

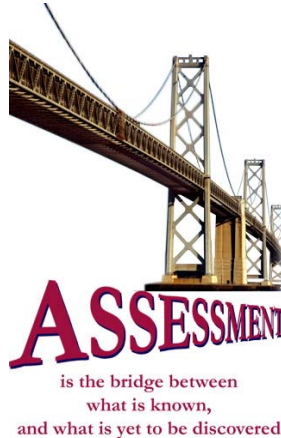
# Technical Skill Assessments

## Your Challenge

To meet the Perkins IV requirement for technical skills assessments in the most cost-effective manner possible.

### How can this be done?

1. Screening available instruments for suitability and applicability to Michigan's schools.
2. Constructing an original instrument.



## Qualities of a Good Educational Assessment

### Reliability

Consider:

- stability of assessment over time (test/retest correlational data)
- internal consistency of assessment measure (Internal Consistency Coefficients- KR-20 or coefficient alpha)
- precision of measure in decision making (standard error of measurement)



**IMPORTANT!** When multiple versions of assessment are available, equivalence of forms will be important to examine. When written responses are used, the accuracy of scoring should be reported.

**WHY?** Ensures that if a student takes a test twice or if two students take different forms of a test, any difference in results will reflect actual changes in skills, not measurement error.

### Validity

Does the assessment test what it is supposed to test?

Consider:

- **Content Validity:** Does the assessment adequately cover the content standards that it is designed to measure?

- **Predictive Validity:** Are students who do well on the assessments actually able to perform the tasks, jobs, or other skills that the assessment is designed to measure?

**WHY?** To be useful for decision making, a test must measure the skills decision makers want to know about.

### Bias-Free

Assures that the assessment permits **all** test-takers to show what they know and can do to the fullest extent possible.

Freedom from bias can be determined through:

- statistical methods
- human judgment

### Feasibility

Good assessments should:

- not be too time-consuming
- be easy to administer
- be easy for students to respond to

Feasibility can be determined through:

- student responses to questions that are asked of them during pilot testing
- by determining whether such factors were considered when pilot tested elsewhere
- examination of the assessment administration manual

### Affordable

A matter of relative judgment, in part, tempered by:

- the nature of the assessment
- how intricate the program area is
- the nature of the student standards
- usefulness of the assessment results



**IMPORTANT!** Some assessments that are more costly may be more affordable.

## Utility

Dependent on:

- How the assessment is administered
- How quickly the scoring takes place
- How quickly the results are returned to both students and program administrators
- How extensive the results are for the student, teachers, and school administrators, as well as the Michigan Department of Education
- Cost of program—in terms of money and effort needed to collect information

## Michigan's Process Preliminary Activities

### Information to be Collected for Available Assessments

### Materials

- Examination copies of assessments (it may be necessary to sign a security/confidentiality agreement to secure permission to review actual assessments)
- Administration Manual
- Technical Report

### Information

- Official title of the assessment
- Test publisher
- Publishing date
- Grade/age span of the assessment
- Suitability for secondary students completing a program
- Type of assessment (norm-referenced versus criterion-referenced)
- Testing mode (paper and pencil or online, or both)
- Number of test items of each type per grade/age
- Testing time per grade/age
- When assessment can be administered during the school year
- Cost (testing materials, scoring, and reporting)
- Restrictions on use (if any)

### Identify the Purpose of the Assessment

Can be found in:

- legislation
- program regulations
- policy created by the sponsoring agency or board



**IMPORTANT!** The sponsoring agency should consider **all** of the purposes for the technical skills assessments before they are selected or developed.

## Identify the Programs for Which Measures are Needed

### Career Clusters

The broader, more general the program area selected, the larger the set of content standards that apply to the program.



**PRO:** More likely that an assessment will cover the more important content standards.



**CON:** The more general the program level, the more challenging it will be to demonstrate the validity of the assessment.

### Federal Cluster Pathways

The narrower the scope of the program being measured, the greater the fit between the program's content standards and the assessment.

