this scope and reach has great visibility, and as such, requires a greater level of rigor in development to withstand scrutiny than a state or locally developed effort. Recognized by experts and the courts in defense of valid tests, there is one conclusive set of standards for test development—*Standards for Educational and Psychological Testing*, 1999 edition (hereafter referred to as "Standards")—published jointly by the American Educational Research Association (AERA), the American Psychological Association (APA), and the National Council on Measurement in Education (NCME). It is assumed that users of the items developed and the tests derived from the collaborative item bank will consult this document to guide development and deployment decisions and that validity and reliability requirements also can be met through appropriate application of the Standards practices. There are many decisions concerning the design and use of the items and derived tests that will affect how the Standards are applied to the development and deployment process. It is impractical to identify all of those decisions in this report. Where appropriate, however, notes have been made about how the decisions presented in the Design Drivers section would influence the application of the Standards.

Section 3: Test Item Bank Design and Implementation Models

There are numerous, complex design features and options that will need to be resolved before a national test item bank can be constructed. The following section opens by presenting a set of proposed design drivers to guide item bank development. These are followed by a set of competing alternatives for item bank design, presented in tabular form, which summarizes how different implementation models meet the design criteria and their associated benefits and challenges. The section closes with a proposal for an implementation model for a test item bank and estimated costs of a two-year design and rollout.

Design Drivers

The following section summarizes key design drivers that must be considered in creating the item bank and identifies the most feasible options for adoption in the implementation model. The options identified and the responses to each were shaped in large part by a review of current research on testing, education measurement, and the design of test item banks in other subject areas. Feedback offered by state CTE leaders at various events in 2008 is also incorporated into this section.

Driver 1: Test Purpose

Although this report focuses on the design and feasibility of a national test item bank, ultimately, design decisions are contingent on the use of tests derived from data bank elements. As such, the process used to identify existing items, train writers to adapt existing or create new test questions, complete a psychometric analysis of items, secure and administer the data bank, and deliver items on a test will vary depending on the purpose of the test.

It is important to emphasize that one test cannot serve multiple purposes very well. For example, though the SAT may function well as an indicator of high school seniors' college preparedness, it is a poor test (and would likely be legally challenged) for assessing teacher knowledge as a means of certifying teachers in a school district. The implications for the proposed test item bank cannot be understated. The items in a national item bank cannot be used for multiple, varying purposes and still maintain

13

their reliability and validity. While it may be possible to use the items for very similar purposes, that decision will need to be made on a case-by-case basis by experts in the test development industry.

To clarify, in test development, there is an identified test user and test taker (along with other roles not considered here). It has been established that the assumed test user is the state CTE leadership. It can be assumed that the test taker will be a student participating in a CTE program (whether the test is for concentrators or any enrolled student). In simple terms, the greater the impacts on the test taker, the more rigorous are the requirements to develop and deploy a test. And as a general rule, as rigor increases, so does the cost of associated test development.

Making decisions about test purpose, though, should not be based on cost considerations or reporting needs. The purpose of a test should come from an identified need, which may include the following:

- Measuring student progress in a curriculum or course of study;
- Establishing a "passing" threshold for a student in an educational program;
- Providing instructors with data needed to make changes in instructional decisions;
- Supplying system leaders with data to evaluate the need for changes in policies, curriculum, or programs;
- Gauging the effectiveness of a teacher, program, school, and so on;
- Assuring that required job skills have been acquired and the test taker can be trusted to be competent in completing a job; or
- Supplying institutions with information needed to make admittance decisions.

Each identified need calls for a unique test purpose, which, in turn, has implications for test development and deployment. Table 4 below identifies six commonly mentioned purposes for the tests derived from a national item bank and reviews each in relation to their feasibility for providing accountability data as called for in federal Perkins legislation.

Table 4.	Purposes for tests derived from test item bank					
	Formative	**End-of-Course Summative (secondary)**	**End-of-Program of Study (POS) (Summative (secondary or post-secondary)**	**Alignment Indicator (secondary to post-secondary)**	**Program Effectiveness Indicator**	**Individual Certification**
Primary Purpose	Assist instructors with current feedback on student understanding.	Measure student acquisition of course knowledge and skills.	Measure student acquisition of POS knowledge and skills.	Provide tools to grant course credit and/or make placement decisions.	Measure effectiveness of an instructional program.	Assure public that individual has skills necessary to complete a job.
Effect on Test Taker	Little impact (very low stakes).	May be used to influence grades (low stakes).	May be used to determine graduation (high stakes) and/or grades (low stakes).	Impact on credit and placement in secondary/ postsecondary setting (high stakes).	None (very low stakes).	Great impact. Creates or denies job opportunities (very high stakes).
Impact on Implementation and Feasibility	Not feasible. Data collected not valid for accountability measures.	Probably feasible. May not meet accountability measures.	Feasible. May need to define scope of use to assure proper use of test scores.	Feasible. Requires greater rigor/costs in development.	Feasible. May not meet accountability measures.	Not feasible. Standards not reviewed to the level of rigor needed. Costs/process too great.

Two options emerge as most feasible and likely to be in alignment with a stated goal of assisting with accountability measures:

- Provide a summative assessment of the cluster and pathway knowledge and skills included in a Program of Study (low stakes); and

- Provide a credit-earning or course placement decision for students in a Program of Study for secondary to postsecondary transition (high stakes).

Of these two options, the former would likely require a less rigorous process for item creation, review, analysis, and test specification design, because the impact on the test taker is relatively low, compared to one involving a credit-granting or an advanced placement decision. It is also likely that the policies and procedures established by states to ensure the validity and reliability of high-stakes tests would place too many restrictions on the use of test items, making the test item bank less feasible without significant financial investment in item development and security.

With an item bank potentially as large as the one identified here, it can be helpful to think of increases in rigor and scope as exponential versus incremental. That is, the addition of just one hour of effort per item can translate to 5,000–10,000 hours' worth of work, depending upon the size and scope of the database. If, for example, in order to increase the rigor of the development process, expert counsel recommends that four, rather than three, subject matter experts review and document their sign-off to any revision to the test item bank, the additional 30 minutes for a fourth expert to review each item, assuming a 9,000 item base, would add 4,500 hours of work time to the overall project. Test users must therefore take into consideration the overall value—encompassing the effort, financial cost, benefits to student, and benefits to CTE field—of any changes intended to address an identified need.

