

How to assess student performance in history: going beyond multiple-choice tests

By: [Julie Edmunds](#)

Edmunds, J. (2006). *How to assess student performance in history: beyond multiple-choice tests*. SERVE.

Made available courtesy of SERVE: <http://www.serve.org/>

*****© SERVE. Reprinted with permission. No further reproduction is authorized without written permission from SERVE. This version of the document is not the version of record. Figures and/or pictures may be missing from this format of the document. *****

Abstract:

The purpose of this document is to help teachers implement effective assessment strategies when teaching history. The primary audience for this publication is third through twelfth grade teachers who cover the topic of history in their instruction. Secondary audiences include school administrators, district and state curriculum staff, or individuals providing professional development.

Keywords: teachers | assessment | history

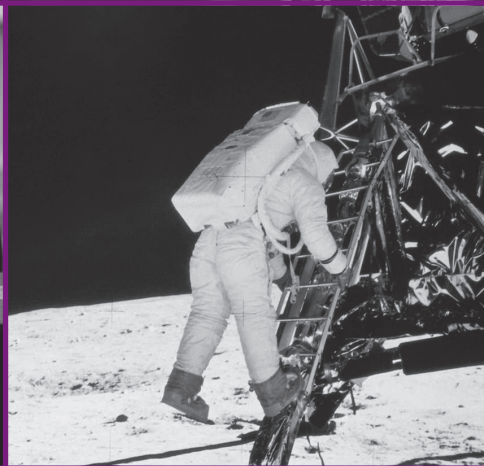
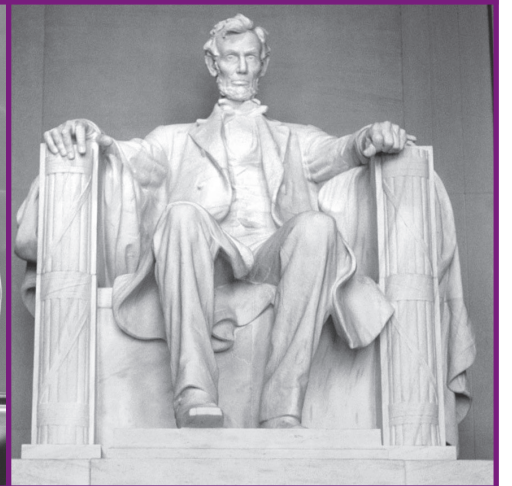
Article:

*****Note: Full text of article below**



HOW TO ASSESS STUDENT PERFORMANCE IN HISTORY:

GOING BEYOND MULTIPLE-CHOICE TESTS



SERVE Center
at the University of North Carolina
at Greensboro

HOW TO ASSESS STUDENT PERFORMANCE IN HISTORY: GOING BEYOND MULTIPLE-CHOICE TESTS

© SERVE, 2006

SERVE Center
at the University of North Carolina
at Greensboro

Associated with the
School of Education,
University of North
Carolina at Greensboro

PRODUCED BY

The SERVE Center at the University of North Carolina at Greensboro

WRITTEN BY

Julie Edmunds

EDITED BY

Donna Nalley

DESIGNED BY

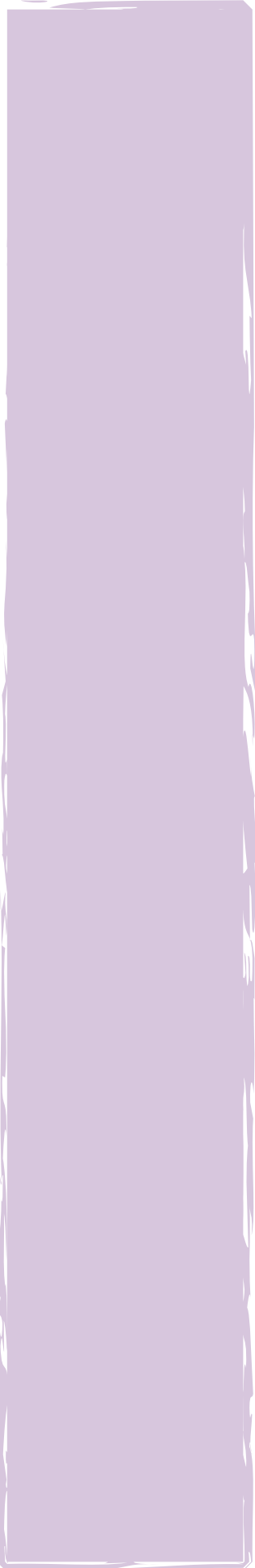
Tracy Hamilton
Jane Houle

This publication does not necessarily reflect the views or policies of the Institute of Education Sciences, U.S. Department of Education, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government.

This document was originally developed with funding from the Institute of Education Sciences, U.S. Department of Education, under contract no. ED-01-CO-0015.

TABLE OF CONTENTS

CHAPTER 1:	Introduction and Overview of Assessment..... 1
	1Introduction
	2The Purpose of This Manual
	2What is Classroom Assessment?
	4Anatomy of an Assessment
CHAPTER 2:	Desired Student Outcomes in History 6
	9Articulating Course Objectives
CHAPTER 3:	Methods of Assessment 14
	16Dialogue and Oral Response
	20Essays and Open-ended Responses
	26Projects and Investigations
CHAPTER 4:	Establishing Criteria 31
	33Checklists
	34A Point System
	37Rubrics (Holistic and Analytic)
CHAPTER 5:	Using Assessment Information: Feedback and Grading 47
	47Formative Assessment: Providing Feedback
	49Summative Assessment: Grading
CHAPTER 6:	Some Concluding Thoughts..... 52
	52Appropriate Assessments in an Era of Accountability
	53Getting Started
	References..... 57
	Resources 59



INTRODUCTION AND OVERVIEW OF ASSESSMENT

Introduction

“When teachers’ classroom assessments become an integral part of the instructional process and a central ingredient in their efforts to help students learn, the benefits of assessment for both students and teachers will be boundless.” (Guskey, 2003, p. 11)

SERVE, in its role as the federally funded Regional Educational Laboratory for the southeast, has worked since 1990 to provide professional development to teachers in the area of assessment. The work has focused on translating what is known from the measurement, applied research, and practitioner worlds into concrete ideas for teachers to implement to improve their assessment practices. The particular focus has been on how to improve the assessments teachers use and the way they use them to improve student motivation and learning. SERVE has learned over the years that working on assessment in the classroom is best accomplished within the context of the content area taught. This publication builds on work SERVE has done with districts that are piloting new approaches to professional development in U.S. History through the U.S. Department of Education’s Teaching American History grant program.

How to Assess Student Performance in History: Going Beyond Multiple-Choice Tests addresses some real assessment challenges that teachers have identified:

- Figuring out what really is important for students to know and be able to do in history.
- Teaching the skills of “doing history” in a world of testing that often seems to value only factual knowledge.
- Identifying and using assessments that provide teachers with better information than only multiple-choice exams.
- Getting students motivated to do a good job on essays and other written work.
- Helping students learn to improve their own work and produce quality products.
- Holding students accountable for quality work as opposed to just turning in something.

THE
OVERVIEW
OF
ASSESSMENT
PRACTICES
IN
U.S. HISTORY

The Purpose of This Manual

This publication is designed to get you ready to meet those challenges by helping you:

- Understand the role of assessment in improving student learning;
- Consider different learning outcomes for students in history and pick those that are most important for your students;
- Determine some of the best ways of assessing student learning and tracking and evaluating their progress toward history outcomes;
- Develop methods of scoring and grading students' work that will provide information and accountability; and
- Integrate your assessment program with the statewide testing programs currently in place.

This publication is not intended as a text but as a **self-study resource**. We hope you will interact with it, respond to questions posed, and use the manual as an opportunity to reflect on your assessment practices. We suggest you work through this manual with at least one other teacher, if possible, because of the valuable sharing of ideas that would result.

What is Classroom Assessment?

**"The aim of assessment is primarily to educate and improve student performance, not merely audit it."
(Wiggins, 1998, p. 7)**

Assessment plays a role at many different layers of education. The federal government and many states see assessment as the linchpin for educational reform. In this manual, however, we focus on classroom assessment or assessments *used by teachers in their own classrooms*. Although many other assessments (state-level tests, National Assessment of Educational Progress, nationally standardized tests) are given to students in classrooms, we do not consider them *classroom assessments*.

We see classroom assessment as having four main purposes. The first three include:

1. Diagnostic or needs assessment purpose: To determine what students already know so teachers can decide the topics and approaches to use.
2. Formative purpose for teacher: To assess student knowledge or performance on some key topic or dimension to inform instructional plans.
3. Summative purpose: To judge or evaluate student performance (i.e., give a grade).

In addition, research is increasingly clear that the quality of the feedback teachers give students relative to how to improve is an absolutely critical aspect of classroom assessment. (Think about how feedback shapes performance for musicians or athletes during their formative years).

This leads to the fourth purpose:

4. Formative purpose for students: To help students develop the skills to reflect critically on their own work. By asking students to assess themselves, teachers encourage students to engage in the type of higher-order thinking necessary for life today.

Formative Assessment

Formative assessment is focused on improving student motivation and learning with the goal of producing higher-quality work or thinking. It's important to realize that there are two different audiences for formative assessment. One audience is the teacher. That is, many teachers may check for student understanding by asking questions or by observing students as they discuss a topic in small groups. These teachers are informally "collecting data" that will help them determine what needs to happen next in instruction. So the teacher is the data user. The second audience for formative assessment is the student. Students need to know what would move their essay answer on a particular question from a "C" to an "A." They need to know what it means to read content deeply for understanding and how their strategies for studying content can be improved.

Research shows that providing students with effective feedback can increase student achievement significantly (Marzano, Pickering, & Pollock, 2001). Feedback is most effective when it:

- Is timely, occurring within one to two days of the work;
- Provides feedback specific to the student's work; and
- Is relative to a criterion or standard.

Summative Assessment

Summative assessment looks at whether a student has achieved the desired learning goals or met standards. In the classroom, summative assessments usually occur at the end of instruction and document what students have learned. Looking at the grades in a teacher's grade book should give an idea of what the key instructional goals or outcomes were for a grading period. These grades most likely represent summative assessments (tests, quizzes, projects, reports, written assignments, etc.) that tell the teacher whether the student has mastered the skills or learned the content. A key aspect of summative assessment is determining the level to which students need to "master" the content and thinking. Tests that define "mastering" content at the level of memorizing events, names, and facts are less likely to be building students' thinking skills than tests that ask students to write about big conflicts or themes that recur over time.

Good summative assessments are:

- Useful. The assessment must provide you with useful information about student achievement in the course. The assessment must be tied to the learning goals you have and those learning goals must be important. If you assess unimportant or trivial concepts or just use chapter tests without really looking at the items critically in terms of whether they reflect your teaching, what have you learned about what your students know?

- Valid for your purposes. The assessment must measure what it is supposed to measure. For example, if you ask students to draw a map reflecting the change in U.S. borders from 1789-1820, you will need to ensure that the assessment is scored based on students' understanding of the concepts not based on their ability to draw. Sometimes, the way the test is presented (e.g., small print with lots of complicated or confusing directions or too many items) can make it a less valid measure of the content being tested. It may be more a measure of student persistence than a measure of their knowledge of the content. As a teacher, taking a test yourself before giving it to your students will help ensure that the items reflect content you actually taught. It will also help you to decide if there are some aspects of the questions or layout that are content irrelevant, representing extraneous hurdles for students that could be simplified.
- Reliable. Reliability has to do with the extent to which the score you give a student on a particular assessment is influenced by unsystematic factors. These factors are things that can fluctuate from one testing or grading situation to the next or from one student to the next in ways that are unrelated to students' actual achievement level (e.g., luck in guessing the right answer, lack of time to complete the assessment on a particular day, teacher bias or inconsistency in scoring of essays across students or from one test to the next). Thinking about how to reduce these factors such that the scores given are likely to be the most accurate reflection of students' true achievement levels on the task or test should be an ongoing process for teachers.
- Fair. The assessment must give the same chance of success to all students. For example, a large project that is done at home can be biased against low-income students, favoring students whose parents have extra time to help them over those whose parents need to work.

We will be revisiting these concepts throughout this publication. The important thing to understand is that assessment is much broader than tests or grades in a grade-book. As explained above, formative assessment is emerging through research as critical to students' continued improvement in school but it is often not well-understood or used in classrooms.

Anatomy of an Assessment

Each time you use an assessment in your class, you are really answering four questions:

1. What do you want students to know or be able to do? (the purpose or goal of the learning and, hopefully, by extension, the purpose of the assessment)
2. What is the best assessment method to use given your instructional goals? (the kind of assessment)
3. How are you going to evaluate the students' responses? (the analysis of the results)
4. What are you going to do with the information? (pre-determined use for the assessments)

Often, teachers answer these four questions based on how they have always done things; they have not had the time or resources to think them through with colleagues or content experts. For example, the most traditional classroom assessment situation is a multiple-choice chapter test. In this situation:

1. The learning goals include the content in the textbook so the objects of the assessment are the facts/concepts that someone else (textbook developer) has decided are important;
2. The assessment method is a paper and pencil, multiple-choice/short answer test;
3. The analysis involves taking the number right out of the total number of items and providing the student with a score on a scale of 100 points; and
4. The test is entered into a grade book and summarized with four other tests and homework to provide a grade for the semester.

It is not to say that this is not a valid way of assessing, but that there perhaps could be a much broader set of formative and summative assessment tools to use that will provide you and your students with better information.

In the next sections, you will find help in answering these four questions. **CHAPTER 2** focuses on the establishment of learning goals that are relevant for history. **CHAPTER 3** includes an overview of assessment methods and a discussion on matching appropriate methods to different types of learning goals. **CHAPTER 4** answers the third question as it looks at making sense of the information you get from students through checklists and rubrics. **CHAPTER 5** examines how to use assessment information through providing feedback to students and incorporating a variety of assessments into your grading practices.

DESIRED STUDENT OUTCOMES IN HISTORY:

WHAT DO WE WANT STUDENTS TO KNOW AND BE ABLE TO DO?