**PRO:** The validity will be higher due to a greater match between the standards and the assessments.



**CON:** Fewer assessments may be available which increases the number of assessments that Michigan needs to construct.



### CIP-Level Assessments



**PRO:** Can more accurately assess students' knowledge and skills in the smaller range of skills contained in any CIP area.



**CON:** More existing instruments will need to be located or, potentially, more assessments will need to be created.

## Identify the Standards of Each Program

### Common Standards:

Set of standards for the program area was written by an industry group and have been agreed on by Michigan Department of Education.

- Will be most likely to find an existing instrument

### Custom Standards:

Michigan has developed a unique set of student standards that differ significantly from those developed elsewhere.

- Assessments developed to measure other programs' standards might be customized for Michigan use
- If the extent of customization is not too great, the reliability and validity attributable to the original instrument may be used for the customized version

## Standards Do Not Exist:

Michigan has neither adopted content standards from others or developed ones of its own.

- It is highly recommended to find a partner that might be available to help carry out the work

Ways interested organizations and individuals can help:

- Serve as members of working committee that draft content standards statements
- Serve as reviewers of draft statements as they are produced
- Serve as conduits to keep their organizations and/or constituencies informed of the progress in developing content standards and recruit individuals to serve as reviewers of the standards manual
- Serve to collect evidence of the suitability of the content standards for postsecondary users

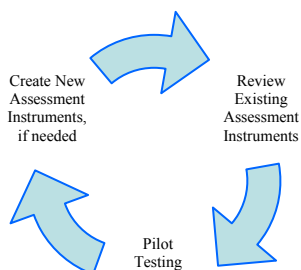
## Locating Suitable Assessment Instruments

Sources include:

- Journals that cover career technical programs
- Organizations that have developed career content standards
- Industry-based organizations that have helped to create content standards and skill certification efforts
- Higher education faculty
- Internet

## Collect Examination Copies of Existing Measures

**Intention:** To identify those programs for which existing measures are available.



## Step 1: Review of Existing Assessment Instruments

Once one or more assessment instruments have been identified for review and information on them has been collected, there are a series of steps that should be used to review the materials in order to determine their suitability for use in the state's technical skill assessment program.

## Thoroughly Review the Assessment Materials

1. Start by examining the review form that was sent to the publisher (See Appendix A).
2. Review the assessment information sent along by the publisher.
3. Review the assessment administration manual along with a copy of the student assessment booklet:
  - Purpose: To help the reviewer better understand the assessment administration process, as well as the nature of the assessment
4. Look at the technical report and other technical data provided by the publisher:
  - Goal: Determine the completeness of the technical information provided by the publisher
  - If the technical data is sparse or non-existent, the publisher should be contacted to ascertain whether any data regarding the reliability or validity of the assessment has been computed (on pilot or operational testing)

## Complete Assessment Review Form for the Assessment



An individual with expertise in educational measurement should review the assessment and related materials and complete the *Michigan Technical Skill Assessment Review Form* (See Appendix B).

## Determine Whether the Assessment is Well Aligned with Program Content Standards

Informal alignment methods rely on human judgment.

What the judges will be looking for:

- Are each of the key skills measured by the assessment?
- Is the number of assessment items per skill appropriate for the importance of the skill and/or the complexity of it?
- Do the assessment items measure the important portions of each of the content standards?
- Does every assessment item measure one or more the content standards?

## Determine Whether the Assessment Meets Minimum Technical Standards

Ideally, this determination will be based on data that results from the use of the assessment in pilot testing, field testing, or operational use. See *Determining Whether the Assessment Meets Minimal Technical Standards* form (Appendix C).

## Reliability

Test-retest reliability:

- A value of .90 is desired

Internal consistency:

- May be lower, especially where the instrument measures more than one skill, but reliability should still be .5 or .6 or higher
- Some assessments are designed to cover several content standards within one measure (that is, the score on the assessment is a combination of measures of several skills)
  - one would not expect to see high internal consistency, since several skills are being measured and reported on together within one test
- Most often, assessments individually measure and report on each content standard so the internal consistency of these individual measures is important

The final type of reliability that may be important concerns the administration of any performance items and the scoring of constructed-response items.