It should be noted that policymakers may have used end-of-course or end-of-program tests for placement/alignment decisions in the past, but likely have done so at their own discretion. Such a decision is usually not supported by a test publisher if the test was not designed for such an explicit purpose. The publisher of a commercially available welding test, for example, probably does not support the use of test results to determine if a student receives college credit because the test was not designed or proven for such a purpose. If a student, in this example, legally challenged the use of his or her test results for credit granting, the test publisher would likely make it clear that the test was used outside its intended scope and that the decision to use the test for such a purpose was made by the policymaker. Similar situations could emerge with a test item bank if its main purpose is to provide a summative assessment of the cluster and pathway knowledge and skills included in a Program of Study, but some states also choose to use the item bank for placement purposes. States opting to do so would retain liability for such action.

Driver 2: Standards and Benchmarks for Item Alignment

The next most critical factor driving item bank design is defining test content. The ability to make inferences on the results of the test lies squarely on users' ability to defend the methods used to define the content, delineate measurable structures, and determine the amount of sampling adequate for the scope. Content definition is aided greatly when the standards, competencies, or tasks have been developed over time using a process that involves the judgment of qualified professionals with expertise in the domain. Generally, the performance elements of the SCCI Career Clusters and Pathways knowledge and skills charts provide this basis. The level and type of review required for these statements to be defensible, however, depends upon the pur-

pose of the test and its impact on the test user. Upon initial impression, the SCCI Charts would likely be defensible for use in end-of-program tests, but likely not for individual certification tests because certification tests usually require a job test analysis, the involvement of more subject matter experts, and a more documented development process. The options in Table 5 indicate the use of the SCCI Cluster and Pathway performance elements and/or the use of other defined standards. Using such a database-driven system opens the possibility of each state adding its "own" benchmarks or standards to the system.

17

Table 5. Possible uses of the SCCI Cluster and Pathway performance elements and other defined standards				
	SCCI Cluster and Pathway Performance Elements Only	**SCCI Cluster and Pathway Performance Elements + State-Added Benchmarks**	**Industry Standards (as agreed upon by participating state members)**	**New Set of Standards and Benchmarks Determined by Participating State Members**
Description	Item bank only houses items that align to the 3,000+ performance elements in the SCCI K&S charts. States select only those that apply to their curriculum.	Item bank houses items that align to the 3,000+ performance elements in the SCCI K&S charts and any state-added benchmarks. States can only access their own benchmarks. States select SCCI benchmarks and their own benchmarks to define content.	SCCI performance elements not utilized. Test owner(s) choose to use one or more collections of industry-defined standards instead. States select only those that apply to their curriculum.	SCCI performance elements not utilized. Test owner(s) choose to create a new set of benchmarks to define desired scope of content in the item bank.
Perceived Advantages	Work already done. Statements are current; revised in 2006 and 2007. Commonly agreed-upon set for all to use. Drives adoption of Career Clusters model.	Same as those listed to left. Additionally, allows states to devise tests that cover state-specific content in the same system.	Industry associations usually conduct more rigorous review of standards. Work already done.	Scope of content can be more closely aligned to existing curriculum and instruction.
Perceived Disadvantages	Scope of content in SCCI benchmarks may not align with current curriculum and instruction.	System becomes more difficult to manage. Requires extra state effort. Less uniformity in item alignment.	Requires consent on which standards to use. Cost to use copyrighted standards. Scope of standards may not align with current curriculum and instruction. Not all industries have identified standards.	Requires significant time, effort, and resources. Aligned to current curriculum and instruction not desired curriculum and instruction.
Impact on Implementation and Feasibility	Feasible. May require additional review depending on test purpose.	Feasible. May increase system costs to provide functionality.	Likely not feasible. Too many barriers to create tests that match value of existing industry certification tests.	Likely not feasible. There are great cost and effort requirements. Time to develop standards delays construction of items.

The two most feasible options, as listed below, use the SCCI performance elements as a starting place for states to select and define the content applicable to each state-identified program of study.

• Using SCCI performance elements only; and

• Using SCCI performance elements and allow states to add their own standards.

The task of defining the constructs for each cluster and pathway has been completed over the past decade and the definitions have undergone recent review and revision. Likewise, states have invested time and effort in developing and reviewing curriculum standards.

To the degree that the SCCI performance elements are of an appropriate "grain size" for item development is open to discussion. Grain size refers to the amount of content expressed in a statement. For example, "solve a variety of algebraic equations" is a much larger grain of content than say, "express equations using numbers and variables." The former would require demonstration of test results, compilation of homework assignments and quiz scores, etc. to measure a student's ability. The latter could likely be measured using test items that provide opportunities to demonstrate knowledge. In the 2007 revisions of the SCCI performance elements, a concerted effort was made to define the statements at a grain size that allowed for test items to be appropriate measures of assessment.

The degree to which a stated benchmark and a corresponding test item are actually aligned is an important consideration in item bank and test design. Any item added to the item bank would need to undergo a formal alignment analysis by a professional with content knowledge and training in test item construction to assure the legitimacy of the content and items incorporated in a test bank.[7] However, care must be taken when constructing an item bank. For example, though it would seemingly be easy to create a category of "transportation" items, into which test items about automotive service, collision repair, diesel technology, etc. could be added, a test derived from such a diverse item bank might provide little information about a student's actual transportation knowledge in relation to the field. The presence of clearly delineated constructs will be essential for the test item bank to have usefulness beyond serving as a repository of questions used for formative assessments in a classroom setting.

---

[7] For example, Norman Webb has developed a model for assessing the depth and scope of knowledge alignment that can be used to assure that items match the construct and level of cognition intended in a given benchmark.

Driver 3: Item Type

The type of item used in a test is driven, to an extent, by the purpose of the test and the definition of the content. To illustrate, though the knowledge of how to apply a skill can be measured in a multiple-choice question, such a question cannot actually assess students' ability to perform an indicated task. Performance items are required when the skill can only be demonstrated or when there is no evidence that knowledge of the application of a skill is an adequate substitute for the demonstration of the skill itself (i.e., a written test is as equally able to discern ability as a judged performance). The nature of many career skills, like those identified in the SCCI Cluster and Pathway charts, will likely require hands-on performance, which will introduce cost and administration issues.

Judging performance, however, is only necessary when the purpose of the test is to provide some reasonable assurance that the test taker possesses the skill required for certification or licensure. If the test purpose has little impact on the test taker, the requirement to demonstrate performance may be reduced, especially if the user can show evidence that the knowledge tested correlates to the skills identified in the content definition. As is often the case when test users are faced with the task of providing a large-scale test with limited resources and time, pragmatism and feasibility will influence assessment design. The use of constructed response items (fill-in-the-blank, essay questions, short-answer, etc.) is appealing, for example, but the cost of bringing hundreds or thousands of subject matter experts together to score student responses would quickly reach hundreds of thousands, if not millions, of dollars annually. This is the reason constructed item responses are used on such a limited basis in public education or are restricted to tests costing $100 or more per test taker. Large-scale testing efforts of math, reading, and writing by states for No Child Left Behind (NCLB) requirements rely almost exclusively on selected-response (multiple-choice, true-false, matching) items.

