

**A Guide to Aid Nebraska School Districts in Addressing the
Quality Components in the District Assessment Plans**

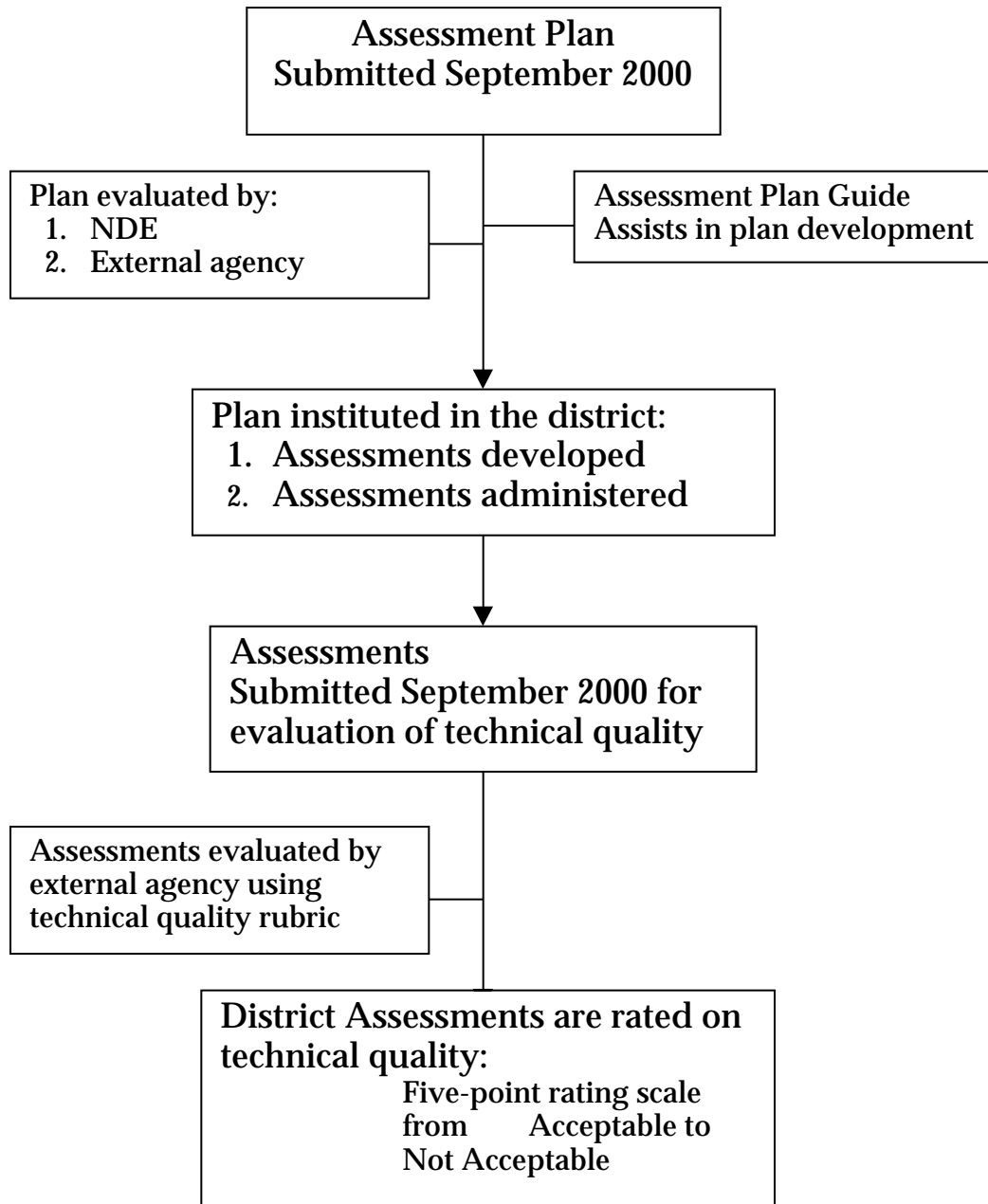
**Prepared for the Nebraska Department of Education by the
Buros Center for Testing**

April 2000

(Draft 3/23/00)

Buros Center for Testing, April 2000
This document may be copied and distributed to school districts within the State of
Nebraska. Other uses are prohibited.