“People who have a clear idea about where they’re going are more likely to get there. Teachers who have a clear idea of the learning outcomes their students should achieve will be more likely to help their students attain those outcomes.” (Popham, 2001)

The first step in helping students succeed in history is determining what we want them to know and be able to do. This principle underlies much of the standards movement today.

There are many different discussions of what students should know and be able to do in history. One representation includes three dimensions of historical literacy (Drake and McBride, 1997):

1. Knowledge of historical facts, themes, and ideas
2. Historical reasoning—ability to analyze, synthesize, and evaluate historical evidence
3. Communication of historical knowledge and reasoning to a wider audience

Although other groups may use different terminology, it is clear that students in history need to know and understand historical events (content) and also be able to apply specific kinds of thinking skills commonly used in the study of history. Unfortunately, in many cases, the content/thinking skills discussion is set up as a dichotomy as if you could only do one or the other. For example, much of history is taught traditionally as a collection of isolated historical facts. Sometimes, these facts are combined into a narrative to give some context. Yet, how often do teachers have their students use this information as they engage in the messy, yet rewarding, work of investigating and interpreting historical issues and events? If students don’t actually use and work with historical information, they won’t remember it far beyond the test and they won’t see any purpose behind it.

In a description of two very different history teachers teaching a civil rights unit, one author highlights how having students engage in the process of “doing history” results in students who “hold views of history that are consistently more thoughtful, nuanced, and complex...(and students who) view history as a more vibrant and powerful influence on their lives” (Grant, 2003, p. 60). Whereas the first teacher below represents a more traditional lecture approach, the second teacher’s unit is more experiential.

Blair’s students experience a seemingly coherent and engaging narrative centered on the beginning of the civil rights movement. Yet neither that narrative nor the ones that Blair presents in other units seem to inspire students’ engagement with or their understanding of history.

DO
WE
KNOW
WHAT
WE
WANT
STUDENTS
TO
KNOW
AND
BE
ABLE
TO
DO?

Instead, Blair's students seem to equate what they learn about history to school knowledge...information that helps them pass tests but means little to them otherwise...

Strait's unit is decidedly broader, less storylike and more experiential than Blair's. Her students read, write, listen, view, and interact in a range of instructional settings that seem far more intellectually open-ended than Blair's narratives. Strait's students...unlike Blair's students...use history as a way to make sense of their lives. (pp. 80-81)

It is critical that teachers of history articulate learning goals for students that include both content and thinking skills. How can they establish those goals? For both sets of goals, a state's curriculum standards are a good place to start. In addition, there are national standards for history, although these standards do not have the same widespread acceptance as standards in other areas such as mathematics or science¹. Good teachers, however, take standards primarily as a starting point; then they prioritize them and integrate them with what they know about their students' interests. (See box on "Unpacking State Standards" at the end of the chapter.)

Teachers may feel that the curriculum standards in their state, or the state tests that assess the standards, focus heavily on specific content—facts and themes—while giving short shrift to the thinking skills necessary to work effectively with the content. Some teachers may follow suit and include only multiple-choice items on their classroom assessments. However, the more recommended approach, by state departments and assessment experts, is for teachers to clearly identify their instructional goals for students, which should include teaching students to reason, analyze and critically evaluate historical information, and then match their assessment methods to instructional goals.

Most educators believe that, in order for students to retain information, they need to work with it in some way—far beyond traditional lecture or quick reading of a history textbook chapter. A research summary reported that having students generate and test hypotheses in history (or other subject areas) was related to achievement increases (Marzano, Pickering, & Pollock, 2001). Having students do more cognitively challenging and interesting work that involves using historical thinking skills to consider content may increase the retention of that content. Most teachers would agree with the need to find ways to engage students in critical thinking about content, not just learning facts. It is often the time, leadership, and resources for accomplishing this that are missing. Hopefully, this document will provide some needed support.

TO THINK ABOUT:

1. Why do you think students should study history?
2. How are you helping students understand those reasons?

¹ When the history standards were released, there was considerable debate over the content standards. There was minimal debate over the historical thinking standards. See Stern, 2000 for a discussion of the controversy surrounding the history standards.

The successful teaching of history is about much more than just having your students do well on the state test. Restricting learning goals in history to the memorization of facts or concepts reduces the richness of historical learning and perhaps, shortchanges students in terms of their long-term understanding of this country's institutions and our place in the world. The larger purposes behind the teaching of history are outlined in the National History Standards, which argue that history is critical for three primary reasons:

1. To maximize the potential of a democratic society.
“...*knowledge of history is the precondition of political intelligence...*without historical knowledge and the inquiry it supports, one cannot move to the informed, discriminating citizenship essential to effective participation in the democratic processes of governance and the fulfillment for all our citizens of the nation's democratic ideals.”

What does it mean to reason and think historically?

One of the key goals for history is to help students reason or think “historically.” Different groups stress different aspects of thinking; however, there are some commonalities. In particular, researchers and disciplinary organizations stress the ability of students to:

- Work with historical evidence;
- Understand the context of historical evidence and the perspective the original author(s) brought to bear; and
- Weigh the evidence to produce defensible interpretations of history (Van Sledright, 2004; Wineburg, 1999; National Centre for History Education, 2004).

The National History Standards presents five specific components of historical thinking. These or similar components are included in many state standards.

- I. Chronological thinking or the ability to understand time in relation to history.
- II. Historical comprehension, which includes understanding and evaluating historical documents. This includes understanding the context behind the document and the perspectives of individuals who wrote the document.
- III. Historical analysis and interpretation. This includes students being able to use the skills of historical comprehension to analyze historical evidence and draw conclusions based on that evidence. As part of this, students recognize that written history is tentative and changing.
- IV. Historical research capabilities, which includes the ability to formulate historical questions, and obtain and interpret historical information.
- V. Historical issues-analysis and decision-making, which includes the ability to examine issues in the past (and present) and make informed decisions about them.

The standards give specific examples of what each of these means.

2. To develop an understanding of the diversity of today's world. "Especially important, an understanding of the world's many cultures can contribute to fostering the kind of patience, mutual respect, and civic courage required in our increasingly pluralistic society and our increasingly interdependent world."
3. To nurture the individual. "Historical memory is the key to self-identity, to seeing one's place in the stream of time, and one's connectedness with all of humankind." (National Center for History in Schools, 1996, p. 13)

Samuel Wineburg (1999) builds on the second point about diversity, arguing: "...history holds the potential, only partially realized, of humanizing us in ways offered by few other areas in the school curriculum." He goes on to make the case that true historical understanding helps individuals recognize that people in the past thought and did things differently than we do today. "It is for this reason that the study of history is so crucial to our present day and age, when issues of diversity dominate the national agenda. Coming to know others, whether they live on the other side of the tracks or the other side of the millennium, requires the education of our sensibilities. This is what history, when taught well, gives us practice in doing." (p. 497)

Articulating Course Objectives

Teachers must identify learning goals that are both based in historical content and incorporate historical thinking (see box for discussion on historical thinking). One tool that helps teachers sharpen their thinking about this mix of content and thinking objectives is a matrix such as the one shown on page 10. (It is similar in concept to a test specifications matrix—a tool used by test developers to ensure that test items reflect all important objectives.) The historical thinking skills are listed across the top while the required content information is listed down the side. The teacher can use a matrix such as this in any number of ways. For example, in each cell, the teacher could fill in assessments to be used. Alternately, the teacher could describe instructional practices or activities that will help students achieve each goal. Each thinking skill may not be covered with each specific content area but they should be included several times throughout the semester. One way to get started in using this tool is by reflecting on a unit you have already taught. That is, think about how you taught and tested students in a specific history unit you recently completed (including the homework and class work assignments and where they fit).

Use a matrix like the one below to think about assessments that would meet both the content and thinking skills goals you have. Examples of different assessment tasks are included in the matrix. As can be seen, factual assignments or assessment tasks are the building block for each of the five content topics (rows). The historical thinking assignments are varied to fit the particular topic. Sharing such a matrix with students at the beginning of a unit may give them a better road map for how they will be examining the content.

**FIGURE 2-1:
U.S. History—High School Unit
Learning Goals Matrix**

GOAL*: TRACE ECONOMIC, POLITICAL AND SOCIAL DEVELOPMENTS FROM 1945-1980 AND ASSESS THEIR SIGNIFICANCE.	DESCRIBE CHRONOLOGICAL RELATIONSHIPS.	EVALUATE PRIMARY SOURCES.	USE MULTIPLE SOURCES OF HISTORICAL EVIDENCE TO DRAW CONCLUSIONS.	IDENTIFY, INTERPRET, AND SYNTHESIZE APPROPRIATE HISTORICAL EVIDENCE TO ANSWER QUESTIONS.	ARTICULATE AND DEFEND AN INFORMED POSITION ABOUT AN ISSUE IN HISTORY.
Effects of the Cold War	Develop timeline of main events in US- Soviet relations.	Read primary sources concerning U-2 incident (official pronounce- ments, notes from embassy). Identify the reasons behind the different stories presented by each.		Research and collect information to answer the question: “How would life today be different if the Cold War had never happened?”	
Civil Rights movement	Trace key developments of the Civil Rights movement.		Read primary accounts of first attempts to integrate schools. Create diary entry of person trying to integrate a high school.		Write a response to the question: “How successful was the Civil Rights movement?” Students will defend their answers using specific information from the time period and today.

GOAL*: TRACE ECONOMIC, POLITICAL AND SOCIAL DEVELOPMENTS FROM 1945-1980 AND ASSESS THEIR SIGNIFICANCE.	DESCRIBE CHRONOLOGICAL RELATIONSHIPS.	EVALUATE PRIMARY SOURCES.	USE MULTIPLE SOURCES OF HISTORICAL EVIDENCE TO DRAW CONCLUSIONS.	IDENTIFY, INTERPRET, AND SYNTHESIZE APPROPRIATE HISTORICAL EVIDENCE TO ANSWER QUESTIONS.	ARTICULATE AND DEFEND AN INFORMED POSITION ABOUT AN ISSUE IN HISTORY.
Social movements, including women, youth, and the environment	Develop a flow chart for each movement, highlighting key aspects and effects.			Choose a specific social movement. Ask a key question about the social movement. Find primary and secondary sources to answer the question.	
Vietnam War	Create a timeline describing key events during Vietnam War.	Interview community members who were alive during the war. Evaluate their memories relative to other documents.	Read multiple interpretations of causes of Vietnam War. Participate in a debate over the causes of the war.		
Technological innovations	Choose a technological innovation. Trace the history of the innovation.				Identify the most important technological innovation from 1945- 1980. Defend your choice.

**This example is based on North Carolina's eleventh-grade standards for U.S. History*

ACTIVITY:

Experiment with creating your own matrix for a unit or topic area detailing historical content you need to cover and historical thinking skills your students need to be able to use.

Unpacking State Standards

How do you know what the standards actually mean? Wiggins and McTighe (1998) in their work with *Understanding by Design*, suggest “unpacking” the standards. We’ll build on their work and apply it specifically to our discussion in this chapter of learning goals in U.S. history:

- Look at your curriculum goal or standard.
- Circle the nouns and put boxes around the verbs.
- Identify the “enduring understandings” or main concepts represented in the nouns or adjectives. (“Enduring understandings” are significant themes that appear over and over again. These can help you pick out what is really important about this standard. You can then frame this enduring understanding as a question to address in your lesson planning.)
- Identify specific historical knowledge students need to gain to develop these enduring understandings and meet the goal.
- Identify historical thinking skills students need to demonstrate to meet the standard.
- Identify the historical communication skills students need to display to meet the goal.

You are then ready to identify assessments that will work to assess the knowledge and skills that you want your students to gain. Let’s examine this with an eleventh-grade history standard from Alabama.

Alabama Learning Standard: 11th-grade Social Studies: United States History and Geography, 1900 to the present.

Learning Standard 2.) Evaluate the presidential leadership of Theodore Roosevelt, William Howard Taft, and Woodrow Wilson in terms of their effectiveness in obtaining passage of reform measures.

Examples: antitrust laws, labor reforms, income tax, stock-market regulation, conservation movement, economic reforms, Federal Reserve System

Step 1: Circle the nouns and adjectives and put boxes around the verbs. You end up with something that looks like this.

Evaluate the presidential leadership of Theodore Roosevelt, William Howard Taft, and Woodrow Wilson in terms of their effectiveness in obtaining passage of reform measures.

Examples: antitrust laws, labor reforms, income tax, stock-market regulation, conservation movement, economic reforms, Federal Reserve System

Step 2: Identify the enduring understandings or large concepts. In this standard, an enduring understanding might be “the impact of the president on social and economic issues.” A large question arising from this understanding might be: “What impact can presidents have on social and economic issues facing the United States?” You could use this question to organize your lessons concerning the administrations of these three presidents.

Step 3: Identify specific historical knowledge. Take a look at the nouns and adjectives in the standard. In this case, you know students need to look at the leadership of three specific presidents. Leadership is defined as the ability to get reform measures passed. Thus, you will have to discuss the reform measures, whether they passed, and the role of the president in getting those measures passed. The Alabama standard is helpful in giving specific examples of historical knowledge. In this case, that knowledge would include the social and economic reforms promoted by these presidents—Theodore Roosevelt, William Howard Taft, Woodrow Wilson—such as, antitrust laws, labor reforms, income tax, stock-market regulation, conservation movement, economic reforms, federal reserve system. All standards might not be so useful in providing specifics. If that is the case, you may want to consult other standards to identify a consensus on what specific historical knowledge is necessary for this enduring understanding during this time period.

Step 4: Identify historical thinking skills. Take a look at the verbs in the sentence. The primary verb is “evaluate.” Thus, students have to be able to take the knowledge about the three presidents and the economic and social reforms and determine how effective each of those presidents were at getting those reforms passed. This will involve cognitive skills such as comparing and contrasting different Presidents and their roles with the passage of reform measures. Students will also have to synthesize information to develop a specific interpretation (how effective each President was) and defend that interpretation with factual information on reform measures.

Step 5: Identify historical communication skills. When students are required to “evaluate” something, they need to make a case for a specific decision. For example, historical communication skills might include the ability to present a clear argument in favor of a particular point of view and provide a clear factually accurate defense of the argument. This could be done as an oral presentation, in a written paper, or as part of a multimedia presentation, but it is clear that it requires an extended response.