While item type is driven by purpose and content, it is also affected by the mode and means to deliver the test. Though a test delivered in paper-and-pencil mode cannot include items that use video media, such options are available in a computer-based testing format, which the item bank may accommodate. The ability to produce adequate numbers of test items in a desired time period may also play a role in determining which types of questions are most feasible. Four item types are presented as options in Table 6. More are available, especially with computer-based testing (i.e., drag and drop, hot spot, sequential ordering, etc.), though many have yet to be fully vetted in research. As such, some options likely present too great a risk for the intended scope and reach of the item bank. Additionally, some items require the in-

volvement of graphic designers and/or programmers, adding great cost to the item development process.

Table 6.	Test item type options			
	Multiple-Choice: Text-Only	**Multiple-Choice: Media Included**	**Multiple-Choice: Scenario-Based or Case Study**	**Expert Evaluated Performance Open-Ended/Essay Oral Interview Judged Work Observed Performance**
Description	A "typical" test item with a one-sentence stem and four choices.	Relies on a media item (image, video, or graph) to test knowledge.	Relies on information presented in a scenario or case study to test knowledge.	Constructed response tasks to measure complex constructs. Evaluation requires judgment of a subject expert.
Cost to Develop	Relatively low.	Varies greatly, but more than text only. Depends on type and amount of media, especially to license existing resources.	Varies greatly, but more than text only. Depends on effectiveness of item writer training.	Relatively very low as evaluation skill provided by reviewer.
Availability from Existing Sources	May be many items available from existing sources, but quality will be an issue.	Some items may be available from existing sources.	Likely few items available from existing sources.	May be many items available from existing source, but feasibility will be an issue.
Impact on Implementation and Feasibility	Most feasible to create many new, quality items and/or secure from other sources.	Feasible. Will require additional time and resources.	Feasible. Will require additional time and resources to develop.	Development is feasible. Scoring items on a large-scale basis likely not feasible.

Constructed response items scored by a subject area expert, while fairly easy to develop, are probably not a feasible option given the need to train multiple evaluators, over a broad range of content areas and geographic areas, to score statewide assessments. It can be argued that the bank should include such items, with states given the option of adopting them for use. There are two problems with this approach. First, the inclusion of performance items implies that they should be included in a derived test if the test is to be judged valuable. This is not the case. Numerous research articles support the use of multiple-choice questions for a wide range of testing purposes. Second, few states, if any, will have the financial resources available to score a large-scale performance test. Those states that do will likely have the resources in place to develop their own items and need not rely on the collaborative item bank as an item resource.

Therefore, the decision is rather what types of multiple-choice items are most appropriate and feasible. If the type of item is of particular importance to the test users, the

criteria will become a driver of item bank design. If, however, other design criteria are of greater importance—cost, item reusability, etc.—then the item type may be determined based upon other decisions.

Driver 4: Populating the Data Bank

Determining how the test item bank will be populated with sufficient items to be useful and feasible to test users is an important driver of system design and process requirements. While this design driver influences item bank implementation and deployment more than the test design, there are political, feasibility, and perception issues that will affect the adoption rate and use of the item bank. Feedback from potential state test users also indicates a desire to have some portion of the item bank functional within two years of its initiation, and to have all items be current, relevant, and aligned to the Cluster and Pathway performance elements.

This driver is presented with the assumption that there will be a cost to access finished items from the item bank. This cost could be incurred at the state or local level and may depend on a number of factors. The decisions around pricing structure and a revenue model that sustains the long-term operation of an item is not a topic of this report, but it is recognized that whatever organization manages the item bank, if it is deployed, will need to spend considerable time addressing these questions.

Each of the options presented in Table 7 is feasible if the condition merits, and each has distinct advantages and implications for test item bank implementation and deployment. It may also be possible that item bank users will be served best by some combination of the options to maximize the availability of items in some areas, while item development continues in areas where no questions currently exist.

Table 7.	Options for test item bank implementation and deployment				
	Create— Managed process using participant state subject matter experts to develop items and local programs to field-test items.	**Create— Contract with custom test development companies to develop and field-test items.**	**Create— Contract with CTE-related test publishing companies to develop and field-test new items.**	**Access— Access available state and local assessment items for use in the item bank.**	**Access— Access available industry and test company items for use in the item bank.**
Description	The managing organization uses instructors and industry representatives in each participating state to develop and review items. Psychometric services would be contracted to lead field-test design and analysis.	The managing organization contracts with one or more companies and organizations that write test items on behalf of others. These groups would work with available subject matter experts (SMEs) in each state to assist in development and field-testing.	The managing organization contracts with one or more test publishers that publish tests in related areas. The groups work with their own SMEs to write and field-test items.	The managing organization works with participating states to gain access to and right to use existing test items at no cost. Professionals would be contracted to ensure alignment to SCCI standards.	The managing organization works with associations and companies to license the use of items in the item bank for some type of revenue-sharing agreement. Items are already tested and aligned.
Cost Considerations	Depending on the efficiency of the managing organization, simple MC items may cost $200–$400 each. Cost could be offset depending on in-kind resources provided by participating states.	Depending on the quantity and timeline, simple MC items may cost $300–$600 each.	Depending on the quantity and timeline, simple MC items may cost $200–$600 each. Initial cost may be offset if test publisher is willing to consider revenue-sharing once items are in use.	Initially, no cost except for management time to acquire and assemble available items. Additional costs to review and align items may drive total costs to $25–$50 each.	Perhaps no initial cost, rather a revenue- sharing arrangement where 50–80% of the revenue for each item flows back to licenser.
Oversight/ Quality Assurance	Provided by external, contracted entity.	Required as part of contract with item development companies.	Required as part of contract with test publishing companies.	Requires review of process used to develop and field-test item in each state. May require external entity to assist.	Likely complete, but should include a review of the process used for each.
Advantages	Allows managing organization to target areas for development. Provides a "new" set of items available to the field. Builds assessment expertise of field.	Assures professional service, quality items, and no conflict of interest by developing items for item bank.	Assumes quality items and professional service. Requires less effort to coordinate SMEs.	Low initial cost. Likely the option that could deploy the most quickly.	Low initial cost. High-quality, vetted items aligned to industry standards.
Impact on Implementation and Feasibility	Feasible. Will require the greatest effort for managing organization.	Feasible. Will require greater financial resources. May limit number of topics that can be addressed.	Feasible. Will require greatest political skills for managing the organization. May require greater costs.	Feasible. Will require extensive coordination to assimilate items from so many different sources.	Feasible. Will require greatest negotiating skill for managing the organization.