Assessment Plan and Assessment Instruments Process



Guide to Addressing the Quality Components in the District Assessment Plans¹

Districts are required to submit an assessment plan to the Nebraska Department of Education and eventually to an external review agency. The Plan is to be submitted every five years. For the District Assessment Plan to be approved it must contain certain components. Many of these components are checked to insure they are in the plan. In addition, certain components are evaluated to insure that they meet criteria associated with sound assessment practices.

This guide is intended to provide districts with general information about the kinds of information they will need to include in their plan to show that sound measurement practices will be followed in the development and/or selection of the assessments that will be used in the district.

The following 6 components of the District Assessment Plan will be evaluated:

1. The assessments reflect the state standards (see pages 3-4)
2. Students have an opportunity to learn the content. (see pages 5-6)
3. The assessments are free from bias (see pages 6-8)
4. The level is appropriate for the students (see pages 8-9)
5. There is consistency in scoring (see pages 9-10)
6. Mastery levels are appropriate (see pages 10-11)

Evaluating the District's Assessment Plan

¹ Note that there are two evaluations. The first is the evaluation of the proposed assessments and the ways those assessments **will be** developed/selected and used. This evaluation is prior to using any assessments and looks at the proposed procedures that will be undertaken to put an assessment program in place. The second evaluation will occur after the assessments are developed and used. This second evaluation will address the same quality components to insure that the quality of what **was actually done** (rather than what was planned) reflects sound measurement practices.

According to the STARS document, a district's Assessment Plan will describe the following assessments (pages 50-51):

- A norm-referenced, standardized achievement test (selected from an approved list).
- A statewide writing assessment.
- District selected or developed criterion-referenced assessments.

The quality components described above apply to both the norm-referenced and the criterion-referenced assessments, but the major focus of these components is on the criterion-referenced assessments selected or developed by the district.

For each of these six quality components this guide provides

- A general description.
- General ways that information could be gathered for that component.
- Documentation required for the norm-referenced, standardized achievement test used in the district plan.
- Documentation required for district selected or developed criterion-referenced assessments.

Quality Components

1. **The assessments reflect the standards**

General Description: This component asks for a documentation of the methods that will be used to decide what standards will be covered in the assessment and that the coverage of the standards is adequate (i.e., there are enough items or tasks to provide a valid inference about the extent that the standard has been achieved).

General ways that the information could be gathered for that component.

There are a number of methods that could be used to provide evidence of the match of items to standards. Several methods are illustrated briefly below.

- Convening panels of teachers to judge how well the questions on the assessments match the relevant content standards. Districts, for example, could bring together panels of teachers (or other “experts”) to match (align) the content standards to the questions or tasks on the assessments.
- For commercially available tests, some publishers have examined their test and provided documentation of which standards, in their view, are covered by their test. This information may be found in the test manual or may be obtained by contacting the publisher directly.
- The adequacy of coverage might also be judged by panels to insure that the major aspects of the standards are measured.
 - For objectively scored assessments (e.g., those containing multiple choice or short answer questions) at least 5 items must be included.

- For subjectively scored assessments (e.g., writing samples) the assessment and scoring must be adequate to make a reasonable inference about the standard.

The goal is to show that a reasonable method will be used to determine that the content of the assessments matches the content standards that are to be measured and that there is sufficient evidence (e.g., number of items) to judge mastery.

Documentation required for the norm-referenced, standardized achievement test used in the district plan.

For the norm-referenced, standardized achievement tests used by the district, the plan may reference the studies conducted by the Buros Center for Testing to align selected standardized tests with the standards in several areas and at several grade levels. This information is available in the STARS Update, dated May 2000.

For tests not included in these studies (or subject areas and grades not included), Districts must provide evidence of how these assessments reflect the state standards (e.g., by using methods such as those described above).

Documentation required for district selected or developed criterion-referenced assessments. Districts will need to describe briefly how they will determine the match between the criterion-referenced assessment and the standards and how they will assure that the assessments represent adequate coverage of the standards.

- For example, districts may propose to convene panels of teachers in the respective content/grade level to examine the match of the test content with

- the standards the test is intended to measure. If districts convene one or more panels to make these determinations, then panels should be described in terms of membership and why these panelists are qualified to serve.
- If teacher panels are not used, then some other method may be proposed that will result in a trustworthy determination of the match between the test content and the standards.

2. Students have an opportunity to learn the content

General Description. The intent of this component is to describe a method to examine how the content covered in the assessment is delivered in the instruction students receive prior to their administration.

General ways that information could be gathered for that component. There are a wide variety of strategies that could be proposed, either singly or in combination, to investigate the extent that students have an opportunity to learn the content of the assessments. Several of these methods are described briefly below.