Unpacking this standard in this way helps you understand how to focus on the standards. It makes it correspondingly easier to identify the appropriate assessments to use to measure the skills and knowledge, which is discussed in **CHAPTER 3**.

METHODS OF ASSESSMENT

In **CHAPTER 1**, we discussed why assessment is important. **CHAPTER 2** focused on the identification of student outcomes in history or what students need to be able to do. In **CHAPTER 3**, we look at important assessment methods for history.

When most people think about assessment, they think of multiple-choice tests. Multiple-choice tests are the most frequent type of assessment, primarily because they are easy to grade (although good ones are not easy to develop) and presumably because they are objective². However, multiple-choice tests are not good for assessing everything. They are very good for some things, such as assessing specific factual knowledge, but they will not give you all the information you need to make decisions in your classroom. In addition, students on too much of a multiple-choice “diet” will not develop their writing skills, which they will need in their post-secondary work.

In order for teachers to have a good understanding of what their students know and are able to do, they must use a variety of assessments. The method of assessment depends on the learning goals you have for the students.

Stiggins (1994) suggests that there are four primary types of classroom assessments: selected response, essay response, personal communications, and performances. We have adapted these categories to make them particularly appropriate for history:

- Selected response—The most common type of assessment, these include multiple-choice, true false, and matching questions, as well as short answer questions where students select an answer from a word-bank.
- Dialogue and Oral Responses—This includes conversations, seminars, question and answer sessions, and interviews.
- Essays and open-ended responses—These are extended open-ended written responses, such as essays. They also may include creation of a graphical organizer or a diagram as well as short answers with no word bank.
- Projects and investigations—Students engage in in-depth activities to explore and respond to a particular situation.

² Some groups today argue that multiple-choice tests may be objective in their scoring but they are certainly not objective when they are created (Wiggins, 1990; National Center for Fair and Open Testing, n.d.).

FIGURE 3-1 matches up the types of assessments with the historical understanding goals described at the beginning of **CHAPTER 2**.

FIGURE 3-1: Matching Desired Student Outcomes with Assessment Methods

	SELECTED RESPONSE	DIALOGUE AND ORAL RESPONSES	ESSAY AND OPEN-ENDED RESPONSES	PROJECTS AND INVESTIGATIONS
Knowledge of historical facts, themes, and ideas	Good match	Good match	Can be a good match, depending on the type of question	Not a good match
Historical thinking and reasoning (chronological thinking, comprehension, analysis and interpretation, research, issues-analysis and decision-making)	Can be a good match depending on the type of question	Good match	Good match	Good match
Communication of historical knowledge to a wider audience	Not a good match	Good match	Good match	Good match

Adapted from Stiggins & Knight, 1997.

As the table above shows, selected response assessments are primarily good for assessing specific knowledge, although they may be able to assess some levels of historical thinking. Therefore, if you want to determine whether students know specific information, a multiple-choice or short answer test would be a good way to go. If, however, you also want to assess students' ability to analyze historical evidence, a multiple-choice exam would be a poor choice.

Since multiple-choice tests already abound and are used by many teachers, and since the other three methods are actually more flexible at assessing a variety of learning goals, we will focus our discussion on the three most informative assessments for history instruction:

1. dialogue and oral response,
2. essays and open-ended responses, and
3. projects and investigations.

We discuss each of these types of assessments, their purposes, and how to create a quality assessment. **CHAPTER 4** focuses on the effective use of checklists and rubrics with these types of assessments.

Before we focus on each of these in more detail, it is important to highlight one thing. As you read through the description of each assessment, you may think that these are really instructional tasks. They may be. The

tie between instruction and assessment should be very close. What makes an instructional task an assessment is if you intend to use information from the task to get information about a student's performance.

1. Dialogue and Oral Response

What they are: Dialogue and oral responses may actually be one of the most common ways of assessing students informally. These include classroom discussions, responses to questions, as well as more formal structures such as classroom presentations or debates. Although teachers may use these as informal assessments, there are also ways of making that information more useful and more systematic for formal assessment purposes.

Purpose: Dialogue and oral responses can be used to check student knowledge. Teachers frequently ask questions to check for understanding of specific content. Less frequently do they use these strategies to assess students' thinking.

The type of dialogue and response you set up depends greatly on the purpose you have for it. If you are assessing students' factual knowledge, quick question and answer type responses are fine. To keep track of students' responses, you may want to develop a quick checklist like the one in **FIGURE 3-2**. If, however, you are trying to assess students' ability to use historical reasoning, then you will want to think much more carefully about the types of discussion you have or the presentations and debates you establish.

FIGURE 3-2: A Discussion Checklist for Factual Knowledge on European Explorers

Mark a check next to each student's name when a comment demonstrates knowledge of a specific fact or concept.

STUDENT NAME	PRINCE HENRY, THE NAVIGATOR	COLUMBUS	CORTES	PONCE DE LEON	DE SOTO

Creating High-Quality Dialogue and Responses

The basis of quality discussions, or what Lewin and Shoemaker (1998) call "substantive dialogue" is quality questioning. Good discussions revolve around real questions of interest. Students know when questions are designed to lead them to the teachers' answer or when they are genuine opportunities to explore and learn. The former may lose their interest quickly while the latter can engage

(Based on Alabama Standards)

them over time. In addition, high-quality dialogue can be shaped around a text. The text can be a letter, a speech, a diary entry, a cartoon, a piece of art, or a poem. This text gives students something they need to interpret and to which they need to respond. It anchors the discussion and dialogue in something solid. Discussion and dialogue around text is the basis of work such as the Paideia seminar (www.paideia.org).

“I hate it when teachers ask questions they already know the answer to. That’s not a real question.”

—high school student

As you develop good questions for your students, you need to think about the following issues (Dantonio, 1990):

1. What is the content focus of the question?
2. What is the process focus of the question?
3. Is the question open-ended, including the opportunities for multiple responses? (In quick content knowledge and recall questions, this may not be applicable.)
4. Is the question clearly stated?

Questions can be used to assess different thinking processes. **FIGURE 3-3** shows historical thinking processes you might want to assess, questioning stems for those processes, and sample questions for an elementary history unit on the Constitution and the United States and then for a high school unit on the Federalist Era. Please note that there will be substantial overlap in the questions. Thus, a question that revolves around historical issues-analysis will also include components of historical knowledge, comprehension and other thinking skills.

FIGURE 3-3:
Questioning for Historical Thinking

HISTORICAL THINKING SKILLS	SAMPLE QUESTION STEMS	MODEL QUESTIONS
Content Knowledge— Recalling facts or concepts, justifying answer with supporting details	What can you remember about _____? Describe _____. List _____. Why did you say this? How do you know?	Elementary: Describe the Bill of Rights and its purpose. High School: What is the concept of judicial review? How was it established?
Chronological Thinking— an understanding of when events occurred and in what temporal order, an understanding of narrative flow and the ability to see patterns	Place these events in correct order. What led to _____?	Elementary School: What led to the writing of the Constitution? High School: How long was it until the United States could function as a democratic country? Why would you say that?

HISTORICAL THINKING SKILLS	SAMPLE QUESTION STEMS	MODEL QUESTIONS
<p>Historical Comprehension—students' ability to understand and make use of historical narratives, documents or artifacts</p>	<p>Summarize what you read....</p> <p>What is this person saying?</p> <p>Who is the author? Assess his or her credibility.</p> <p>How did the historical situation contribute to the author saying this?</p>	<p>Elementary: What does the First Amendment say? Why was this written?</p> <p>High School: What is Benjamin Banneker saying about the idea of equality in the early United States? Why is he saying this?</p>
<p>Analysis and Interpretation—the ability to use historical information to develop an interpretation of why or how things happened or to develop theories for possible alternatives</p>	<p>Compare and contrast.....</p> <p>What would have happened if _____?</p> <p>How is this (historical experience) like another (historical experience)?</p> <p>Whose interpretation of (this historical event) is (better, more accurate, more believable)? Why?</p> <p>Which part of this narrative is fact and which is interpretation?</p>	<p>Elementary: Compare and contrast government under the Articles of Confederation with government under the Constitution.</p> <p>High: How is the formation of new democracies in the world today the same as the start of the United States? How is it different?</p>
<p>Historical Research—the ability to ask historical questions, find, and examine historical information, and construct a sound interpretation</p>	<p>What new questions do you have after learning this (reading this, viewing this)?</p> <p>What do you want to know more about?</p> <p>How would you find the answer to those questions?</p> <p>What story would you tell about _____?</p> <p>What conclusion did you reach from your research? Why?</p>	<p>Elementary: What more do you want to know about life in the new United States of America? How would you find answers to your questions?</p> <p>High: Tell a story about what you found out about the interactions between American Indians and European settlers in the early 1800s. Be ready to defend your version of the story.</p>
<p>Historical Issues—Analysis and Decision-Making—the ability to investigate the dilemmas of history, identify the historical antecedents of current social issues, and evaluate results of different decisions</p>	<p>What did people (in this time period) think about _____?</p> <p>Why did this (controversial event) happen? What would you have done if you had been there?</p> <p>Why does (this situation) exist today?</p> <p>Describe the different choices people have today in ways they respond to (particular situation). Which would you recommend? Why?</p>	<p>Elementary: Why did many of the founding fathers have slaves?</p> <p>High School: How did the early founders see the relationship between religion and government? How do these views influence the debate today?</p>
<p>Historical communication skills</p>	<p>Any question will help you assess this.</p>	

Preparing Students to Engage in Dialogue and Discussion

Just as with writing, students need to be taught how to engage in substantive dialogue and respond to complex questions. Research has shown that about half the time there is no relationship between the complexity of the question and the complexity of the answer that follows it (Strother, 1989).

There are several ways teachers can increase the quality of students' responses:

1. Limit the amount of talking done by teachers. For dialogue to play an important part in students' cognitive growth, the teacher should talk no more than 30% of the time (National Paideia Center, 2003).
2. Teachers need to follow up with students and probe their responses. Asking them to build on their answers will increase the quality of the responses. In addition, authenticity may help students think more about their responses.

Things to be careful about:

1. Questioning and dialogue are such natural occurrences that teachers don't think much about them. Yet, most discussions revolve around factual and recall level questions (Strother, 1989). If you want to use dialogue and discussion effectively to assess higher order thinking skills, you need to prepare for it. You can develop questions ahead of time and run them by your peers.
2. True dialogue and discussion occurs when people listen to each other. Students may need to be taught to listen and encouraged to take risks in responding. Teachers also need to be careful to listen to the students themselves and restrict the amount of time they talk.
3. It is easy for some students to dominate the discussion. Teachers can use various strategies to make sure everyone is involved, including requiring everyone to make some sort of a comment or drawing names out of a hat or tossing a ball around for comments. Placing students in small groups can also encourage more discussion.

Practice Developing Good Questions to Guide Discussion

With a partner, practice developing some high-quality questions. Complete each step:

- Step 1:** Establish the content focus of the question. What do you want your students to know and be able to discuss?
- Step 2:** Establish the process focus of the question. What thinking skills do you want them to demonstrate? You also may want to build certain process skills, such as explaining, debating, or justifying an answer with evidence.
- Step 3:** Make sure the question is open-ended and includes opportunities for multiple responses. Imagine some of the responses students could give. Would every student give the same response or does the question allow for different responses?
- Step 4:** Make sure the question is clearly stated. Simple is generally better here.
- Step 5:** Develop follow-up prompts in case the initial question doesn't work.
- Step 6:** Consider the best context for the discussion. Is it whole class? Partners? Small groups? What will you do with responses if it is done in partners or small groups?

2. Essays and Open-Ended Responses

What they are: Essays and open-ended responses are responses written to a prompt of some sort. They can ask students to interpret primary sources, to compare different kinds of information, to draw a diagram illustrating relationships, or to solve a problem. They are useful for encouraging students to organize and consolidate their thinking in a way that goes beyond remembering events. A really good essay question is like a puzzle where you're trying to help students put some pieces together. The document-based questions that show up on Advanced Placement exams and some state tests are examples of good, thoughtful essay questions (sample AP questions can be found at www.collegeboard.com and sample document-based questions at different levels for New York State can be found at <http://emsc32.nysed.gov/osa/>).

Purpose: Essays and open-ended responses are good at assessing students' reasoning and their ability to communicate historical knowledge. They also can be good at assessing students' historical knowledge, depending on the type of question that is asked. Since the most common type of open-ended response is an essay question, this section will focus primarily on essay questions.

There are two main kinds of essay questions: Restricted Response and Extended, Open-Ended Responses (Kubiszyn and Borich, 1996). Restricted response essay questions are really an extended version of

the short answer. They ask students to recall specific factual information and can be used to assess students' knowledge, or application or comprehension of information. These questions generally begin with verbs such as "Describe," "List," or "Summarize." Some restricted response essay questions might include:

- Describe the monarchical system of government and give three examples.
- Summarize the provisions of the Treaty of Ghent.

These types of questions do not assess the more complex historical thinking or reasoning goals. Why use these instead of multiple-choice tests? Criswell and Criswell (2004) present three reasons:

- The written expression reveals more of students' underlying logic.
- A student cannot guess the correct answer. A certain amount of guessing or bluffing certainly could be part of the essay process, however.
- The teacher believes that students should be able to recall the answers rather than just identifying it.

All of these justifications can also be reasons to use open-ended short answer questions instead of short answers with a word bank.

The real power of essay questions comes in the use of extended, open-ended responses, which can measure students' ability to engage in higher order thinking, such as their ability to synthesize and evaluate information. These next several sections will address creating high-quality, extended, open-ended responses.

Developing Quality Essays and Open-Ended Responses

Good essay questions and other open-ended responses are actually harder to write than most people realize. Nevertheless, they are well worth the effort since they provide a window into students' thinking. In addition, writing is also a key method of communicating information and students must be able to communicate content information with clarity and precision.

Stiggins (1997) highlights three steps to take as you are creating high-quality essay questions:

1. Specific desired content knowledge. You need to clarify what it is exactly you want students to know. What is the content that students are expected to discuss?

A not so good example: Write everything you know about the Black Death.

(This is much too broad of a topic. Also, technically, there is no way a student can get this wrong since you are asking them to write everything they know. If they know nothing, they write nothing.)