- When constructed-response items are used, some differences in human scorers will be observed



**IMPORTANT!** If no reliability information is available for an assessment, the Department may wish to propose a pilot test of the assessment in order to gather student results that can be used to estimate the reliability of the assessment.

## Validity

Content Validity

- The most prevalent form of validity evidence that achievement measures must have

How to determine content validity:

- The manner in which the assessment was created, especially the manner in which the accurate measurement of the standards was emphasized through item development, item editing, item reviews, pilot testing, and final item revisions
  - the types of alignment studies described in the previous section can also be used to verify the content validity of the assessments
- Expert panels can review the assessments to assure that they faithfully measure the content standards

Predictive validity is important for assessments that claim to predict future events, such as readiness for postsecondary training or readiness for work (success at work).

Construct validity will not occur often for technical skill assessment programs.

## Free of Bias

Can be ensured by:

- Mentioning in a technical report how item writers were trained to avoid such biases, whether and how the items were specifically reviewed for bias and sensitivity, and what statistical means were used to flag potentially biased items
- Differential Item Functioning (DIF). DIF statistics examine whether any sub-group scored statistically significantly lower on any test item

### Determine the Feasibility of the Assessment

- How complex is the assessment itself? How many parts does it have, how long is each of these, and how many types of assessment exercises are there in the entire assessment? Are some parts group administered and other parts individually administered? Anything that adds complexity to the assessment can reduce the feasibility of it in the minds of users.
- How difficult is it to learn how to administer the assessment?
- How challenging is it to score the assessment?
- What provisions have been made for the participation of students with disabilities and English language learners?

### Determine the Cost of the Assessment

In order to determine the actual costs associated with the use of an assessment, the user should carefully delineate the numbers of students to be assessed, the number of unique situations (classrooms) where the assessment will occur, the number of individuals who will receive results (i.e., teachers who will receive classroom results, administrators who will receive school results, administrators who will receive district or ISD results), which types of results are desired, and whether any ancillary features of the assessment are to be ordered.

### Step 2: Should the Assessment be Pilot Tested?

#### Reasons to Pilot Test:

- New assessment and no technical data exists
- When the assessment is not new, but the existing technical data is not applicable to the state's intended use of the assessment
- the existing data for the assessment is dated and there is a need to obtain updated information for the assessment
- the state's schools aren't sure whether to use the assessment, so pilot testing will allow them to "see the assessment in action"



## Determine Whether to Use the Assessment

If not clear cut, the relative merits or strengths of the assessment need to be weighed against the areas of weakness.

### Step 3: Creating New Assessment Instruments (see Appendix D)



**IMPORTANT!** This is not a step that should be taken lightly because of the time, effort, and expense involved in creating (and maintaining) a new assessment.

## Others to be Involved in the Project

Content Specialist

Understands the breadth of the content area and can make sure that it is suitably represented in the development of the assessment.



**IMPORTANT!** Should be someone who will not impose his or her own values in determining the balance and inclusion of various assessment topics in the new assessment.

## Collaborating with Other Agencies

One of the ways of reducing the costs of developing an assessment is to create a collaborative that shares in the work to be done and the costs for doing so.

A second approach might be to work within the state to network intermediate and local school districts to share in the costs of developing the assessment.



The Michigan Assessment Consortium (MAC) was formed to foster the development of high school course assessments. It could easily turn to the development of technical skill assessments by interested intermediate and local school districts.

## Preliminary Development Activities

There are a number of steps that need to be taken when determining how to develop assessments of a technical skill area.

## Assessment Purposes

The following is an example of possible purposes for a technical skills assessment:

- Determine whether students have learned enough to earn credit for the course/credit area
- Predict how well students will do in subsequent instruction, such as in community college or employment
- Program improvement
- Hold educators accountable for student learning

The assessment developers will need to determine whether the assessment purposes can be accomplished with a single examination or whether more than one assessment will be needed.