An unanswered question is whether test items currently held by states, test publishers, and industry associations are suitable for the SCCI Cluster and Pathway statements. For many years, the focus of occupational testing has been on job-specific tasks; however, Cluster and Pathway statements are designed to address knowledge and skills that cut across many jobs. As such, it may be that the content tested in most existing state and commercial tests will not be aligned with the content defined in the SCCI statements. If further review indicates that there is limited alignment, options for developing new items probably will be most feasible.

There are intangible benefits of the first option—using teachers in each state to develop new test items. First, there is some greater amount of field "buy-in" when peers create items. Secondly, while many teachers will need extensive training and coaching to become adequate item writers, an additional benefit is that acquiring these skills and assessment knowledge could have numerous positive, residual effects in the CTE field.

Driver 5: Test Delivery and Administration

How a test is to be delivered and administered influences not only the design of the test but also the measures taken to increase the validity and reliability of the test and/or test items. The options listed in Table 8 are generalized representations of the myriad of sub-options included in each option. The intent of this section is to introduce four starting places for further discussion.

All of the options discussed are computer-based and database-driven, because paper-and-pencil management of the item bank and test delivery system is not feasible given the scale requirements of the proposed system. Due to the specific purposes associated with the system, it may be worthwhile to develop a new, customized computer application to manage the item bank, test delivery, and test administration. However, the time and financial costs of designing software may make this an unattractive option. As technology advances, it will become increasingly difficult to garner a consistent meaning for the term "computer-based." This report does not attempt to define all of the methods and tools available, but rather to provide a general overview of major options.

Table 8. Test delivery and administration system options				
	Item bank only—no test delivery system	Item bank with "off-the-shelf" test delivery and administration system	Item bank with "off-the-shelf + upgrades" test delivery and administration system	Item bank with custom-built test delivery and administration system
Description	System only manages the input, storage, and export of test items. Tests cannot be constructed or delivered on the system. States select items for use and receive data file. States would have to deliver tests using own system/method.	System manages the input, storage, and export of test items. States can derive and deliver tests, create and manage users, and generate reports. System would be used "as-is" like purchasing software for your computer—no custom functionality.	System manages the input, storage, and export of test items. States can derive and deliver tests, create and manage users, and generate reports. An existing "as-is" system would be modified to meet the unique needs of test users by adding functionality and/or new applications.	The managing organization would contract a Web application company to custom-build a system to manage the input, storage, and export of test items; to derive and deliver tests; to create and manage users; and to generate reports.
Cost Considerations	Least expensive option. Likely could build/lease for three years for $50k–$250k.	Cost can vary depending on features built into system. Likely build/lease for three years for $100k–$500k. Additional system use/hosting fees vary with use.	Cost can vary based on the cost of initial package and number of additional upgrades. Likely build/lease for three years at $250k–$1.5mil. Additional system use/hosting fees vary with use.	Likely most expensive option. Could cost $500k–$3mil to build. Would need to prepare for maintenance and "patch" costs. Additional system use/hosting fees vary with use.
Impact on Implementation and Feasibility	Feasible. This option, though, creates major challenges for states without test delivery systems currently in place. May require involvement of third-party test publishers/providers.	Feasible. Likely the quickest option, too. Likely will not be able to find a system that meets the requested functions of all participating test users, though. Many states may require use of own test delivery system; vendor rendering this option is not needed.	Feasible if funds available. May allow for additional perceived "essential" functions and applications in the system (i.e., automatic generation of accountability measurement reports). Many states may require use of own test delivery system; vendor rendering this option is not needed.	Likely not feasible. A development effort of this magnitude is challenging even for experienced organizations. May not be best use of funds. Many states may require use of own test delivery system; vendor rendering this option is not needed.

At the other end of functionality, a system that only manages the input, storage, and export of test items greatly reduces time and costs, but also fails to address the test delivery needs of end users. This, however, may not be an issue for states that have already invested in a test delivery system for statewide academic tests. In many cases, a statewide CTE test would be delivered using existing state assessment platforms so

24

that secondary student data could be readily collected. An intangible benefit to delivering tests through state systems is that it could reinforce the alignment of CTE with academic instruction, while indirectly improving the potential for collaboration and cooperation between CTE and academic instructors.

The functionality of the chosen system will drive decisions made about the process used to develop test items. Item security issues, functionality available to assist in item creation, review, and testing, and the tools available to assist in test proctoring and delivery will all drive the design of processes required to ensure validity of the test that is constructed.

Item Bank Implementation Model

This section examines the most feasible and potentially successful implementation model given the information at hand. This proposal is intended to provide sufficient information for a managing organization to begin to draft an RFP or for an interested vendor to construct a business plan; however, neither would be viable without further research, vetting, and input from stakeholders, potential users, and members of the CTE community. It should be noted that a variety of implementation models could provide reliable tests that allow for valid use of test scores. While no one implementation model is inherently more valid and reliable than any other, it is fair to say that the selection of a final model will require thoughtful consideration, professional judgment, and expert guidance to provide a quality, meaningful, and useful testing experience for students.

The proposed implementation model would accomplish the following:

- Define the test purpose as primarily used to derive summative end-of-program tests that indicate student achievement.

- Use SCCI Cluster and Pathway performance elements as a primary source of content definition. Allow for addition of state-identified benchmarks to support flexibility and future extension.

- Rely upon "typical" multiple-choice test items to populate the item bank, but allow a small number of items to include images, charts, or scenarios in the item stem. Acknowledge that derived tests will be unable to directly measure some performance skills listed in the SCCI Cluster and Pathway performance elements.

- Populate the item bank with new items written using state-identified resources under the direction of the managing organization and/or by one or more test developers.

- Use or contract to build an item bank system that will easily export data to a variety of test delivery systems used by states and state-contracted vendors to assemble, deliver, and manage tests.

- Develop the training materials and programs needed for state CTE leaders to facilitate the test development activities needed to derive tests from the item bank and to work with the test delivery system of their choice or of their state's choice. Recognize that many states will enact their own policies and procedures, but will make them available to all as a suggested resource. Training topics include, but are not limited to, the following:

  - Selecting and retrieving items from the item bank;
  - Testing item security issues;
  - Adding standards and new test items to the item bank;
  - Creating a test specification/blueprint;
  - Coordinating with state testing coordinators; and
  - Developing RFPs for statewide test delivery.

This proposed model supports a collaborative approach to building a test item bank that meets many of the needs of interested states, while providing the most feasible approach given the financial conditions, expectations for test validity and reliability, and availability of existing resources.

Estimated Costs of Proposed Implementation Model

Estimating the cost of enacting the proposed implementation model is complicated by a lack of detailed specifications and deployment options. Nonetheless, providing cost estimates can help launch a discussion about the possible funding sources for development, the scope and length of development, and the type of revenue model needed to support the long-term sustainability of the collaborative item bank.