A much better example: In the 1300s, the Black Death decimated one-third of Europe's population. Describe the impact of the Black Death on various aspects of European society, including food supplies, political structures, and religious and philosophical beliefs.

2. Specific kind of historical thinking, reasoning and problem-solving you want students to demonstrate. If you want to maximize the potential of essays, you will want to ensure that the reasoning is focused on higher-order thinking skills.

A not so good example: List three restrictions put on African-Americans in the South after the Emancipation Proclamation.

(Although this very clearly specifies the knowledge to be included, this is really a recall question and does not require any higher order thinking. It may be good as a homework or class work item rather than as a test item.)

A much better example: After the Emancipation Proclamation, many states in the South placed formal and informal restrictions on the lives of African-Americans. Identify three restrictions. For each restriction:

- Describe the restriction.
- Identify the purpose behind the restriction.
- Evaluate the impact of the restriction on African-Americans at the time.
- Use your knowledge of the African-American experience today to suggest some possible long-term impacts of this restriction.

(This much more detailed question includes a mix of lower-order thinking skills—simply recalling the restriction and possibly identifying the purpose of the restriction—with higher-order thinking—evaluating the impact of the restriction, making a connection to life today).

3. Guide students in the direction of an appropriate response without spelling out the answer. The question needs to be specific and provide enough guidance to students so that they understand what they are supposed to write or do:

A not so good example: Describe the role of religion in the early United States. (This is a very typical essay question. It is too broad, however, and provides little guidance to the students.)

A much better example: **The Role of Religion in the New Republic:** You are member of a state legislature in 1800. Someone has proposed a law that would allow the state government to purchase lands for churches. You know that the issue of the appropriate role of religion has been an important struggle in the early years of the United States. You also know that many states at this time period actually have provided official support to specific religions. To inform your decision, you have read, among other pieces, the works of two of the founding fathers, George Washington and James Madison (see below).

1. Summarize Washington's and Madison's perspectives on religion and government. In your summary, bring in additional information from the time period to describe why the writer might have this perspective.
2. How will you vote? Defend your decision, using the texts and other information you know from the time period. Remember, you are legislator in 1800, not today.

3. Now, pretend you are a legislator today voting on legislation authorizing the display of the Ten Commandments in government building. What would your decision be? What justification would you give? You must at least partly defend your decision by using historical information from the early Republic. You may also bring in more recent information as appropriate.

George Washington. *Farewell Address, 1796*

Note: This was written as George Washington completed his second term as President.

Of all the dispositions and habits which lead to political prosperity, religion and morality are indispensable supports. In vain would that man claim the tribute of patriotism who should labor to subvert these great pillars of human happiness — these firmest props of the duties of men and citizens. The mere politician, equally with the pious man, ought to respect and to cherish them. A volume could not trace all their connections with private and public felicity. Let it simply be asked, Where is the security for property, for reputation, for life, if the sense of religious obligation desert the oaths which are the instruments of investigation in courts of justice? And let us with caution indulge the supposition that morality can be maintained without religion. Whatever may be conceded to the influence of refined education on minds of peculiar structure, reason and experience both forbid us to expect that national morality can prevail in exclusion of religious principle.

James Madison. *Memorial and Remonstrance Against Religious Assessments 1785*

Note: This was written in opposition to a bill proposed by the General Assembly of Virginia to collect funds for teachers of religion. The bill was not passed.

...The Religion then of every man must be left to the conviction and conscience of every man; and it is the right of every man to exercise it as these may dictate. This right is in its nature an unalienable right. It is unalienable, because the opinions of men, depending only on the evidence contemplated by their own minds cannot follow the dictates of other men: It is unalienable also, because what is here a right towards men, is a duty towards the Creator. It is the duty of every man to render to the Creator such homage and such only as he believes to be acceptable to him....

...it is proper to take alarm at the first experiment on our liberties. We hold this prudent jealousy to be the first duty of Citizens, and one of the noblest characteristics of the late Revolution. The free men of America did not wait till usurped power had strengthened itself by exercise, and entangled the question in precedents. They saw all the consequences in the principle, and they avoided the consequences by denying the principle. We revere this lesson too much soon to forget it.

...(the) proposed establishment is a departure from the generous policy, which, offering an Asylum to the persecuted and oppressed of every Nation and Religion, promised a lustre to our country, and an accession to the number of its citizens. What a melancholy mark is the Bill of sudden degeneracy? Instead of holding forth an Asylum to the persecuted, it is itself a signal of persecution. It degrades from the equal rank of Citizens all those whose opinions in Religion do not bend to those of the Legislative authority. Distant as it may be in its present form from the Inquisition, it differs from it only in degree. The one is the first step, the other the last in the career of intolerance.”

(This question provides much more detailed guidance on what is expected. Providing information on how the essay will be scored can also help in providing guidance.)

To Do

**Practice writing essay questions.
Pick a topic area from your history class.
Follow the three steps described above:**

1. Specific desired content knowledge. What is the content knowledge you want students to explain? Think about the information you have covered in class.
2. Specific kind of historical thinking, reasoning and problem-solving you want students to demonstrate. What kinds of historical thinking skills do you want to assess? Do you want them to be able to read and interpret primary sources? Do you want them to make an argument and back it up with historical facts?
3. Guide students in the direction of an appropriate response without spelling out the answer. What would the structure of a good answer look like? Try to produce a variety of answers students might develop and see if you need to revise your question to better point to the appropriate level of detail.

Preparing Students to Write Essays

Needless to say, students need to be taught how to answer these questions. You cannot just provide students with an essay question as an assessment and expect that they will understand what to do. Instead, you must carefully scaffold instruction to get students ready to answer essay questions. Since answering questions in clear writing supported by details is a critical skill in history (or any field), this is a case when teaching to the test is actually a useful instructional activity.

Below are some steps for scaffolding students' writing skills. For a more complete description, consult some of the resources in the bibliography.

1. Have students write frequently. Get them used to writing, even if it is just short answers or brief reflective writing. Each piece of writing does not need to be large and complex.
2. Have them read what they write to one another to get them used to writing for an audience.
3. Focus on 2-3 specific writing skills at a time. Don't try to teach students everything they could possibly need to know about writing at one time. This will make it easier for the students and for you as you look at their work (McCoy, 2003).
4. Develop a rubric or checklist for assessing the writing. Show students how to use the rubric using "anchor" papers or samples of real student work to show them examples of good historical communication as well as examples of less successful work (Lewin and Shoemaker, 1998).
5. Give struggling students specific feedback about their responses on specific questions. Have them rewrite based on the feedback.

Things to be careful about:

Essay questions are not the perfect assessment. Below are some potential pitfalls in the use of essay questions.

1. Don't confound writing proficiency with reasoning skills. In many cases, a well-written essay will receive a high score, regardless of the content of the essay. It is helpful, in these cases, to make sure that you have separate scores for communication or writing skills and for content knowledge or reasoning skills. We'll investigate this in more depth in the next chapter.
2. Be realistic about the resources you have. Scoring essays will require more time than multiple-choice tests. Make sure you build in enough time for grading and giving feedback with the essays. If you have many students, you may want to stagger essay assignments so you don't have a large amount to grade at the same time. In all cases, you should think carefully about what you want to accomplish with the essays so that they are well worth the time and effort.
3. Make sure that expectations are clear up front. Essay questions can be open to bias, particularly in the scoring process, so make sure that the scoring expectations are clear from the beginning. We'll talk more about that in the next chapter.
4. Don't give away the answer, such that every student has exactly the same response. Make sure the question and the expectations are general enough to allow for substantive thought on behalf of the students.

3. Projects and Investigations

What they are: These are in-depth projects or investigations that have students consider a complex issue over multiple days. Projects or investigations are often done as the culminating experience of a unit and are designed to allow students to synthesize information and skills together. Projects and investigations may take many different forms. Research projects and presentations are popular with many teachers. Other projects may ask students to solve an ill-defined problem connected to the area of study. They may ask students to research and prepare for a debate. They may have students research and prepare a re-enactment of a particular event.

Purpose: As indicated in **FIGURE 3-1** at the beginning of this chapter, projects and investigations are useful for assessing students' historical reasoning and historical communication skills. Projects and investigations are generally not as useful for assessing specific content knowledge.

The real strength of projects and investigations is that they provide one of the few opportunities to assess students' work in a context more closely resembling the adult world or the world of work. For example, there are very few times when adults in the workforce are required to take multiple-choice tests (drivers license and credentialing exams being notable and rare exceptions); however, there are numerous times when people need to make sense of data, do presentations based on research, or write reports. Well-designed projects and investigations provide the opportunity to engage in the sorts of complex thinking, problem solving, and communicating that are required in many workforce situations. They also provide an opportunity for teachers to assess any goals they may have in the affective and social domains.

Developing Quality Projects and Investigations

Just like essays, projects and investigations require some investment of time to do well. To truly take advantage of projects, they should tap into some of those higher-order thinking skills. A popular research project that asks students to go onto the Internet and find 5-10 facts about an individual is actually fairly lower level, since it simply requires students to summarize factual information. Below are some things to think about as you plan the use of projects and investigations for assessments.

Step 1: Determine the goals and objectives of the project. What are you trying to assess through this project? As with any type of assessment, you need to start by identifying the learning objectives you want to assess. Identify the content area you are going to include in the project. Look at the types of historical thinking skills you want students to demonstrate. Identify any objectives in the affective or social domains you may want to assess. For example, you may want to look at the ability of students to work cooperatively in a group. You may want to assess their persistence in the face of a challenging assignment. These are all valid objectives for a project or an investigation. After you have identified the objectives for the assessment, you need to design a project or investigation that aligns with those objectives.

Step 2: Determine the substance of the project. Coming up with a good project or investigation does require some creativity. The key, however, is thinking about the real world of historians and the type of work in which they are engaged (Wiggins, 1992). What is the actual work of historians? Historians examine multiple primary sources from the time period. They consider the work of other historians. They analyze and synthesize data to create a story that answers a question. They present their conclusions to other historians, who may or may not believe them. They defend their conclusions or revisit those conclusions in light of new information. A high-quality project and investigation will thus incorporate these characteristics into the assessment task itself.

Step 3: Lay out the specific components of the task, including the actual work students will do and the project's scope and duration. Projects and investigations should be of a reasonable scope and duration so that everyone has some chance of success. This will come with practice. It might be helpful to develop a day-by-day guide or checklist showing how students' time will be spent.

Step 4: Develop the assessment criteria and provide them to students up front. Students will know what they are trying to achieve and what they need to do to do quality work. It may also be useful to show students actual examples of quality work so that they have something to judge. We'll talk more about this in the next chapter.

FIGURE 3-4 shows the process of developing an historical investigation assessment task.

RESOURCES

Web Resources for History Projects

<http://www.college.hmco.com/education/pbl/> Houghton Mifflin's site on project-based learning is designed to be used with their texts, but it also can stand alone.

www.edutopia.org This site from the George Lucas Educational Foundation includes a topic on project-based learning with research references and connections to schools implementing projects.

www.bie.org Buck Institute provides professional development and online resources on project-based learning.

FIGURE 3-4: Developing an Historical Investigation for a Unit on the Industrial Revolution (Steps 1 through 3) — A Teacher’s Thought Process

Step 1: Determining the content. According to my curriculum standards, my students need to learn about “the transformation of the American economy and the changing social and political conditions in the United States in response to the Industrial Revolution” (California, grade 8). In particular, they are supposed to understand the implications of industrial revolution (development of cities, conservation, immigration) as well as the growth of the big business and the labor movement. I spend some time exploring the History Matters website (www.historymatters.gmu.edu), which has a great collection of primary sources and activities, as well as links to other websites. I see that there are a lot of interesting primary source documents out there on the labor movement, sweatshops, urbanization, immigration, and the conservation movement. Since these are all issues that also resonate today, I believe they will be good areas for investigation. I want the project to allow them to investigate one of these areas in depth.

Step 2: Determining the process skills. My curriculum standards also say that students should be able to make connections between what happened in the past to social and economic conditions today. Thus, I want the project to focus on understanding the issues in the past and connecting them to what we see today. Because I am familiar with expectations of historical thinking, I also want students to be able to examine primary sources and make conclusions based on those primary sources.

Step 3: The components of the investigation. The overall structure of the investigation should involve looking at the historical aspects of various pieces of industrialization that are represented today. I want to include choice so each student will choose an aspect he or she is interested in and explore its historical roots, making connections to the current situation and then perhaps developing recommendations. We’ll work towards having a summit of prominent thinkers (the students) doing presentations on: Understanding the Past to Prepare for the Future.

Here is what each student will do:

1. Each student will pick a topic out of the following:
 - a. High Gas Prices? Why the U.S. Relies on Fossil Fuels
 - b. Urban and Suburban Sprawl: How Cities Came to Be
 - c. Mass Transit: Transportation and its Impact
 - d. Migrant Workers: Immigration and the Economy
 - e. Workers’ Rights: The Labor Movement

- f. Money and People: Capitalism and Social Concerns
 - g. Industry and the Environment
2. Each student will begin by identifying a specific real situation that exemplifies this topic today. They must start with at least two documents that present two different perspectives today.
 3. The students will then trace the history of their issue from 1867 to 1914. What was happening connected to the issue in that time period? Students will examine their textbooks and summarize how the issue is addressed in the textbook. They will also identify at least two primary sources from the time period that represent different perspectives. They will compare the primary sources with each other and with the textbook to develop their interpretation of the main aspects of that issue during that time period.
 4. Using the information, they will develop a presentation and supporting paper that presents a historical understanding of the situation today. They will then develop recommendations for resolution of this issue based on their understanding of the historical situation.

Preparing Students to do Quality Projects and Investigations

Just like any of the other assessment methods, students need to be taught the steps involved in doing quality projects and investigations. You can't assume that students will know how to do research or how to compare and contrast things. Fortunately, this instructional process can easily be built into the development of an ongoing project or investigation. One way to do this is to clearly identify each of the steps necessary to do the project. The first time they do a project or investigation, the students may need very specific help from identifying questions they need to ask to identifying sources to creating first drafts of the product. It is better to err on the side of too much detail at the beginning and then gradually allow the students more ownership in the project as they become more used to the process.