- Representative panels of teachers representing the grade and content area could examine district curriculum, including scope and sequence, to determine whether the content measured on the assessments will be delivered to students in the district prior to the administration of the assessment.

- Grade/content appropriate teachers could be surveyed regarding where in their lesson plans the test content is addressed.
- Selected samples of teachers' classroom assessments could be examined to ascertain whether the content is covered in these assessments.

Documentation required for the norm-referenced, standardized achievement test used in the district plan. The district's Assessment Plan must include how information about opportunity to learn will be gathered at the district level for the norm-referenced, standardized achievement test content.

Documentation required for district selected or developed criterion-referenced assessments. The district's Assessment Plan must include how information about opportunity to learn will be gathered at the district level for the criterion-referenced assessment content. If districts convene one or more panels to make these determinations, then the panels should be described in terms of membership and why the panelists are qualified to serve.

3. The Assessment are Free from Bias/Offensive Language or Situations

General Description. The intent of this component is to describe a method to examine how the assessments will be examined to ensure that they are free from bias or offensive language/situations. A test item or task is free from bias or offensive language/situations if a) understanding the content of the item or task is not based on the students' socio-economic, racial, cultural, or ethnic

background, and b) the content of the item or task is not likely to be offensive to students.

General ways that information could be gathered for that component. There are several strategies that could be used, either singly or in combination, to ensure that assessments are free from bias or offensive language/situations. Some of these methods are illustrated below.

- During the assessment development process, item/task writers could be given an orientation to this issue and presented with examples of bias or offensive language/situations that should be avoided when writing the test questions. A list of examples of bias or offensive language/situations that might be used for this purpose is available in Appendix A.
- After the assessments are drafted, panels of teachers representing various socio-economic, racial, and ethnic groups could review the assessments to insure that no biased or offensive language or situations are included.
- Statistical analyses may be conducted to examine the extent that the items do not demonstrate statistical bias.

Documentation required for the norm-referenced, standardized achievement test used in the district plan. Information about bias or offensive language/situations might be included in the published materials describing the development of the test. A reference to these materials (e.g., the technical manual) is one way to document that the assessments are free from bias or offensive language/situations.

Documentation required for district selected or developed criterion-referenced assessments. The method(s) to be used to gather information about freedom from bias or offensive language/situations should be described for the district selected or developed criterion-referenced achievement test content. If districts convene one or more panels to make these determinations, then the panels should be described in terms of membership and why the panelists are qualified to serve.

4. The Level is Appropriate for the Students

General Description. The intent of this component is to describe how the district will determine the developmental appropriateness of its assessments.

Developmental appropriateness includes a) the appropriateness of the cognitive requirements, and b) the reading level.

General ways that information could be gathered for that component. There are several strategies that could be proposed, either singly or in combination, to ensure that assessments will meet the requirement of being developmentally appropriate. Several methods are described briefly below.

- Convene a panel of grade/content level teachers, sometimes augmented by specialists in reading and/or cognitive psychology, examine and evaluate the assessment instruments from a developmental perspective.
- Conduct readability analyses on assessment passages and items.

Documentation required for the norm-referenced, standardized achievement test used in the district plan. Information about the developmental level of the test might be included in the published materials describing the development of the test. A reference to these materials (e.g., the technical manual) is one way to document that the assessments are developmentally appropriate.

Documentation required for district selected or developed criterion-referenced assessments. For criterion-referenced assessments, districts may convene one or more panels to make these determinations. If this is done, the panels should be described in terms of membership and why the panelists are qualified to serve. If readability analyses are to be conducted the specific methodology should be described.

5. **There is Consistency in Scoring**

General Description. The intent of this component is to describe how the district will determine the extent that the assessment scores are consistent and, to the extent feasible, free from random error. There should be descriptions of procedures for both objectively scored assessments (e.g., assessments that include only selected response questions) and for assessments that are scored more subjectively (e.g., writing samples or other constructed response items or tasks).

General ways that information could be gathered for that component. There are several strategies that could be proposed, either singly or in combination, to ensure that assessments will meet the requirement of being scored consistently. Several of these methods are described briefly below.