The costs identified in Table 9 were devised from research presented in current test industry publications, experience of professionals engaged in test development, and education measurement research. As with most other purchases, the cost per unit will

decrease as the number of units purchased increases due to economies of scale. The annual cost of leasing the technology system, for example, will likely be less expensive if it is part of a five-year lease rather than a two-year lease. A test publisher may be willing to lower the per item cost of development if contracted to develop 1,000 items instead of 100 items. This is one reason the cost estimate is presented in ranges. Another is that as it takes longer to develop the item bank, those costs associated with overhead or ongoing expenses increase.

Table 9. Estimated costs of the item bank implementation model		
Activity	Low estimate	High estimate
Requirements Gathering and RFP Development Use professional(s) to assist in developing a definition for the technical requirements of the system, process requirements for item development training, review and analysis, and initial project plans.	$15,000	$30,000
Item Development Includes item creation @$200–$350 for each item, item review, item editing, item analysis, and item alignment. Note that if a minimum of 3 items for each of the approximately 3,000 performance elements are developed, then 9,000 items are needed to populate the item bank.	1,800,000	3,150,000
Item Bank System Lease or build a computer-based online system to input, store, and export items and manage users. Assumes two years of development and one year of initial delivery. Design requirements to ensure applications provide desired functionality in system.	50,000	250,000
Coordination, Communication, and Project Management One or more entities will be required to commit human and other resources to manage the two-year project development plan, invite and involve participating states, engage development partners, provide communications to stakeholders, etc. Likely requires 0.5 to 3.0 FTE to complete tasks. Additional costs include travel, technical assistance, meeting expenses, overhead, materials, and communication efforts.	200,000	600,000
Training Materials and Programs for Test Users (State CTE Leaders) The state CTE leader will be responsible for test development functions critical in assuring proper use of items. If items are used improperly, if the security of the item bank is compromised, or if the item bank seems too cumbersome for practical use, then the initial efforts and investment in the collaborative item bank may be diminished. Training could include online learning modules, Webinars, workshops at national events, technical support services, manuals, connection to reliable training sources, and sample tools/templates.	25,000	100,000
Technical Assistance and Oversight Educational measurement experts will provide guidance in design, feedback on suggestions by developers and users, and assurance that the Standards have been addressed. These experts would have no direct involvement in training item writers, writing items, or reviewing or analyzing items. This will assure an unbiased review of the process.	20,000	50,000
Total Cost* Figure is a rough estimate given current information and assumptions. The number of items developed and the cost per item to develop them will be the primary driver of costs. Further efforts will be required by the managing organization to plan for costs involved with the development of the collaborative item bank upon further definition of system and process requirements.	2,110,000	4,200,000

* Note that if 30 states were to access the item bank, the estimated average first-year development and two-year maintenance cost would range from roughly $23,500 to $47,000 per state.

Section 4: Clearinghouse Design and Implementation Models

Survey results indicate that, in addition to building their own custom set of assessments using a national test item bank, states also wish to use existing national credentialing and certification or state licensing exams to assess the technical skill attainment of CTE concentrators for some instructional programs. This option is relatively more popular among educators at the postsecondary level, where students are often prepared for workforce entry into a more discrete occupational field. This section reviews design criteria that can be used to structure the assessment clearinghouse, identifies the type of assessments that could be incorporated into the database, and arrays technical considerations that will affect system operation. The section closes with a cost estimate for system development and maintenance.

Design Criteria

The assessment clearinghouse will be most effective if it (1) contains validated state- or third-party assessments that are used by employers to document individuals' technical skill holdings; (2) provides detailed information about each assessment to enable states to make informed decisions as to its applicability; and (3) stores this information in a format that is accessible and easy to navigate.

Types of Assessments

Nearly all states have adopted the SCCI Career Clusters and Career Pathways model to classify CTE programs offered in high schools and postsecondary institutions. Because states have flexibility in defining coursework sequences for each pathway within a career cluster, and for discrete occupations within a given pathway, it is likely that no one technical skill assessment will apply across all states. To support states in selecting from existing technical skill assessments, OVAE has developed crosswalks that link CTE coursework within career pathways (defined using CIP codes) with all associated occupations for which a student may be prepared to enter, as well as for specific occupations within a pathway area (identified using SOC and O*NET codes).

The proposed assessment clearinghouse would enable states to connect SCCI-defined pathways and standards with existing state- and industry-validated occupational as-

sessments, providing states a choice of whether to use these existing instruments or to replicate exam development using the test item bank. To accommodate state needs, the clearinghouse will need to contain a broad range of industry-recognized, validated assessments. The following types of assessments should be considered for clearinghouse use, with the expectation that not all of them will meet the quality standards established for clearinghouse inclusion.

Industry Credentialing and Certification Exams—used to recognize an individual's attainment of validated skills, these would likely comprise the bulk of exams included in the assessment clearinghouse. Several states and organizations—including Texas, Virginia, Oklahoma, South Carolina, and the VTECS consortium—have catalogued existing industry credentialing and certification exams, and it may be possible to use information contained within these lists to jump-start clearinghouse design.

Occupational Licensing Exams—used to certify that an individual within a state is qualified to engage in a given trade or profession. Typically, each state develops its own criteria for licensing, meaning that there may be little immediate value of populating the clearinghouse with information on these exams. As an initial activity, the clearinghouse might contain a tally of occupations that have licensure requirements attached to them; an expanded feature might be to provide information about licensure requirements as they relate to industry certifications.

Nationally Developed Third-Party Assessments—include *Off-the-Shelf Technical Skill Assessments* and *Customized Assessments* used to assess whether test takers have achieved a set of generic or more specifically defined technical skills within a discrete occupational area. A number of public and private educational agencies specializing in CTE have developed technical skill assessments that may be candidates for clearinghouse inclusion, including the National Career Assessments Services, Inc., the National Occupational Competency Institute (NOCTI), Skills USA, the Vocational Technical Consortium of States (VTECS), and WorkKeys, among others. Other assessments developed by public and private proprietary firms would also be eligible for database inclusion.

State-Developed Technical Skill Assessments—created almost exclusively at the secondary level, these include end-of-course and end-of-programs assessments used to assess students' skill attainment. Examples of states with such systems include Arkansas, North Carolina, Oklahoma, Utah, and West Virginia. Because state-developed assessments are designed to be aligned with state-established standards, these exams may be given relatively lower priority for clearinghouse consideration.