Things to be careful about:

- Many teachers assign large projects or investigations to students to do outside of class time. This can result in wonderful projects completed by parents or not-so-wonderful projects created by students struggling without resources at home. Teachers can avoid this by having students do the majority of the work during class time, with small specific assignments to be done at home.
- Projects and investigations can be a time sink. How many times have you started to do something that you thought

was a good idea and then never had the students finish it because you ran out of time? Planning the projects carefully so that they are accomplishing and assessing key instructional goals can help minimize this issue.

Assessing your Assessment Tasks

As you begin developing assessment tasks such as those described above, Herman, Aschbacher, and Winters (1992) present some criteria to consider. These criteria apply to all types of assessments.

- Do the **tasks match the important outcome goals** you have set for students? Do these goals reflect complex thinking skills, such as analysis and synthesis?
- Do they pose an **enduring** problem type—the types of problem and situations that students are likely to face repeatedly in school and their future lives?
- Are the tasks **fair and free of bias**? For example, do they favor either boys or girls, students who have lived in a particular location or region, students with a particular cultural heritage, or those whose parents can afford to buy certain materials?
- Will the tasks be **credible** to important constituencies? Will they be seen as meaningful and challenging by students, parents and teachers? Do the tasks rely on quality subject matter content?
- Will the tasks be **meaningful** and engaging to students so that they will be motivated to show their capabilities? Do the tasks involve real problems, situations, and audiences? (pp. 41-42, emphases in original)

ESTABLISHING CRITERIA

“In the absence of criteria, assessment tasks remain just that, tasks or instructional activities. Perhaps most important, scoring criteria make public what is being judged, and, in many cases, the standards for acceptable performance. Thus, criteria communicate your goals and achievement standards.” (Herman, Aschbacher & Winters, 1992, p. 44)

In the previous chapters, we have described some assessment methods that are well suited to examining students’ ability to do history. These methods—dialogue and oral responses, essays and open-ended responses, and projects and investigations—can provide rich information on what students know and are able to do. Yet, as indicated in the quote above, an essay is not necessarily an assessment task; it could be just an instructional activity. What makes an essay (or anything else) an assessment is the establishment of criteria against which the work is measured and the use of those criteria to obtain information about a student’s performance. Even dialogue, which may be assessed more informally, should have some explicit criteria.

Having clear criteria in assessments is important for several reasons. A meta-analysis of multiple studies concluded that the use of clear learning goals and objectives is related to increases in achievement (Marzano et al, 2001). These learning goals and objectives should be articulated in the assessments teachers use. Teachers may assume their assessments go beyond superficial or rote learning goals and reflect their goal of deep understanding; but when examined more closely, they may find this not to be the case. Developing specific criteria can help teachers use assessments that better align to their learning goals. Another large benefit to developing specific criteria is that students become aware of the expectations and are increasingly able to monitor their own learning (Shepherd, 2000).

Having explicit criteria is particularly important for assessments that are not multiple-choice. Multiple-choice items can be scored objectively. The student is offered a fixed number of options and the option selected is compared to a scoring key. Given the scoring key, any teacher would score the items in the same matter. The methods we discuss in this manual—essay and open-ended responses; dialogue and oral responses; and projects and investigations—depend to a greater extent on teacher judgment of a response. Thus it is critical to develop good scoring criteria that:

- “Help teachers define excellence and plan how to help students achieve it.
- Communicate to students what constitutes excellence and how to evaluate their own work.
- Communicate goals and results to parents and others.
- Help teachers or other raters be accurate, unbiased, and consistent in scoring.
- Document the procedures used in making important judgments about students.” (Herman, Aschbacher & Winters, 1992, p. 48)

SCORING CRITERIA

Establishing the performance criteria up front is also exceptionally useful in the planning process. First, it forces you to decide on the actual goals for the activity and what you will stress. Second, it helps focus the assignments more and avoid the “creep” that can happen to projects or activities.

In summary, scoring criteria must:

1. Reflect the goals you are trying to assess; and
2. Communicate expectations in advance.

To explore what happens when these characteristics are not in place, examine the scenario in the box. It is also important to think about the other extreme in the use of criteria, which is also not helpful to students’ development. In this other extreme scenario, a teacher may go too far and actually give the answer to essay questions in advance. Thus, students know “what the teacher expects” in terms of the answer but, because they have been given the actual answer and only have to memorize it, there is no critical thinking or deep understanding being developed.

The Importance of Alignment and Communication

To explore the relationship between an assessment task and the criteria used to assess performance on that task, imagine the following scenario.

You are taking a course at a local college on assessment. Your instructor gives you the following assignment: “Write one sentence in response to the following question: Why is it important to match your assessments with your learning goals?” You try very hard to include meaningful information in a concise manner.

You turn in your paper and receive it back within minutes, a red “F” stamped across it. “But,” you sputter, “I answered the question.” Then the instructor shares the grading criteria with you:

1. The sentence incorporates three capital letters.
2. The sentence includes exactly four pronouns.
3. The sentence was written in green ink.

An “A” meant all three criteria were met; a “B” meant two criteria were met; a “C” meant one criterion was met. An “F” meant no criteria were met, thus your grade of an F.

In thinking about this somewhat unlikely scenario, ask yourself the following questions:

1. Did the criteria assess the presumed learning goal behind asking the question?
2. Based on the assessment criteria used to grade your sentence, what were the actual assessed learning goals? What does this tell you about selecting criteria for assessments?
3. As bizarre as they might have been, if you had known the criteria in advance, could you have written a sentence that enabled you to get an A? What does this tell you about communicating expectations in advance?
4. If these criteria had been communicated to you in advance, what would have been the focus of your sentence-writing? What would you have actually learned from writing the sentence? What does this tell you about the relationship between assessments and learning?
5. Do you think your students might ever feel the way you would have if you had been in this situation?

There are at least three ways to arrange criteria:

1. Checklists
2. A Point System
3. Rubrics (Holistic and Analytic)

Different arrangements of criteria work in different situations. We'll discuss each of these arrangements and their most appropriate uses in this chapter.

1. Checklists

Checklists are the simplest of the different methods of examining students' work. A checklist can be used to indicate that a student has effectively completed the steps involved in a task or demonstration. Checklists may be used to track completion of steps in a project, record observable behaviors in a discussion group, or identify the existence of specific components in an essay. In general, checklists are useful for simply determining the presence or the absence of something. They are also useful in helping students assess themselves or determine whether they are on track. Checklists are not as useful in situations where there can be partial levels of accomplishment or when you need to determine the relative quality of students' work. Below are two specific examples of how checklists could be used in a history class.

- A.** Discussion on topics or questions is commonly expected of students in their history classrooms. However, it may not always be clear to students exactly what constitutes "good" discussion behaviors. To assess students and provide them more guidance, checklists can be helpful. Teachers can record students' responses to specific questions, as the example on page 16 showed. Checklists can also be helpful as a debriefing tool to encourage students to take responsibility for their behavior. For example, **FIGURE 4-1** shows a checklist developed by the National Paideia Center for helping middle grade students self-assess their participation in a seminar. In addition to the checklist, there are also open-ended questions that ask students to reflect on their contributions to the seminar.

MY POSITIVE BEHAVIORS:	MY NEGATIVE BEHAVIORS:
<input type="checkbox"/> I was courteous to others.	<input type="checkbox"/> I interrupted others.
<input type="checkbox"/> I paused and thought before speaking.	<input type="checkbox"/> I did not look at the person speaking.
<input type="checkbox"/> I listened to others tell their ideas.	<input type="checkbox"/> I raised my hand to speak.
<input type="checkbox"/> I kept an open mind concerning what was said.	<input type="checkbox"/> I did not speak at all.
<input type="checkbox"/> I used information from the text to support my statements/questions.	<input type="checkbox"/> I talked too much.
<input type="checkbox"/> I felt comfortable speaking in seminar.	<input type="checkbox"/> I talked about subjects other than the text.
<input type="checkbox"/> I gave my opinions clearly.	<input type="checkbox"/> I did not listen to others.
<input type="checkbox"/> I acted as a positive role model for others.	<input type="checkbox"/> I was discourteous to others.
<input type="checkbox"/> I was prepared for the seminar.	<input type="checkbox"/> I was not prepared for the seminar.

Source: National Paideia Center, 2003, Used by permission

- B.** Debates and project work are often used to help students develop in-depth knowledge about a critical issue in a time period. Such long-term projects often require many steps to completion. Checklists that lay out all the steps along the way can help students keep track of the work they are doing. **FIGURE 4-2** provides an example of a checklist for a typical research project in middle school.

FIGURE 4-2:
Research Project Checklist

TASK	DATE COMPLETED
Selected topic	
Developed four questions to ask about the topic	
Identified three different sources of information	
Wrote answers to questions on note cards	
Planned report using concept map	
Wrote first draft of report	
Met with fellow student to edit	
Revised report	
Word processed final report and turned into teacher	

2. A Point System

A point system assigns points for certain features of the student's response. Open-ended questions or projects are often scored with this approach because points can reflect partial as well as full credit. **FIGURE 4-3** shows an example of how one teacher used a point system to evaluate an essay written for a final exam.

FIGURE 4-3 Using Points for Essays

Assignment: It is February 1968. The US and South Vietnamese forces have pushed back the VC forces after the Tet Offensive. But the political damage has been done. You are a top policy advisor to the President. Draft a memo to President Johnson outlining what policies the United States has used to win the war in Vietnam - which ones have worked, which ones have not worked. Then, advise the President on what course of action he should take: Should the United States withdraw from Vietnam, increase and widen the war, or maintain the current policies and status quo? In your recommendation, give reasons for your suggested course of action and possible outcomes. Here is a suggested format for the memo (you can use any format you wish):

February 15, 1968

TO: President Lyndon B. Johnson
FROM: (YOUR NAME), Foreign Policy Advisor
SUBJECT: US Policy in Southeast Asia

(INSERT YOUR TEXT HERE)

Your essay will be graded using the following rubric:

Writing Style (50 Points)

Sentence and Paragraph Structure - 10 points
Grammar and Spelling- 10 points
Clarity - 10 points
Organization - 10 points
Introduction and Conclusion - 10 points

United States Policies to Win the War (50 Points)

Policies that have worked - 25 points
Policies that have not worked - 25 points

Course of Action President Should Take (50 Points)

Clarity of Description of Action- 25 points
Reasonableness of Actions - 25 points

Reasons for Recommendations (25 Points)

Possible Outcomes (25 points)

One of the shortcomings of many point systems is that there are not specific criteria delineating what gets full credit for a certain number of points. This can be remedied if teachers clearly articulate the criteria for a certain number of points. **FIGURE 4-4** provides an example of a point system for short-answer questions developed to evaluate content knowledge for a Teaching American History Grant in Roanoke Rapids, North Carolina. The rubric builds substantially on work done by the Center for Research on Evaluation, Standards, and Student Testing (CRESST).

FIGURE 4-4:
Short Answer Rubric—Conceptual Identifications

0 - NO KNOWLEDGE OR UNDERSTANDING

Student provides no answer or a completely wrong answer that indicates no understanding of the concept. This can also be a simple reordering of the same words.

1 - MINIMAL KNOWLEDGE OR UNDERSTANDING

The response shows that the student has some inkling of knowledge or understanding of the concept(s). This can be demonstrated in two ways:

A. A correct association is provided but conveys no understanding of the significance of the concept(s).

Assimilation: becoming similar

Capitalism: money

An association should NOT be counted if it is simply a reordering of the words of the term with no other words added. Both of the answers below would be scored as 0.

Ex. westward movement—moving west

Industrialization—industry

B. An incorrect definition that suggests some correct association

Ex: Urbanization—people moving out of cities (urbanization does involve people and cities but they are not moving out of cities)

2 - PARTIAL KNOWLEDGE OR UNDERSTANDING

The student indicates some partial understanding or some knowledge of the concept; however there is no attempt to connect the information to a period in history or to provide a factual example.

Ex: State's rights—states can decide on things

3 -MOSTLY COMPLETE UNDERSTANDING OR KNOWLEDGE OF CONCEPT

The student:

A. Has complete understanding of the concept and knows how to connect it to the time period but provides no factual example

Or

B. Provides an accurate factual example of the concept but presents a limited understanding of the concept.

Ex: Industrialization—U.S. had dramatic shift from agriculture to making things by machines. Contributes to growth of cities/ source of difference between north and south.

Imperialism—U.S. going overseas, such as involvement in the Philippines and Cuba.

4—COMPLETE KNOWLEDGE OF UNDERSTANDING OR CONCEPT

Student has complete understanding of the concept, connects it accurately to the time period, and provides at least one factual example of the concept. Ex. Imperialism: By end of 1800s, U.S. had completed expansion of territories in United States. Pressure began to mount to become a world power by reaching overseas. Business community wanted more markets and religious leaders wanted to spread Christianity abroad. At the turn of the century, U.S. was sending forces to Philippines, East Asia, and Cuba.

Adapted from Baker, Aschbacher, Niemi, and Sato, 1992

When teachers begin laying out specific criteria for the achievement of specific points, they are beginning to create a version of a rubric, the most informative way of presenting assessment criteria.

3. Rubrics

Rubrics are rating scales that are used with performance assessments. Defined as “scoring guides, consisting of specific pre-established performance criteria, used in evaluating student work on performance assessments” (Mertler, 2001), rubrics describe different levels of student performance. These kinds of rating scales are like grades (A-F) but with the gradations in quality described in detail for students relative to the expectations for the task. Such criteria are especially important when teachers are trying to engage students in reflecting on or assessing the quality of their own work.

There are two types of rubrics: holistic and analytic. Holistic rubrics require judgment of the overall performance as a whole while analytic rubrics involve examination of specific aspects of the performance (Mertler, 2001).

To Do

Try out this short answer assessment point scale with some of your student’s work. Does it work? Does it help you assign points more accurately?