- For objectively scored tests, the procedures might include any of the accepted methods of estimating test score reliability.
 - Test-retest, alternate forms, or internal consistency estimates of score reliability).
- For more subjectively scored tests the methods used to enhance consistency might include:
 - having all, or a sample, of papers double scored to estimate the percent of scorer agreement. This would be preceded by having a panel screen and pre-test the clarity and adequacy of the scoring rubric, training scorers to a certain criterion level of accuracy using “marker” papers (papers that have been prescored by “experts” that are used to train and calibrate scorers),
 - Using outside scorers who have been trained to a criterion of scoring accuracy, and who are audited to insure their continued accuracy.

Documentation required for the norm-referenced, standardized achievement test used in the district plan. Information about the score consistency (reliability) might be included in the published materials describing the development of the test. A reference to these materials (e.g., the technical manual) is one way to document the level of scoring consistency.

Documentation is required for district selected or developed criterion-referenced assessments. The method(s) used to determine consistency of scores should be described for the district selected or developed criterion-referenced achievement test content.

- For objectively scored tests a reference can be made to any one or more conventional estimates of test reliability.
- For more subjectively scored tests documentation should include a) procedures for developing the scoring rubric, b) plans for training scorers (teachers), c) policies for who can serve as scorers (e.g., whether teachers can score assessments from their own students), and d) methods for determining the extent of agreement among scorers.

6. Mastery levels are appropriate

General Description. The intent of this component is to describe how the district will determine the procedures used to make “mastery” decisions (i.e., meeting or not meeting standards) about students based on the students’ performance on assessments.

General ways that information could be gathered for that component. There are several strategies that could be proposed, either singly or in combination, to ensure that mastery determinations will be reasonable. The major consideration for making the mastery decision is that the decision is not arbitrary. For example, simply saying that a score of at least 70% (or some other arbitrary value) of the total possible score will be considered mastery is an arbitrary

decision. This value does not take into account how hard or easy the assessment items/tasks are and therefore is not an appropriate way to determine mastery.

Several acceptable methods are described briefly below.

- Systematic examinations of the assessment that take the difficulty of the items/tasks into consideration (e.g., methods described by Angoff or Ebel)
This method involves a process where grade/content level teachers anticipate how students who have minimally met the standards would perform on the test questions
- Using teachers' professional judgments to classify students into mastery categories, and then match the students' scores with the different mastery classifications (these methods include "contrasting groups" and "borderline group."). This method involves grade/content level teachers classifying their students, prior to knowing how the students performed on the test, into performance categories. One of these performance categories is for students who have just minimally met the content standards. The scores of these "borderline" students are then used to identify an appropriate passing score (usually either the average or median score for this group of students is used).
- Other approaches are also acceptable, as long as they take into consideration the difficulty of the test and the anticipated or actual performance of students in the district who have minimally met the standards.

Clarification about what documentation is required for the norm-referenced, standardized achievement test used in the district plan.

For the norm-referenced, standardized achievement tests used by the district, the state has determined that the students scoring above the 50th percentile can be classified as masters of the appropriate standards.

Documentation required for district selected or developed criterion-referenced assessments. Districts will need to describe how they will determine the performance levels on the assessments that will be defined as determining mastery. If one or more panels make these determinations, the panels should be described in terms of membership, why the panelists are qualified to serve, and the activities the panelists will use to make the decision.

Appendices will be incorporated:

Appendix A: Bias Review

Appendix B: Developmental Appropriateness Review

APPENDIX A

Guidelines for Bias and Sensitivity Review of District Assessment Instruments

BIAS AND SENSITIVITY

Specifications to Ensure A Multicultural Perspective and to Avoid Bias and Controversial and Sensitive Topics

(Adapted from Bias Issues in Test Development, National Evaluation Systems, Inc.)

Purpose of this appendix:

- To suggest methods that Nebraska school districts might use to insure that their district assessment strategies avoid biases against race/ethnic, religious, cultural groups and biases against gender, age, disability, or sexual orientation
- To avoid issues that are sensitive; i.e., may be considered offensive to a large segment of the population

Guidelines

The following categories are to be considered when developing test materials to detect potential bias issues.

Language Usage

Use language that has the same basic semantic content for all persons regardless of race, gender, ethnicity, age, sexual orientation, or physical or mental condition.

Consider both the denotation and connotation of words, expressions, and symbols and avoid negative and potentially inflammatory connotations.