## Assessment Reporting

Some of the reporting levels that should be considered:

- Individual student reports for teachers
- Student and/or parent reports
- Classroom roster reports
- Item analysis report
- School summary reports
- Special reports

## Assessment Design

There are a number of important considerations when designing an assessment:

- Types of assessment (broad spectrum of skills or narrow set of skilled in greater detail)
- Speed or power test (AKA: timed vs. untimed)
- Length of the assessment (can be number of test items and/or test time)
- Types of items used in the assessment
  - Multiple Choice
  - Constructed-response items
  - Performance items
  - Portfolios
- Test administration mode
- Student response mode

## Developing the Assessment Blueprint

A good blueprint should contain:

- An indication of which content standards will be measured by the assessment
- A precise determination of how each of these skills will be measured
- An indication of how many items of each type will be developed for each skill
- An indication of the total length of the test, the number of test sections, and the composition of each test section
- The manner in which results will be reported and an indication that the test design will permit the types of reporting planned.
- An overview of the process

**Content by:**  
**Edward Roeber, Michigan State University**  
**Funded under a technical assistance grant from the**  
**United States Department of Education**  
**Office of Vocational and Adult Education**  
**July 2008**

**Contact:**  
**Jill Kroll, Ph.D.**  
**Office of Career and Technical Education**  
**Michigan Department of Education**  
**P.O. Box 30712**  
**Lansing, MI 48909**  
**517-241-4354**  
**krollj1@michigan.gov**



Michigan Department Education  
Office of Career and Technical Education  
**Statement of Assurance of Compliance with Federal Law**

The Michigan Department of Education, Office of Career and Technical Education complies with all Federal laws and regulations prohibiting discrimination and with all requirements and regulations of the U.S. Department of Education. It is the policy of the Office of Career and Technical Education that no person on the basis of race, color, religion, national origin or ancestry, age, gender, height, weight, marital status, or disability shall be subjected to discrimination in any program, service or activity for which it is responsible, or for which it receives financial assistance from the U.S. Department of Education.

Contact: Civil Rights Coordinator  
Office of Career and Technical Education  
P.O. Box 30712  
Lansing, MI 48909  
(517) 241-2091

## Michigan Technical Skill Assessments Survey

Attached is a copy of the technical skill standards adopted by Michigan. In order to select technical skill assessments for adoption by the State of Michigan, the Office of Career and Technical Education requires the following information on assessments under consideration for adoption. For the assessment specified below, please provide the following:

1. Information on the alignment of this assessment to the Michigan technical skill standards. What analyses have been carried out to demonstrate alignment to the standards? (See attachment showing the standards).
2. Information on the level and importance of the skills covered by assessment.
3. Describe the results of analyses of the test-retest reliability of this assessment.
4. Describe the number of equivalent alternate forms available, the frequency with which new alternate forms are developed and the analyses carried out to ensure that the alternate forms (including any subscales) are equivalent and reliable.
5. Describe the measures used to maintain the security of the test forms and items.
6. Describe how individual student assessment results may be obtained by school districts and/or the State of Michigan and the date by which results are available to the state and school districts.
7. Describe the administration schedule for the assessment. Can the assessment be administered according to district or state timelines? If not, how frequently is the assessment administered?
8. Describe the results of analyses of the internal consistency of this assessment.
9. Provide a copy of the inter-item correlation matrix showing all items, including those within subscales, and the item-total correlations for each item within the assessment.
10. Describe the results of analyses of the Inter-rater reliability (where applicable)
11. Describe the processes, rules and procedures for administration. Address how easy or difficult it will be for school districts to administer the assessment.
12. Itemize the costs associated with administering this assessment.
13. Describe how the results of the assessment will be useful to students.
14. Describe how the results of the assessment will be useful for program improvement.
15. Describe how the results of the assessment are utilized by business and industry for hiring, promoting or evaluating employees. Address the value of the assessment results to business and industry.
16. Describe the results of analyses examining the extent to which the assessment accurately discriminates between students with greater mastery compared to students with lower mastery (discrimination index).
17. Describe the difficulty level of the assessment. Do you consider the assessment to be appropriate for evaluating the technical skills of secondary students?
18. Describe the frequency with which the assessment is revised, updated and re-aligned with standards.
19. If applicable, describe the results of analyses of the validity of any subscales.
20. Describe the results of evaluation of the content by business, industry and postsecondary institutions.
21. Describe the results of the bias reviews conducted on the items.
22. Describe how the assessment results are provided to students and parents in an understandable manner.
23. Describe the results of analyses of the predictive validity of the assessment. (Do students who do better on the assessment do well in jobs or in postsecondary education in the same field?)
24. Provide any additional test statistics that may help in evaluation of this assessment.