Holistic Rubrics

Rather than assigning separate scores for each important aspect of task performance, holistic ratings consider all the criteria simultaneously and result in a single summary or grade. This approach may be most appropriate when the purpose is to provide students with an overall index of their performance on a task or product. For example, teachers often choose to provide overall scores for essays to indicate the extent to which a student was able to write an effective argument. Suppose a sixth-grade social studies teacher asked her students to describe the relationship of the United States' government to the governments of Greece and Rome (South Carolina Curriculum Standards). For this teacher, a "quality" essay might include accurate content, logical organization with main ideas and supporting details, and clear writing. She may use a holistic rubric something like the one in **FIGURE 4-5**.

FIGURE 4-5:
Holistic Rubric for Rating Student Writing Samples in Social Studies (Grade 6)

4	<ul style="list-style-type: none">● Consistently uses accurate data.● Demonstrates a logical plan of organization and coherence in the development of ideas.● Develops ideas fully using such things as examples, reasons, details, explanations, and generalizations that are relevant and appropriate.● Consistently expresses ideas clearly.
3	<ul style="list-style-type: none">● Generally uses accurate data.● Develops the assigned topics using a general plan of organization.● Demonstrates satisfactory development of ideas through the use of adequate support materials.● Generally expresses ideas clearly.
2	<ul style="list-style-type: none">● Uses some accurate data.● Attempts to develop the assigned topic, but demonstrates weakness in organization and may include digressions.● Demonstrates weakness in the development of ideas with little use of support materials.● Has difficulty expressing ideas clearly.
1	<ul style="list-style-type: none">● Uses little accurate data.● Minimally addresses the assigned topic but lacks a plan of organization.● Does not use support materials in the development of ideas or uses irrelevant material.● Does not express ideas clearly.

(Chicago Public Schools, n.d.)

FIGURE 4-6 provides an example of a holistic rubric for assessing historical knowledge at the elementary level.

**FIGURE 4-6:
Holistic Scale for Assessing Historical Knowledge
(Elementary)**

5	<ul style="list-style-type: none"> ● Offers accurate, comprehensive and complete analysis of the information and issues. ● Provides a variety of facts to explore major and minor issues. ● Extensively uses previous historical knowledge to provide an in-depth understanding of the problem and to relate it to past and future situations.
4	<ul style="list-style-type: none"> ● Concise, accurate analysis of the documents. ● Facts provided relate the major and minor issues involved. ● Recalls previous general historical knowledge to examine historical issues.
3	<ul style="list-style-type: none"> ● Relates only major facts to the basic issues with a fair degree of accuracy. ● Analyzes information to explain at least one issue with essential support. ● Uses general ideas from previous historical knowledge with a fair degree of accuracy.
2	<ul style="list-style-type: none"> ● Provides only basic facts with only some degree of accuracy. ● Refers to information to explain at least one issue or concept in general terms. ● Limited use of previous historical knowledge with only limited accuracy. ● Major reliance on the information provided from text.
1	<ul style="list-style-type: none"> ● Repeats one or two facts without complete accuracy. ● Deals only briefly and vaguely with concepts or issues. ● Indicates little or no previous historical knowledge. ● Relies heavily on the information provided.

(Chicago Public Schools n.d.)

Analytic Rubrics

Analytic rubrics are rating scales that describe performance along a continuum. They are used to describe a product or a performance along multiple dimensions. They are most useful when teachers want to focus on specific components of a larger product. For example, to assess an essay analyzing primary sources, the dimensions or criteria that might be rated include: quality of content knowledge (accuracy, importance), quality of historical thinking (defensible interpreta-

tion of texts, connection of texts to time period), and quality of writing (focus and clarity). Each dimension would have a description of what it meant at different levels of implementation.

Let's take a look at one of the essay assignments from **CHAPTER 3** (p. 22). This assignment and the accompanying rubric were developed for high school students as part of a Teaching American History Grant in Horry County, South Carolina.

Assignment: The Role of Religion in the New Republic

You are member of a state legislature in 1800. Someone has proposed a law that would allow the state government to purchase lands for churches. You know that the issue of the appropriate role of religion has been an important struggle in the early years of the United States. You also know that some states at this time period have provided financial support to specific religions. To inform your decision, you have read, among other pieces, the works of two of the founding fathers, George Washington and James Madison.

- Summarize Washington's and Madison's perspectives on religion and government. In your summary, bring in additional information from the time period to describe why the writer might have this perspective.
- How will you vote? Defend your decision, using the texts and other information you know from the time period. Remember, you are legislator in 1800, not today.
- Now, pretend you are a legislator today voting on legislation authorizing the display of the Ten Commandments in government building. What would your decision be? What justification would you give? You must at least partly defend your decision by using historical information from the early Republic. You may also bring in more recent information as appropriate.

The goals of this assessment are to measure students' knowledge of the time period, their ability to interpret primary source documents and place those documents in a context, and their ability to make a connection between history and events of today. Thus, the rubric has the criteria as laid out in **FIGURE 4-7**. As you look at the rubric, note that the criteria in this rubric are limited in number and fairly general. Good quality rubrics don't try to assess everything; instead, they focus on a small but critical set of skills or knowledge. It is often desirable to assess skills that are transferable to other topics (e.g., skill in summarizing main points) or can be applied beyond the specific item being assessed (e.g., skill in defending a position using evidence). In some cases, teachers may want to include task-specific criteria (such as knowledge of a specific historical event) in the rubric (Moskal, 2000).

FIGURE 4-7: Essay Rubric: US History—High School

The essential question is whether students are able to apply the thinking and tools of historians to critical historical issues.

CRITERIA	4	3	2	1
Historical Knowledge: Knowledge outside of text	Essay includes substantial information (more than 4 facts or concepts) not found in the text.	Essay includes three or four facts or concepts not found in the text.	Essay includes one or two facts or concepts not found in the text.	Facts and concepts in essay come only from text.
Historical Knowledge: Accuracy and Importance	The most important historical concepts or information necessary to answer the question are included. Concepts and facts included are accurate and used appropriately in the essay.	The writer includes some of the most important historical concepts or information, although one or two key ideas are missing. Any misconceptions/factual errors are minor and do not affect the overall accuracy of the argument.	The writer is missing most of the important historical concepts or information necessary to answer the questions. Specific details included may be accurate but could be seen as insignificant. Alternately, there are no misconceptions but factual errors are prevalent enough to negatively impact accuracy.	Information included is either irrelevant or insignificant to the question. The writer does not include any of the most significant historical concepts or terms. Key statements reveal misconceptions; details and examples contain many factual errors.
Historical Thinking: Interpreting texts	Essay presents a valid and defensible interpretation of the text. In the interpretation, the writer clearly describes how the text and the author are products of the specific time period.	Essay presents a valid and defensible interpretation of the text. There is a limited attempt to place the text or author in the context of the time period.	Essay presents a somewhat questionable interpretation of the text. There is little attempt to connect the text or the author to the time period.	Writer demonstrates little understanding of the text. There is no attempt to describe how the document or author is related to the time period.
Historical Thinking: Connection to today	The essay clearly describes how the events of today have their roots in historical experiences.	The essay articulates a connection between events today and historical events.	The essay discusses events of today, making a superficial or loose connection to historical events.	Writer discusses only events of today without making any link to historical events.

CRITERIA	4	3	2	1
Historical Communication	The essay's organization is logical and appropriate to the topic; the organizational scheme may be chronological or thematic. The essay has a core argument that responds to the question. The writer includes key statements that contribute to the argument. Each key statement is supported by relevant specifics. The organization and the focus of the essay make interpretation of the purpose/ thesis/main idea easy for the reader.	The essay's organization is generally logical and appropriate to the topic. The essay has a core argument that responds to the question. Key statements, and supporting details, are generally in support of the topic, although the connection may not be obvious at times. The organizational structure is strong enough to move the reader through the text and help the reader interpret the purpose/ thesis/main idea without too much confusion.	The writing addresses the topic, but contains many extraneous or loosely related ideas. The writing responds to the question, but rather thinly or unevenly. There may be a central argument but the key statements and supporting details do not always further that argument.	The essay responds to the topic but does so in either a very general or a very haphazard way. Ideas, details or events appear strung together in a loose or random fashion. There is no distinction between key statements and supporting details or the details do not support the statements.

The significant advantage of analytic rubrics is that they offer diagnostic information to students about the strengths and weaknesses of their performances on a variety of dimensions so they can better target performance areas that need to be improved. The dimensions chosen and the descriptive categories used for the rating scales need to be chosen so that they communicate to students what it is important to do well.

Creating and Using Quality Rubrics

Creating a quality rubric does take time and effort. However, taking this time up front results in substantial savings during the actual assessment of the work. In addition, the time spent up front in developing a quality rubric that is used to guide instruction also results in time saved as the instructional process becomes more clearly focused.

Herman, Aschbacher, and Winters (1992) describe an iterative process for developing the criteria inherent in rubrics. We have adapted the components of this process as follows:

1. Think about the specific skills and knowledge you want to assess. Make sure these are important skills and knowledge on which you are willing to spend some time. You would not want to spend a large amount of your valuable instructional and assessment time on unimportant or insignificant skills or knowledge.

2. Investigate and think about what the field of history defines as good work by historians in these skills and knowledge. We have included some resources at the end of this book that will be helpful in investigating how educators translate historians' skills into the classroom.
3. Gather model rubrics (such as some of the rubrics in this book) in different areas that you might want to assess. You can use these as starting points, but you should feel free to adapt them to meet your needs.
4. Describe the high end performance for each of the important dimensions you will assess. What would be high implementation in each set of skills or knowledge? For everything except a checklist, think backwards from that high performance point to identify two to three levels of weaker performance.
5. Consider the rubric in the light of actual student work samples. Try it and revise as necessary.

Online Resources for Rubrics

<http://school.discovery.com/schrockguide/assess.html>

—Kathy Schrock's Guide for Educators includes lots of links to sites on rubrics. There is a technology slant, but much of the information is applicable for all subject areas.

<http://www.nwrel.org/assessment/toolkit98.php> —The

Northwest Regional Educational Laboratory has developed a toolkit full of information about assessment, including rubrics. Virtually the entire kit can be downloaded for free.

<http://rubistar.4teachers.org> —Rubistar is an

online rubric generation site. You can select from existing categories or modify them to make your own.

<http://www.uwstout.edu/soe/profdev/rubrics.shtml> —

the University of Wisconsin-Stout has a good site with samples of general rubrics in different skill areas.

http://intranet.cps.k12.il.us/Assessments/Ideas_and_Rubrics/ideas_and_rubrics.html —Chicago

Public Schools has good information about creating rubrics, as well as a bank of sample rubrics.

<http://www.rubrics4teachers.com/> —Teacher

Planet, which has free membership, offers many resources, including a large collection of rubrics.

<http://pblichecklist.4teachers.org/> —This site will

help teachers develop checklists for project-based learning.

There are some specific things to think about in creating rubrics. First, when used to help students learn, rubrics need to be written in language students can understand and need to avoid vague terms, such as "creative" or "challenging." Andrade (1999) discusses how a pair of middle school teachers, working on an oral presentation rubric, avoided using the term "starts with a creative beginning" by listing possible ways that students might demonstrate a creative beginning. Good quality rubrics also avoid unnecessarily negative language, such as "boring."

If you are developing an analytic rubric, coming up with different points on the scale can be challenging. Andrade (1999) recommends using the following strategy developed by a fifth grade

teacher: thinking about the levels as a “Yes,” “Yes but,” “No but,” and “No.” **FIGURE 4-8** provides an example of what this would look like.

FIGURE 4-8:
One Criterion of an Analytic Rubric for a Research Report—Middle School

CRITERION	QUALITY			
Use of multiple primary and secondary sources	Yes, I used more than one primary source and more than one secondary source in my report.	Yes, I used both primary sources and secondary sources in my report, but I had only one primary source and/or one secondary source.	No, I didn't use both primary and secondary sources in my report, but I did use more than one source.	No, I didn't use more than one source in my report.

As you develop your own rubrics, you may want to evaluate them using —what else—an analytic rating scale for rubrics. **FIGURE 4-9** lays out various criteria that are necessary for rubrics: how useful the rubric is, whether it provides students an opportunity to self-assess their work, how valid and reliability it is, the clarity of the descriptors used, and the importance of the key elements in the rubric.

One of the most important things to remember about rubrics is that they are meant to be used. Students should work with the rubrics to help them understand the expectations, to help them create a quality product, and to develop the ability to assess themselves and their work. In fact, having students use rubrics to self-assess their own work may be one of the most powerful uses of rubrics (Black and William, 1998). Descriptive studies have shown that students used rubrics consistently and reliably and that their use did not depend on their academic ability (Hafner and Hafner, 2003). Teachers should work with rubrics to guide their own instruction and ensure they are focusing on key areas, to track student work, and, of course, to assess student performance.

The next chapter discusses how to use assessment criteria to give effective feedback to students and how to translate them into grades.

FIGURE 4-9:
The New and Improved Aunt Olive’s “Show-Me” Rubric on Rubrics by Ned Miller

CRITERIA	4	3	2	1
Useful	Assessment is useful and convenient to the learner from the beginning of the discussion about the task and it concentrates all student energy on what is important to accomplish.	Learners can understand the assessment and begin to use it before the task has begun. It focuses their energy on what it is important to accomplish.	Assessment is available to learners at some point before the task is completed. Students can use it to get an idea of what is important to accomplish.	Assessment not available to learner or learner cannot interpret or understand it.
Self-Assess	Learners required to self-assess and are required to reflect and concentrate on quality of work.	Learners encouraged to self assess and reflect on quality of work.	Learners are invited to self-assess, but may not reflect on quality.	There is no promotion to self assess or focus on quality.
Valid	An explicit set of objectives is measured. Understanding is a prerequisite to scoring well. Students cannot score well or poorly due to factors unrelated to objectives.	A distinct set of objectives is measured. Students must understand the concepts to score well and will score well if they understand.	Vague objectives may be present. Students who understand objectives tend to score higher.	Students are uncertain as to what is expected. Assessment does not measure what it says it measures.
Reliable	Different students & teachers assessing the same task will get the same results.	Different students & teachers assessing the same task will get highly similar results.	Students & teachers are guided to similar results.	There is no consistency of results.
Descriptors	Has explicit descriptors that allow the students & teacher to distinctly discriminate within a range of quality. (4 or 6 are ideal so no “middle” exists)	Has specific descriptors, which allow the students & teacher to discriminate a range of quality that is limited to few enough so that discrimination is distinct. (usually 3-7 levels)	Has descriptors that help students & teacher measure specific items but there may be too many items.	No understandable descriptors.
Key Elements	Measures only vital elements that are critical to the task. No items of low consequence.	Measures key elements that are essential to the task. Few items of low consequence.	Meaningful items assessed, but may be of limited significance.	Inappropriate items are assessed (due to ease of grading?)