Information About Each Assessment

Ideally, the assessment clearinghouse will provide detailed, consistent information about each assessment to permit state administrators to compare and contrast testing options. It is anticipated that the database manager would consult with test developers to obtain detailed information on the following fields, which would be edited for consistency and relevancy:

- Name of testing organization
- Skills addressed
- Occupations covered by exam
- Educational level(s) for which suited
- Recommended educational prerequisites
- Test type (paper, online, performance)
- Scoring criteria
- Process/frequency for updating
- Test administration timelines
- Assessment cost
- Usage: numbers, geographic coverage, etc.
- Test supplier contact information

Those who develop the database will also need to address the availability of test-taker outcomes for state use. Currently, states have limited access to third-party assessment data, in part due to privacy concerns, which limit the sharing of personally identifiable information outside the testing agency. While it is prohibitive to design the clearinghouse to incorporate a score retrieval system, the database would, at a minimum, include a description of the policies and procedures governing state access to test-taker data.

Database Considerations

The assessment clearinghouse could be configured to provide users with some flexibility in using the system. Issues for consideration include the following:

Platforms and Capacity—At the minimum, the assessment clearinghouse needs to be Web-based and able to accommodate commonly accepted Web browsers. Attached files should be in formats that use common off-the-shelf software not requiring special licenses, ideally, with most information provided in .pdf and .html format. The site

should be tested using normal broadband-width access speeds. If resources permit, more effort can be put into Web site design, appeal, and audio-video capabilities.

Reporting Features—The website should provide at least a series of static files that contain an organizing structure whereby each assessment, and all of the information about each assessment, is placed in its appropriate category in the Career Clusters and Career Pathways occupational crosswalk. This can be expanded to include (1) drop-down menus and search capabilities to find and select assessments from a number of perspectives; (2) relational databases that cross reference the assessments within the structure and allow for cross-comparisons; and (3) the ability to create custom-printed and exportable reports using a common interface such as Microsoft Access.

Accessing Information—The information needs to be accessible to state CTE administrators at both the secondary and postsecondary levels, with secure password access that requires regular updating. In an expanded version, access can be granted to administrators at local school districts, community colleges, and possibly to instructors, though doing so would likely compromise system security and dramatically increase maintenance and support costs.

Site Maintenance—The manager of the project (or hired third-party contractor) will need to maintain the site and respond to inquiries from participants, suppliers, and other interested parties. These roles can be shared with the test item bank for greater operational efficiency.

Updating Assessments—A base version would have a protracted update schedule to save on operating costs. Most of the clearinghouse expenses would be up front, allowing for later-stage project resources to be used to identify additional assessments and revise existing information. An expanded version would require a regular update and an aggressive research timeline, so every assessment would need to be reviewed for changes at least once annually.

Clearinghouse Implementation Model

The assessment clearinghouse is intended as a resource for states seeking to identify assessments that are aligned with CTE programs offered in high school and postsecondary institutions. Model components are based on feedback provided by state administrators on informal surveys and feedback provided following presentations at national conferences. While the system developer will determine the ultimate format

and system features of the clearinghouse, it is anticipated that the implementation model would accomplish the following:

- Provide online access to a searchable, easily navigable database of industry-recognized, valid, and reliable assessments.

- Cross reference assessments with the SCCI Cluster and Pathway performance elements.

- Provide detailed, periodically updated information on test attributes to enable users to select from multiple options.

- Develop the training materials and programs needed for state CTE leaders to retrieve information from the clearinghouse. Training topics include, but are not limited to, the following:

  - Selecting and retrieving items from the clearinghouse;
  - Identifying database security issues; and
  - Using assessment information to identify appropriate exams for state use.

Estimated Costs of Proposed Implementation Model

It is worth noting that the creation of an assessment clearinghouse is not a new proposal. In June 2007, a private company approached individual states with a proposal to develop, operate, and maintain a comprehensive database of approximately 1,000 industry skill certifications. Initial development cost estimates ranged from $50,000 to $75,000 per state (depending upon state population size), with costs for each of the two remaining years estimated at $15,000 per state. States choosing to join the project after the early adopter period closed were to be offered consortium membership at higher rates. Assuming the proposed model were adopted by 20 states, start-up costs would exceed $1 million, with annual ongoing costs of $300,000 per year to maintain the system.

Cost estimates of the assessment clearinghouse shown in Table 10 assume that the system will build upon previous work conducted by the Next Step Working Group technical skill assessment taskforce and contributions by state secondary and postsecondary agencies. Given that considerable work has already been done to identify ex-

isting assessments, the cost of producing such a system will likely be at the lower end for project vendors. The content found in the assessment clearinghouse would be equally comprehensive in both the low and high estimates. The high estimate would mainly enhance database considerations, primarily expanded reporting features, more frequent updates, and better support and training.

Table 10. Estimated costs of the assessment clearinghouse implementation model		
Activity	Low estimate	High estimate
Requirements Gathering and RFP Development	$15,000	$30,000
Use professional(s) to assist in developing a definition for the technical requirements of the system, components of assessment profiles, and review and analysis and initial project plans.		
Assessment Cataloguing	100,000	200,000
Assumes compilation of background information for approximately 500 assessments, at four hours per assessment, which translates to 1.0 FTE at $100,000.		
Test Management System	100,000	300,000
Lease of a computer-based online system to store assessments and to generate reports for one year of development and one year of initial delivery.		
Coordination, Communication, and Project Management	200,000	400,000
One or more entities will be required to commit human and other resources to manage the one-year project development plan, update the test database, and invite and involve participating states, engage development partners, provide communications to stakeholders, etc. Likely requires 1.0 FTE to complete tasks. Additional costs include travel, technical assistance, meeting expenses, overhead, materials, and communication efforts.		
Total Cost*	415,000	930,000
Figure is a rough estimate given current information and assumptions. The number of tests profiled and the cost per item to develop them will be the primary driver of costs.		

* Note that while the cost estimate may seem high, the actual cost over a three-year period would be quite reasonable when computed on a per-state basis. For example, if 30 states were to access the assessment clearinghouse, the estimated average first-year development and two-year maintenance cost would range from roughly $5,000 to $10,500 per state.

Conclusion

The technical skill test item bank and assessment clearinghouse models proposed within this report are intended to support states in assessing the skill attainment of students participating in CTE coursework and programs. The proposed systems are both conceptually and technically feasible, given sufficient fiscal resources and development lead times. It is anticipated that the design criteria and implementation models proposed in this report can serve as a starting point for developing an RFP that would stipulate the design features and operational requirements for each component of the assessment system.

Clarifying system ownership and securing start-up funding are likely the two most critical impediments to system creation. Although development and management of the test item bank and assessment clearinghouse need not reside with a single organization, it is crucial that the operator of each is seen as a credible agency that can command the allegiance of all states participating in the project. While it is assumed that the federal government will not take responsibility for financing or for providing project oversight, it is anticipated that federal staff will provide continuing guidance to system developers to ensure that the testing system will support states in responding to Perkins accountability requirements. It may also be advisable to create an oversight board, composed of state representatives and testing experts, to keep the oversight focused, on-track, and fiscally responsible.