## Determining Whether The Assessment Meets Minimum Technical Standards

1. **Reliability** – Please describe how the reliability of the assessment has been established. These include any of the following:

- a. Stability (Test-Retest)
- b. Internal Consistency (KR-20 or Cronbach Alpha)
- c. Split-Half
- d. Alternate Form
- e. Inter-Rater Reliability (for constructed-response or performance items requiring hand scoring)

What analyses have been performed and what data are available to demonstrate the reliability of the assessment. Is the data available suitable for the assessment?

2. **Validity** – Please describe how the validity of the assessment has been established for each type of inference for which the assessment is recommended.

- a. Content Validity – An essential element in the selection of appropriate instruments to assess Michigan technical skill standards is the alignment of the assessments and the Michigan technical skill standards. Alignment will be judged four ways:

1. How many of the Michigan technical skill standards are measured by the assessment?
2. How many assessment items measure one or more Michigan technical skill standards?
3. Does the assessment emphasize the most important skills in the Michigan technical skill standards?
4. Does the assessment assess students at a level appropriate for high school?

What analyses have been carried out to demonstrate alignment to the standards? Please provide information on any alignment study that has been carried out for this assessment to the Michigan technical skill standards. Please describe the results of evaluation of the content by business, industry and postsecondary institutions.

- b. Predictive Validity – Please describe the results of analyses of the predictive validity of the assessment. Is there information to show that successful performance on the assessment predicts success on the job? Do students who do better on the assessment do well in jobs or in postsecondary education in the same field?

Please describe the results of analyses examining the extent to which the assessment accurately discriminates between students with greater mastery compared to students with lower mastery (discrimination index).

- c. Concurrent Validity – Does performance on this assessment correlate with success on other related/ comparable assessments?
- d. Construct Validity – If the assessment is designed to measure a more theoretical construct, is there evidence to support the proposed interpretation of the results?
- e. Sub-Scale Validity – If applicable, please describe the results of analyses of the validity of any subscales.

3. **Bias and Sensitivity Reviews** – Please describe the results of bias and sensitivity reviews conducted on the items.

## Step 3: Creating New Assessment Instruments

### Things to Consider:

1. Will the assessment development work be done with state and local staff or will an assessment contractor be used?
2. Who will actually lead the assessment development project?
3. What entities will be involved in some capacity in the development project?
4. Is it possible to collaborate with other agencies (states or ISDs) in creating and administering the assessment? If so, how?
5. How will the project be paid for?

### Use an Assessment Contractor

Assessment contractors will typically be able to handle much of the development work and the logistics involved in the assessment project, thus relieving local and state educators from much of the work. Downfall is they are expensive

### Leading the Assessment Project

Leading the assessment project should come from someone who is familiar with how assessments can be created and with the assessment training and management experience to successfully plan the project and manage it to successful conclusion

Assessment project managers can come from several types of organizations including local and intermediate school districts, universities, non-profit and for-profit organizations, and retirees from such organizations.

### Others to be Involved in the Project

Project Manager - to assist the director in carrying out the developmental work

Project Clerical Staff

### Collaborating with Other Agencies

#### Paying for the Project

- one or more agencies paying the cost
- collaborative of states or intermediate and local districts sharing the costs
- seeking external funding- from government agencies
  - foundations
  - industry groups

Pro of a group of entities working together: costs for such work can be shared among several participants.

### Development of Project Staffing

- Project Director
- Project Manager
- Several content specialists
- One or more clerical staff
- Two or three technical advisors
- Panel of about 5 content experts

One or two individuals who are aware of bias and sensitivity issues