Adapted by Susan Johnson & Debra Smith

Source: Electronic Learning Marketplace

To Do

Pick an assignment you would like to assess with a rubric. Follow the steps to create a rubric.

- Describe the skills and knowledge you want to assess. Choose no more than four or five. Decide whether you want to assess all of these at the same time (a holistic rubric) or whether you want to measure achievement on each of these independently (an analytic rubric).
- Think about what the field of history has to say about those skills and knowledge.
- Find some model rubrics from this book or any of the resources listed.
- Identify the high end description of performance on the task for each skill or dimension to be assessed. What does “success” look like? Work backwards to create gradations of quality.
- Assess your rubric using Aunt Olive’s.
- Try it out with some students’ work. How did it work? Was it useful to you? Was it useful to your students?

USING ASSESSMENT INFORMATION: FEEDBACK AND GRADING

The previous chapters have discussed identifying appropriate learning goals in history, using assessments aligned to learning goals, and developing criteria that measure the learning goals. This chapter focuses on what you actually do with assessment information. Why are we going to all this trouble?

In **CHAPTER 1**, we talked about two purposes for assessment: formative assessment and summative assessment. In this chapter, we discuss the use of assessment information for formative and summative purposes. Specifically, we look at how you use assessment results to provide feedback for improvement and to provide summative information such as grades or scores.

Formative Assessment: Providing Feedback

“The most powerful single innovation that enhances achievement is feedback. The simplest prescription for improving education must be ‘dollops of feedback.’”
(Hattie, 1992)

Some researchers have argued that formative assessment is one of the best ways to make dramatic gains in student learning, particularly for low achieving students (Black and William, 1998). While formative assessment can help teachers identify ways to modify their instruction, one of the best ways it can help improve student achievement is by providing feedback to students centered on their work and how they can improve it.

The research support for feedback is among the strongest in the educational field. John Hattie’s (1992) review of 7,827 studies found that studies on providing students with specific information or feedback regarding their work in relation to particular objectives reported increased student achievement. In their examination of over 40 experimental and quasi-experimental studies, Black and William (1998) reported that the studies on the effects of formative feedback reported results ranging from effect sizes of .4 to .7. This translates to gains of approximately 25 percentile points.

What is *feedback*? Feedback is information that is given to students concerning their performance. According to Black et al. (2003), appropriate feedback includes the following key components:

- Data on the student’s performance on a criterion
- Data on the desirable performance of that criterion
- Information comparing the two performances and assessing the gap between them
- Information about how the gap can be changed

W
I
L
L
I
A
M
S
H
I
R
T
O
N

Note how the components of feedback are centered on the work the student is accomplishing relative to a standard. The feedback is not making judgments about the student himself or about how the student is performing in relation to other students. As Black and William (1998) recommend, "feedback to any pupil should be about the particular qualities of his or her work, with advice on what he or she can do to improve, and should avoid comparisons with other pupils" (p. 6).

In a meta-analysis of many research studies, Marzano et al (2001) found that the most effective feedback shared the following characteristics:

1. Corrective. It provides a correct answer or an explanation of what is accurate and what is inaccurate. Quality feedback involves explanation, not just an indication of whether the answer was correct or not (Bangert-Drowns et al, 1991).
2. Timely. Feedback should be given soon after the work is completed.
3. Criterion-referenced not norm-referenced. This means the feedback is connected to an established standard of learning, not comparing one student against another student.

In addition, feedback should be both teacher and student-based. Students should get the opportunity to provide some of their own feedback and provide feedback to each other. This improves students' ability to examine their own work critically and to determine themselves what they need to do to get better (Shepherd, 2000).

Examining Your Use of Feedback in the Classroom

Ask yourself the following questions:

1. What kind of feedback do I provide to my students?
2. When do I give them feedback?
3. How well does my feedback connect to the learning standard I have for the students?

Ongoing feedback can be time-consuming, which is why it should be integrated as naturally as possible into the instructional process. Black and William (1998) built on their extensive research to suggest some ways that this can happen:

- Build in opportunities, through discussion and dialogue, to make sure that students have a chance to share their evolving understanding. Make sure those discussions are genuine by asking questions to get at students' understanding not at some predetermined answer. Also, wait long enough for students to think about their answers and probe them to justify their responses.

- Give students tests and other exercises as opportunities to provide feedback. It is better to have frequent short tests instead of infrequent long tests. Any new learning should be tested within a week. Examine the quality of the questions carefully.
- Create a culture that values deep thinking and questioning. In this case, ongoing reflection and assessment become part of the way the class does business. Have students use their responses to questions to share their thinking with others. Many students do work simply to turn it in and get a check. If the work they do is used in subsequent discussions, they may feel more accountable for doing it well.

Summative Assessment: Grading

We know that formative assessment is critical for helping students improve their learning. At some point, however, for accountability purposes, to communicate progress to parents and community members, or, at the very least, to fulfill school and district expectations of report cards, there must be a summative assessment: a picture of what students have learned.

We have discussed how checklists and rubrics can create a better picture of what students know and are able to do. But what is the relationship between rubrics and the grades that appear on report cards? Grading students' achievement status involves the teacher in a series of decisions. In *A Practical Guide for Developing Sound Grading Practices*, Stiggins (1991) describes some of the decisions that must be made if a single grade is to be given over an assessment period.

First, what criteria should be considered in determining a report card grade? A grade over a grading period is usually considered as a composite measure of student achievement on the objectives of the instruction, rather than a measure of student interest, attitude, or effort.

Second, as a measure of achievement, what and how much grading data should be gathered? For the grade to be a valid measure of achievement in a subject or course, there must be a sufficient sampling of student performance on the critical subject or course objectives to provide a fair and reliable assessment of the student's achievement status.

The teacher must be the judge, but at one extreme, using only one measure (e.g., a one-hour, paper-and-pencil exam) to determine a report card grade for a grading period is clearly not a sufficient sample of student performance. At the other extreme, assessing student performance daily would not provide students with the time needed to develop competencies and skills in advance of being assessed.

Third, how are individual grades or scores combined into a single grade at the end of a grading period? Performance-based assessments can be readily incorporated into any grading system.

Suppose you had the following five goals in mind for students. Students will be able to:

1. Demonstrate knowledge of the Great Depression.
2. Analyze primary source documents.
3. Write a report providing a historical interpretation that incorporates multiple sources. Defend that interpretation with evidence.
4. Work collaboratively to identify the impact of the Great Depression on life today.
5. Discuss historical content with accuracy. Provide justification for conclusions.

Given these five instructional objectives, student assessment data as shown in **FIGURE 5-1** might be used to arrive at a report card grade. Each assessment would have a scoring key or rubric that clearly describes what the students need to do to get specific points. The column entitled “weight” refers to the importance of a particular assignment in the overall grading period.

FIGURE 5-1:
Sample Grading Period Weighting System

ASSESSMENTS	INSTRUCTIONAL OBJECTIVE	MAXIMUM POINTS	POINTS EARNED	WEIGHT	STUDENT SCORE/ MAXIMUM SCORE
Paper and pencil test on Great Depression	Knowledge of Great Depression	50	42	2	84/100
Primary Source Analysis sheets (5 x 10 points each)	Analyzing primary source documents	50	35	1	35/50
Historical Interpretation Report	Analyzing primary source documents: Write a report providing an historical interpretation	25	20	4	80/100
Checklist for weekly discussions (5 x 10 points each)	Knowledge of Great Depression: Discuss historical content, justify conclusions	50	45	1	45/50
Group project on connection between Depression and life today	Analyzing primary source documents: Work collaboratively to connect Great Depression to life today.	50	40	1	40/50
Total					284/350

In **FIGURE 5-1**, demonstrating knowledge accounts for 150 points out of 350 or 43% of the overall grade; analyzing primary source documents is addressed in assessments that represent a total of 200 points or 57% of the overall grade; writing a report interpreting historical information is 100 points or 28% of the overall grade; oral discussions account for 14% of the grade; and working collaboratively to connect the Great Depression to life today also counts 14% of the overall grade. When you add up these percentages, they add up to more than 100. This is because some instructional objectives are being measured in multiple ways. Nevertheless, the proportion of the total grade that is addressing specific instructional objectives should be proportional to the emphasis you want to place on each of those objectives. In this case, the assessment distribution indicates that the ability to analyze primary source documents is the most important part of a student's grade. If you believe that another aspect of historical thinking should be more important, then that aspect should receive the majority of points.

The weighting system used in deriving report card grades should be related to course objectives and explained to students so that they know their goals. In the example above, students might be informed at the beginning of the grading period about the five instructional objectives and the assessments to be used. The number of points needed for different grades (A, B, C, etc.) should also be communicated.

In such a point system, it is also important to stay flexible so as not to penalize students when an assessment is poorly designed. For example, if students were told that 300 points constituted an A, but students lost points because of poorly designed items on the test, some adjustment to the point system would have to be made.

Student achievement status on important instructional objectives can be communicated in ways other than a single report card grade in Social Studies or History. Some teachers find that grades, although required by policy, are not particularly helpful in conferencing with students and parents about students' performance on specific goals. Checklists, rubrics, and narratives can be used in addition to grades or as an alternative means of reporting.

The next chapter concludes our discussion of assessments by examining the relationship between classroom assessments and assessments done for accountability purposes and by providing tips for getting started on your work with assessments.

To Do

Using Figure 5-1 as a model, develop a table that shows your own Grading Period Weighting System. Then answer the following questions:

1. What activities count towards the final grade?
2. What skills do those activities measure?
3. What percentage of the grade measures different skills? How closely does the percentage assigned to each skill match your priorities?
4. Are there valued skills or knowledge that are not represented on your table?

SOME CONCLUDING THOUGHTS

Appropriate Assessments in an Era of Accountability

All of the information in this book is all very good and well but you may think it is vastly removed from the reality of teaching today with its focus on high-stakes state-mandated tests. In some cases, your state may not test social studies at all, in which case the subject of history may be neglected at the elementary level in the singular focus on reading and math (Jones et al. 1999). At the high school level, there may be end-of-course tests in history that focus primarily on recall of specific historical events. In this situation, teachers sometimes express feeling pressure to “cover” the entire history curriculum so students have at least heard the terms once before they take the test.

The types of assessments included in this publication are not contradictory to the state-mandated assessments. Instead, they supplement and can actually help your students do better on state tests. The analysis of historical documents and information as discussed here builds students critical literacy skills, which should help them on a variety of standardized tests (Wineburg, 2005).

What is the relationship between the types of assessments discussed in this book and the standardized tests that happen at the end of the year? To answer this question, let’s begin by looking at what teachers need to do to get students ready for standardized tests. First, they need to help students learn the content that would be included on the test. Second, they need to help students learn how to work with the content in the ways that are assessed on the test. Finally, they need to help their students learn test-taking strategies. The last item is probably the least important and needs the least amount of time—although it can be the thing on which teachers focus the most. What teachers need to spend most of their time on is helping students learn the content and skills.

“Can effective teaching take place in the absence of learning? Certainly not” (Guskey, 2003, p. 7). This quote reminds us of one simple rule for teachers: Just because a teacher teaches does not mean that a student learns. In order for students to learn, they need to work with the information. In a broad examination of effective instructional strategies, Marzano et al. (2001) reported on studies that found that students improved their achievement by approximately a quarter to a third when they engaged in strategies such as summarizing information, utilizing graphic organizers, and generating and testing hypotheses. In short, students’ achievement increases substantially when they are cognitively engaged with the content. Thus, although lecture or reading may be the most efficient way to *teach* the content, it will not always be the best way to get students to *learn* the content.

X
S
S
T
E
S
T
I
N
G
C
H
A
P
T
E
R

If students need to learn the content in order to do well on the exams, how do you know whether they have learned the content? You need to assess, of course. Some of the assessments you choose to administer may mirror those that occur at the end of the year or end of the course. Those assessments will provide you with limited information about your students' understanding. You may know how many answers a student got correct, but you may not know how thorough or accurate their understanding is of the event or topic. The kinds of assessments described in this manual give teachers richer information about what students can and can't do. They will help you draw conclusions about students' knowledge and skills and where they need to improve.

In addition, some researchers have hypothesized that students are more motivated and are more likely to engage in learning for learning's sake when they are doing more challenging performance assessments instead of paper and pencil tests (Brookhart and Durkin, 2003). Motivating students to engage in learning is one of the most significant challenges of any teacher; if high quality performance assessments can help in that process, they will also likely contribute to improved achievement.

Getting Started

Teachers who want to begin incorporating more and varied assessments in their history classrooms need to be aware that change is generally a gradual process that occurs through experience, dialogue and reflection.