Stereotyping

Be sure that materials do not convey any of the following:

- that a population group is deserving of a particular fate
- that a population group is by nature dependent on help from the majority culture
- that a population group lacks some positive quality fairly common to humans
- that a population group has an excess of a quality fairly common to humans
- that a population group has a genetic deficit or surplus in some area of intellect, talent, or ability
- that what may be a norm only in Western culture is "truth" or that European civilization is "better than" (as opposed to "different from") other civilizations
- that a causal link exists between membership in a particular population group and poverty, crime, intelligence, physical stamina, mental alertness, etc.

Representational Fairness

Include females and males in nontraditional occupations, roles, and relationships.

Represent women, older persons, religious, ethnic, and racial minorities, and persons with disabilities in many different environments and occupations and in roles of diverse status and power.

Choose names that connote a broad balance of national origins and present names of both sexes in both traditional and nontraditional roles.

Depict older persons in a variety of roles and activities; wherever possible, balance the depiction of older persons as "grandparent, sedate" types with depictions of individuals who are active, intelligent, and even fallible.

Depict persons with disabilities as productive members of society with versatile lives; if the disability is part of a story line or test materials, focus should be on the management of the disability, not on the disability itself for no apparent reason.

Allow for the family living arrangements to include, or at least not to exclude the families as individuals or couples who have no children, single parents, adoptive parents, same-sex parents, group living arrangement, and multigenerational arrangements.

Treatment of Races and Cultures

Reflect the multicultural composition of American society when showing people in everyday situations so as to convey that equal status and non-segregated relationships are the ideal toward which society is striving.

Make no suggestion that any racial, ethnic, or religious group is more or less worthy than any other.

Make no suggestion that any pattern of success or failure is correlated with a particular group or that any single group is more or less talented than any other group.

Convey wherever possible that differences between diverse minority groups is valuable.

Balance depictions of people in what are traditional roles with depictions of people in nontraditional roles.

Make no attempt to play down or hide history; materials should be presented accurately and not evaluated with values of this present day.

Be sensitive to what names minority or cultural groups prefer.

When appropriate use terminology that is used within a particular cultural group

Treatment of the Sexes

Show both sexes obtaining education and training in all fields and at all levels.

Avoid implications that holding a particular job is incompatible with being a man or a woman.

Avoid showing a connection between a set of interests, attitudes, or career choices and concept of masculinity or femininity.

Avoid suggestions that males should earn more than females and vice versa.

Avoid implications that men (or women) are expected to be the only economic support for their families.

Avoid implications that some interests and activities are restricted to or favored by males or females.

Portray both males and females as persons who pursue interests based on their individuality rather than their gender.

Portray all work as honorable and respectable.

Show both males and females in management and non-management positions; show them both in roles of leadership.

Make attempts to balance materials portraying the experiences and achievements of men with experiences and achievements of women, wherever possible.

Discuss the economic, political, and social limitations of women's rights, opportunities, and accomplishments in the context of an historical period; make no attempts to alter or hide the facts.

Attempt to diversify choices of histories or events to create the broadest spectrum of historical and cultural representations.

Treatment of Persons With Disabilities

Represent people with disabilities in text and illustrations; focus should not be on the disability itself but on the interaction of persons with disabilities in everyday life.

Avoid portraying persons with disabilities as objects of curiosity.

Show people with disabilities interacting positively in a variety of interpersonal relationships, do not show them being helped by others unless such help is appropriate (e.g., signing for the deaf).

Show people with disabilities as community leaders, business executives, parents, etc.

Sensitive Topics

The following are sensitive topics in Nebraska. If these topics are covered in the test materials, care must be taken to treat them in light of the guidelines listed below.

Death	Drugs/Alcohol	Family Violence
Murder	Racial Bias	Religious Holidays
Guns	Religion	Suicide
Ghosts/Spirits	Family Problems	Homosexuality
Pregnancy	Homelessness	
Violence	Divorce	

Avoid doctrine regarding specific points of view as they relate to religion, the home, family relationships.

Uphold the values fundamental to American democratic society, such as respect for majority rule, minority rights, privacy of person and property, and equality.

Avoid an undermining of other cultures, value systems, and societies in the world at large.

Uphold those standards of ethical behavior, which enable people to live together in an atmosphere of trust, safety, and respect.

(For specific guidelines relating to reading see below.)

Topics to Be Avoided

Avoid issues that ask students to divulge private, personal and family information:

- Feeling about siblings, parents, other relatives
- Alcohol and drug use

The following topics should also be avoided in district assessments:

Abortion

Birth Control/Condoms

Child Abuse/Child Neglect

Incest

Rape

Sexual Activity (excluding reproduction)

Sexually Transmitted Diseases