Test item bank ownership is complicated because there are many public and private organizations that have a financial stake in administering CTE assessments. Agencies such as the SCCI, which developed the Career Clusters and Pathways model and identified the knowledge and skill statements within pathways, also have a professional stake in system design. To reduce competitive pressures that could lead to the creation of multiple, overlapping assessment systems, test developers and representatives of state CTE associations should be consulted to explore the options for a collaborative approach to system creation. For example, it may be possible for two or more test developers to collaborate in the design of the test item bank and/or assessment clearinghouse, with different agencies taking responsibility for different assessment system features.

Securing implementation funding will also be difficult given the uncertainty associated with system use and profitability. Potential vendors may be hesitant to invest

large sums for upfront development in the expectation that states will sign up to participate once the system is created. State education agencies, in turn, may be unwilling (and unable, given current budget shortfalls) to invest substantial resources or staff time in a collaborative effort to build a national assessment system that someday may meet their assessment needs.[8] Concerns could potentially be reduced if participating states were made co-owners of the assessment system or if system development were underwritten, in full or in part, by a third-party agency.

Suggested Future Action

Creation of a nationwide technical skill assessment system will require securing buy-in on system purposes from a wide range of stakeholders, including the federal government, state secondary and postsecondary education agencies, and CTE professional associations. Consultations with test developers and employer groups will also be necessary to ensure that the proposed system is conceptually, technically, and economically feasible. This report proposes developing a test item bank and assessment clearinghouse to meet identified state needs. While the design criteria and implementation models contained within this report appear viable, additional consultation is needed to review and, where necessary, overhaul or refine proposed system features.

To initiate this process, task force members recommend that OVAE use existing project resources to convene a one-day meeting of assessment experts—including testing coordinators, state secondary and postsecondary administrators, and representatives of CTE professional associations—to review recommendations contained within this feasibility report. Based on meeting conversations, it may be possible to secure agreement on a work plan for system development, including identifying a final, agreed-upon set of criteria for configuring an assessment system and the likely fiscal requirements for initiating development. It is anticipated that the product of this work plan will include an RFP that can be circulated among potential system developers, and shared with potential outside funders, to drive system creation.

---

[8] While Perkins accountability requirements may initially have helped promote state interest in creating a national assessment system, the recent approval of states' five-year Perkins plans may dampen enthusiasm in states whose technical skill attainment measures have been accepted. States' interests will likely increase if OVAE holds states accountable for increasing the number of students tested using industry-recognized, valid, and reliable assessments. Interest in workforce development issues and the continued refinement of the SCCI Career Clusters and Pathways model may also help to maintain momentum for creating national technical skill assessments.

Appendix

Results of State Survey

Purpose of Survey

The purpose of the survey was to gather information and input from all states on their interest in, use of, and support for the proposed national test item bank and assessment clearinghouse. Questions were designed so that results could be aggregated, in order to determine the number of states that might ultimately participate in the project and how their most important needs might influence the design and development of the project. Each state was asked a number of questions concerning these four topics:

- Current status of technical skill assessments—whether the state has a system in place at either the secondary or postsecondary level or both, and if so, its major characteristics.

- General interest in a national system—based on the state's view on whether to proceed nationally with this particular vision, what skills it intends to assess, and what standards it intends to use as a basis for assessments.

- Expected degree of participation in a national system—focusing on the usefulness of various features of the proposed item bank and clearinghouse.

- Support for a national system—in terms of start-up funding, operational funding, and supplying information.

Methodology

Survey Instrument

The survey was developed by the Technical Assessment Master Assessment Plan Task Force, based on feedback the group received after announcing and presenting its ideas to states at a Webinar on February 11, 2008. Specific questions were developed around each of the four topical areas listed above. Most questions were composed with closed-ended, pre-determined categorical responses in order to expedite tabula-

tion and analysis and to assure aggregation. The survey was converted into a format suitable for Survey Monkey, so the collection could occur online.

Collection Process

All states were invited to participate in the survey through a log-in ID and password provided to the State Director of Career and Technical Education. The respondent did not have to be this director, per se, but did have to identify himself or herself. Most states responded once, though they were allowed to have two responses if they chose, one each for secondary and postsecondary. Three states opted for this dual reply. All respondents were asked if their replies applied to both secondary and postsecondary, or just one of the two. The collection window was open for two weeks in March 2008.

Survey Tabulation

Responses were collected into two Excel databases, one that contained the full text for each cell and one that contained numeric assignments of each reply into a predetermined code. The file was edited and cleaned to assure completeness, and states were segregated by secondary and postsecondary responses as needed for proper accounting. Numeric responses were aggregated and/or averaged as appropriate for each question and were tabulated into a master file containing all the questions and the number of responses for each categorical option. Tables A1 through A9 summarize these aggregates into a more presentable format, but maintain all of the potential categories that could have been answered.

Tables A10, A11, and A12 combine information from the survey with other data gathered from states (both prior to and after the survey itself), in order to highlight the interrelationship among states in terms of their interest in the project, the experience states have with technical skill assessments, and whether they are using assessments to meet newly established Perkins reporting requirements. This additional information stemmed from data gathered from states in May 2007 in preparation for the Data Quality Institute and from approaches submitted by states in May 2008 in their final Perkins State Plan. These additional sources of information provide a way to analyze the proposed project from the perspective of the 23 states that did not complete the entire survey, along with the 31 states that did.

Analytical Methods

The method for analyzing these data is exclusively through the use of descriptive statistics. Absent any compelling research questions and given the small number of responses (limited by the number of states), no attempt has been made here to explore

correlations between factors or to test the differences among groups of states. Descriptive statistics are suitable for the purposes of this project, which is mainly to ascertain the aggregate level of interest in various project features.

Summary Tables

Current State Status

Table A1. State survey responses			
Responses	Number of states	Percent of respondents (n = 37)	Percent of universe (n = 54)
All responses	**37**	**100**	**69**
Joint secondary and postsecondary	24	71	44
Secondary only	5	14	9
Postsecondary only	5	14	9
Separate responses*	3	8	6
Did not participate in survey	*17*	—	*31*

— Not available/applicable.

* The three states with separate responses for secondary and postsecondary were merged into a combined response for each of the three states, with answers averaged for each response for those questions that were not clearly aimed at secondary or postsecondary status, interest, and participation.

Interest in National Technical Skill Assessment System

Type of state respondent	Number of states interested in using resources from national item bank or clearinghouse	Number of states **not** interested in using resources from national item bank or clearinghouse	Number of states interested in pursuing national item bank (with degree of interest)	Number of states interested in pursuing national assessment clearinghouse (with degree of interest)
Table A2. Overall interest in technical assessment system and its two main components				
All states	**31**	**6**		
Both secondary and postsecondary	21	3		
Secondary only	4	1		
Postsecondary only	3	2		
Separate replies	3	0		
All interested states			**31**	**31**
Strongly agree			15	21
Somewhat agree			15	10
Somewhat disagree			1	0
Strongly disagree			0	0
Not interested			*6*	*6*

NOTE: Only interested states were asked the follow-up question as to their degree of interest in the two project components and all subsequent questions summarized in these tables.

Reasons Given for Lack of Interest:

• Still uncertain as to how this will address our state competencies. Testing is an expensive endeavor.

• Need to know more about what levels of participation are offered and need to talk to the various segments of our college faculty about their interest.

• We are using industry-recognized assessments and credentials. Unless the assessments were valued by industry, we would see little benefit to them beyond using them as formative measures.

• Cost and relevancy to our local community college programs that include the whole occupation and not just a small segment.

• An item bank will not solve the issue of technical skill assessment.

• Plan to develop assessments for those degree programs that do not have a national/state licensure exam and/or a capstone course.

Table A3.   Level of skills most interested in assessing

Skills	Number of states	
	Secondary	Postsecondary
General work readiness or employability (across all clusters)	12	10
Foundation-level skills common to a cluster area	15	7
Pathway-level skills within a cluster area	19	15
Specific occupational skills within a pathway	19	28
Other	5	2
Not sure*	6	6
All interested states	**31**	**31**

* Some of the "Not sure" replies were from respondents from secondary or postsecondary institutions who could not place an answer in categories that pertained to the other type.

Table A4.   Standards most interested in using as the basis for assessments

Standards	Number of states	
	Secondary	Postsecondary
Cluster and/or pathway, based on knowledge and skills statements identified by States' Career Clusters Institute	12	7
Cluster and/or pathway, established by own state	11	5
Employability and general work readiness, identified by professional and/or trade associations	13	10
Occupationally specific, developed by the state	6	10
Occupationally specific, developed by third-party professionals and/or trade associations	19	22
Other	2	3
Not sure*	5	5
All interested states	**31**	**31**

* Some of the "Not sure" replies were from respondents from secondary or postsecondary institutions who could not place an answer in categories that pertained to the other type.

Participation in National Item Bank and Clearinghouse

Table A5.   States' likelihood of using the item bank and clearinghouse for various activities

Activity	Number of states responding					
	Extremely likely	Very likely	Somewhat likely	Not very likely	Not at all likely	Do not know or not applicable
Creation of standard secondary assessments based on SCCI knowledge and skills statements	4	8	8	3	1	7
Creation of standard postsecondary assessments based on SCCI knowledge and skills statements	1	4	11	7	2	6
Creation of customized statewide secondary assessments	6	9	6	5	0	5
Creation of customized statewide postsecondary assessments	2	3	10	11	0	4
Access test information in clearinghouse to identify secondary assessments	9	11	6	0	0	4
Access test information in clearinghouse to identify postsecondary assessments	4	9	13	2	0	3

NOTE: Based on replies from the 31 states interested in using resources from national test item bank and assessment clearinghouse.

Table A6.   Important uses of the item bank

Use	Number of states responding					
	Extremely important	Very important	Somewhat important	Not very important	Not at all important	Do not know or not applicable
Create customized statewide assessments for secondary	9	13	3	2	0	4
Create customized statewide assessments for postsecondary	3	12	7	5	0	4
Allow access to states to choose items and create own assessments	7	12	8	2	1	1
Allow access to school districts to choose items and create own assessments	0	7	9	8	4	3
Allow access to faculty and teachers to choose items and create own assessments	0	6	10	9	5	1
Provide online delivery / administration	12	10	5	0	0	4
Collect and aggregate data for Perkins reporting	16	7	7	0	0	1

NOTE: Based on replies from the 31 states interested in using resources from national test item bank and assessment clearinghouse.

Support for National Item Bank and Clearinghouse

Table A7.   Likelihood of providing support for the item bank

Area	Extremely likely	Very likely	Somewhat likely	Not very likely	Not at all likely	Do not know or not applicable
	Number of states responding					
Submit questions to the item bank	2	6	9	10	1	3
Provide start-up funding to seed the item bank	1	1	8	8	6	7
Provide annual funding to access the item bank	1	3	11	6	4	6
Other	(*)	(*)	(*)	(*)	(*)	(*)

* Six comments provided.

Table A8.   Amount of support willing to provide for item bank and clearinghouse

Amount available	Number of states for	
	Start-up development	Annual ongoing
Unable to provide resources	3	2
Less than $5,000	2	2
$5,000–$10,000	1	2
$10,001–$20,000	2	2
$20,001–$30,000	0	1
$30,001–$40,000	0	0
$40,001–$50,000	1	0
More than $50,000	0	0
Total with dollar amounts specified	9	9
Unable to specify a dollar amount	22	22

Table A9.   Amount currently being spent on collecting technical skill assessment data for Perkins reporting

Amount	Number of states for	
	Out-of-pocket (state funds plus Perkins funds)	State and local staff time
Less than $10,000	4	5
$10,001–$35,000	5	1
$5,000–$10,000	0	3
$10,001–$20,000	2	1
Unable to estimate expenditure	20	21

Interest in All States by Experience, Approach, and Size

Table A10. Experience with technical skill assessments and interest in project		
Interest in project	Experienced (all but one at secondary level only)	Not experienced
Interested in project	4	27
Interest at secondary level only	2	4
Interest at postsecondary level only	1	1
Interest at both levels	1	22
Not interested in project	8	15

Table A11. Perkins measurement approaches and interest in project				
Measurement approach to be used for technical skill attainment core indicators	All states interested in project	States experienced and not interested	States not experienced and not interested	Total using approach
Secondary				
Assessment: national or state approved	25	8	11	44
Assessment combined with another approach	1	0	0	1
Completion (program, course, or percent of standards)	1	0	4	5
GPA	4	0	0	4
Postsecondary				
Assessment: national or state approved	20	4	9	33
Assessment combined with another approach	5	2	0	7
Completion (program, course, or degree)	1	0	3	4
GPA	5	2	3	10
Total in Category	31	8	15	54

Table A12. Interest in project by size of state			
Interest in project	Large states	Medium states	Small states
Interested in project	10	9	12
Interest at secondary level only	4	1	0
Interest at postsecondary level only	0	2	1
Interest at both levels	6	6	11
Not interested in project	7	9	7
Total in category	17	18	19

NOTE: Based on 2007 total population, with cutoffs of 6 million for large-sized and 2 million for medium-sized populations.