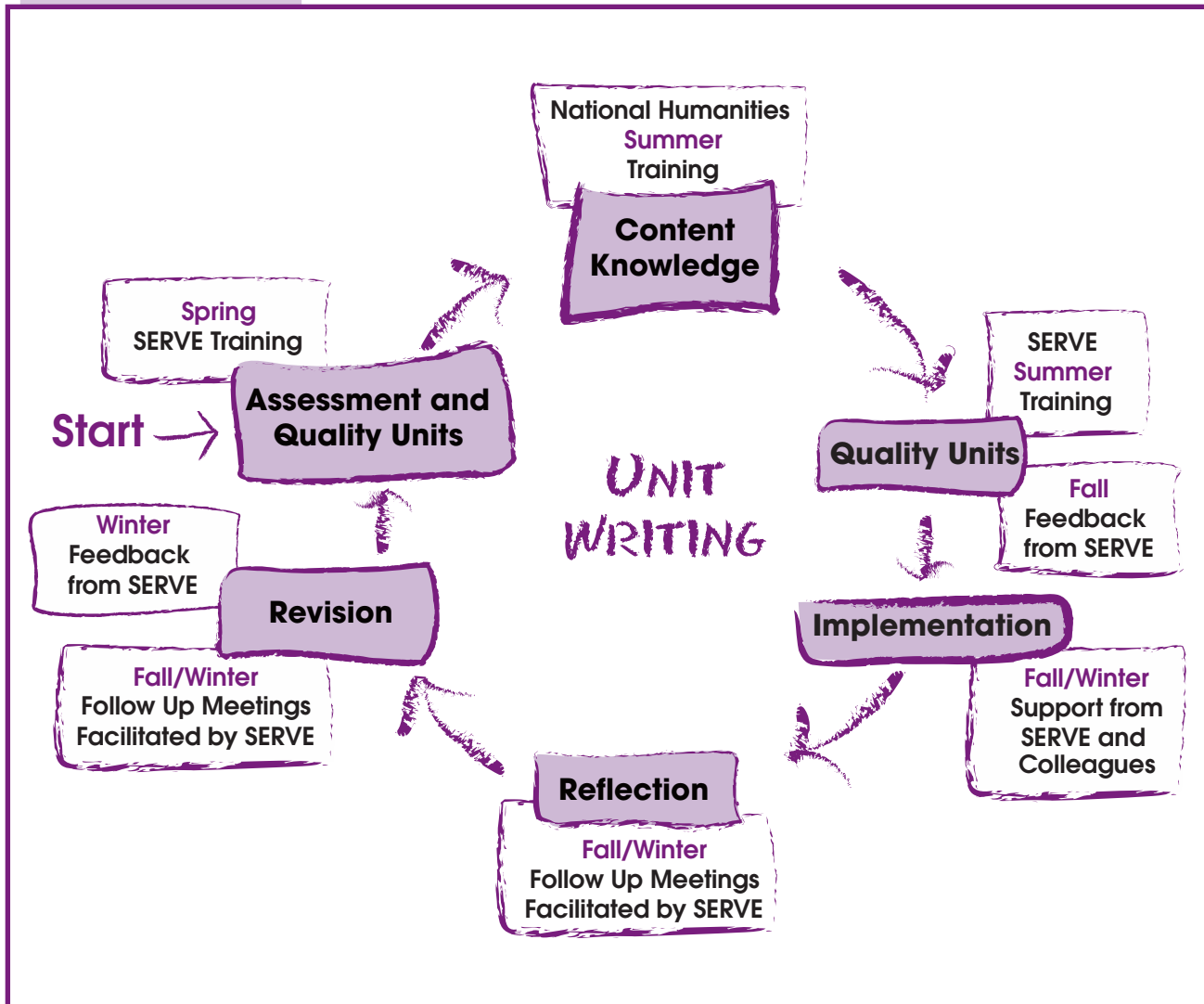
Teachers need time to try out new assessments, reflect on their success or failure, and make revisions. Just as student learning is an individual process that is self-directed and personally constructed, so is teacher learning about assessment practices. Changing assessment practices is not a simple, linear, lock-step process that all teachers follow in a prescribed manner. Rather, it is a process of becoming more purposeful about:

- Desired student outcomes in history
- The use of assessment methods that match well with desired outcomes
- The quality of the assessment including the criteria
- The use of grading systems that reflect student achievement on these outcomes.

How might teachers engage in this more purposeful thinking about student assessments?

1. Some districts have initiated district-wide staff development efforts to strengthen assessment within the context of sound instructional planning (e.g., units, lessons) in history. **FIGURE 6-1** shows one model of staff development in which assessment is being considered within the broader context of quality unit development. This model is being implemented by Roanoke Rapids, North Carolina and surrounding districts as part of a federal Teaching American History grant.

**FIGURE 6-1:
Model of Staff Development**



(Source: Butler and McColskey, 2004 at <http://www.serve.org/Assessment/assessment-publicationh2.php>).

2. If no systematic district or school professional development support is provided, teachers, either individually or in informal groups, could begin to reflect on their assessment practices. Incorporating more and varied assessments into the classroom may be easier if experiences, concerns, and frustrations are shared with colleagues. Sharing successful tasks and methods with other teachers also increases the number of assessments available. The idea of thinking about a teacher's responsibility as that of providing quality work for students to do was advanced by Dr. Phillip Schlechty (2002), who is the founder of the Center for Leadership in School Reform in Louisville, Kentucky (www.schlechtycenter.org). He defined ten dimensions of quality work (e.g., clear and compelling product standards or criteria, authenticity) so that teachers could have specific ideas about how to continually improve upon their assignments and performance assessments. Others have also developed rubrics that teachers can use to examine

the quality of their assignments/assessments (i.e., work given students to do):

http://www.cresst.org/resources/justforteachers_set.htm

<http://www.nwrel.org/assessment/toolkit98/A36H1.PDF>

<http://www.consortium-chicago.org/publications/p0a02.html>

In addition, there is support for teachers in the form of meeting protocols (structured agendas) for groups of teachers to use in conducting meetings for the purpose of “tuning” or improving the quality of their assessments or assignments (<http://www.serve.org/Assessment/IQ/strprofdev.php>). For more information on forming Critical Friends Groups, see:

www.lasw.org/index.html

www.cesnorthwest.org/critical-friends-groups.htm

www.nsrharmony.org

There is no right place to start with assessment. There are many activities, depending on the prior experience, time constraints, interest, and resources of the teacher(s) involved, that represent jumping-off points for changing or at least reflecting on assessment practices.

Listed below are some examples of activities that might get conversations started about assessment practices.

- Articulate one very important desired student outcome (refer to **Chapter 2**). For example, a teacher might be interested in how well a student can apply historical knowledge to an issue today. Review the assessment methods in **Chapter 3** and choose an approach to assessing students’ competence in this area that you have not tried before. Try the assessment approach and see what you learn about student performance and about the assessment method.
- Experiment with a format for a grading period that outlines for students the major goals you have for their performance, how their performance on these goals will be assessed, and how report card grades will be derived.
- Start a list of the advantages and disadvantages of each of the assessment methods described in **Chapter 3**. What do you think you need to know from someone who has tried each method before you go any further?
- Develop a chart (see **FIGURE 5-1**) showing how you can combine assessment data in obtaining student report card grades. What kind of weighting system are you using? Does the weighting match up with your priorities for student learning?
- Analyze the tests you have used in the past. Try to sharpen the focus and eliminate more trivial items based on the information in this manual.
- Start a folder of assessment samples from released state tests or item banks, other teachers, district tests, or published articles and critique them for your own purposes.
- Revise the projects or essays you already do with your students. Identify those that are most essential and experiment with rubrics that could be used to assess student

performance on those tasks. Reflect on the effect rubrics have on the instructional and assessment process.

The process of incorporating and using a broader array of assessment methods can sharpen teachers' thinking about the learning goals they have for their students and how to measure achievement of those goals. It can also result in improvements in the quality of instructional tasks that teachers design for students. Finally, if teachers are explicit and purposeful about their goals, students are more likely to evaluate the quality of their own work. The benefits of experimenting with a variety of assessment methods lies as much in the conversations they engender between teachers and students and among teachers as in the information they provide about student competence. Conversations about the characteristics of quality work can empower both students and teachers, resulting in powerful and engaging learning in history.

- Alabama Content Standards, Social Studies. Retrieved 3/6/2006 from <http://alex.state.al.us/standardAll.php?subject=T1&summary=1>
- Andrade, H. (1999) *Understanding rubrics*. <http://learnweb.harvard.edu/alps/thinking/docs/rubricar.htm>
- Baker, E. L., Aschbacher, P.R., Niemi, D. & Sato, E. (1992) *CRESST performance assessment models: Assessing content area explanations*. Los Angeles: National Center for Research on Evaluation, Standards, and Testing, UCLA
- Baker, E. L. and Linn, R. L. (2002). *Validity issues for accountability systems: CSE Technical Report 585*. Los Angeles: Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Testing, UCLA. <http://cresst96.cse.ucla.edu/reports/TR585.pdf>
- Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. L., & Morgan, M. (1991). "The instructional effect of feedback in test-like events." *Review of Educational Research*, 61:2, 213-238.
- Black, P. & William, D. (1998b). "Inside the black box: Raising standards through classroom assessment." *Phi Delta Kappan*, 80:2, 139-148 www.pdkintl.org/kappan/kbla9810.htm
- Black, P., Harrison, C., Lee, C., Marshall, B., & William, D. (2003). *Assessment for learning: Putting it into practice*. Berkshire, England: Open University Press.
- Brookhart, S. M. and Durkin, D.T. (2003). "Classroom assessment, student motivation, and achievement in high school social studies classes." *Applied Measurement in Education*, 16:1, 27-54.
- Butler, S. and McColskey, W. (2004). "Increasing academic rigor through high-quality professional development." *The Vision*, 3:2, 30-35.
- Chicago Public Schools (n.d.) The rubric bank. Retrieved 3/6/2006 from http://intranet.cps.k12.il.us/Assessments/Ideas_and_Rubrics/Rubric_Bank/rubric_bank.html
- Criswell, J.R and Criswell, S.J. (2004). "Asking essay questions: answering contemporary needs." *Education*, 124: 3, 510-516.
- Drake, F. D. and McBride, L. W. (1997). "Reinvigorating the teaching of history through alternative assessment." *The History Teacher*, 30. 145-173.
- Electronic Learning Marketplace (2005). <http://www.elm.maine.edu/assessments/>
- National Center for Fair and Open Testing (n.d.). *FairTest fact sheet: Multiple-choice tests*. <http://www.fairtest.org/facts/mctfcat.html>
- Guskey, T. R. (2003). "How classroom assessments improve learning." *Educational Leadership*, 60:5, 6-11.
- Hafner, J. C., Hafner, P. M. (2003) "Quantitative analysis of the rubric as an assessment tool: an empirical study of student peer-group rating." *International Journal of Science Education*, 25: 12, 1509-1528
- Hattie, J. A. (1992). "Measuring the effects of schooling." *Australian Journal of Education*, 36(1), 5-13.
- Haulmann, Bruce E. (2003). Vietnam War: Mid-term exam. Retrieved 3/6/2006 from <http://www.instruction.greenriver.edu/Haulman/History250DE/H250DLW6.htm>
- Herman, J. L., Aschbacher, P.R. & Winters L. (1992). *A practical guide to alternative assessment*. Alexandria, VA: Association for Supervision, Curriculum, and Development.
- Jones, G.M., Jones, B.D., Hardin, B., Chapman, L., Yarbrough, T. and Davis, M. (1999). "The impact of high-stakes testing on teachers and students." *Phi Delta Kappan* 81:3, 199-203.
- Kubiszyn, T. and Borich, G. (1996) *Educational testing and measurement*. New York: HarperCollins.
- Lewin, L. and Shoemaker, B.J. (1998). *Great performances: Creating classroom-based assessment tasks*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Marzano, R. J., Pickering, D. and Pollock, J. (2001). *Classroom instruction that works: Research-based strategies for improving student achievement*. Alexandria, VA: Association for Supervision, Curriculum, and Development.



- McBride, L. W., Drake, F. B. and Lewinski, M. (2005). *Alternative assessment in the social sciences*. <http://www.coe.ilstu.edu/jabraun/socialstudies/assess/soc-sci/frames.html>
- McCoy, M. M. (2003) "Incorporating effective writing strategies." *Social Education* 67:4, 200-203.
- McTighe, J. & Wiggins, G. (2004). *Understanding by design professional development workbook*. Alexandria, VA: ASCD.
- McDaniel, E. (1994): *Understanding educational measurement*, Madison WI: WCB Brown & Benchmark.
- Mertler, Craig A. (2001). "Designing scoring rubrics for your classroom." *Practical Assessment, Research & Evaluation*, 7:25. Retrieved May 18, 2005 from <http://PAREonline.net/getvn.asp?v=7&n=25>
- Moskal, B. M. (2000). "Scoring rubrics: what, when and how?" *Practical Assessment, Research & Evaluation*, 7:3. Retrieved May 18, 2005 from <http://PAREonline.net/getvn.asp?v=7&n=3>
- National Centre for History Education. <http://www.hyperhistory.org/>
- National Center for History in the Schools (1996). *National standards for history: Basic edition*. Los Angeles: Author.
- National Paideia Center (2003). *The Paideia Seminar: Active thinking through dialogue in the secondary grades-A manual*. Author.
- North Carolina Standard Course of Study, Social Studies. Retrieved 3/6/2006 from <http://www.ncpublicschools.org/curriculum/socialstudies/scos/>
- Shepherd, L. A. (2000) *The role of classroom assessment in teaching and learning: CSE Technical Report 517*. Los Angeles: National Center for Research on Evaluation, Standards, and Testing, UCLA
- Schlechty, P.C. (2002). *Working on the work: An action plan for teachers, principals, and superintendents*. Jossey-Bass.
- South Carolina Academic Standards, Social Studies. Retrieved 3/6/2006 from http://www.myscschools.com/offices/cso/social_studies/social.htm
- Stern, S. M. (2000). "Why the battle over history standards?" in ed. Stotsky Sandra, *What's at stake in the K-12 standards wars: A primer for educational policy makers*. New York: Peter Lang.
- Stiggins, R. J. & Knight, T. (1997). *But are they learning?* Portland, OR. Assessment Training Institute, Inc.
- Stiggins, R. J. (1997) *Student-involved classroom assessment*. Upper Saddle River, NJ: Merrill Prentice Hall.
- Strother, D. B. (1989) "Developing thinking skills through questioning." *Phi Delta Kappan*, 324-327.
- VanSledright, B. A. (2004). "What does it mean to think historically...and how do you teach it?" *Social Education*, 68: 3, 230-234.
- Wiggins, G. and Mc Tighe, J. (1998). *Understanding by design*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Wiggins, G. (1998). *Educative assessment: Designing assessments to inform and improve student performance*. San Francisco, CA: Jossey-Bass
- Wiggins, G. (1992). "Creating tests worth taking." *Educational Leadership*, 49:8, 26-34.
- Wiggins, G. (1990). The case for authentic assessment. *Practical Assessment, Research & Evaluation*, 2(2). Retrieved November 29, 2005 from <http://PAREonline.net/getvn.asp?v=2&n=2>
- Wineburg, S. S. (1999) "Historical thinking and other unnatural acts." *Phi Delta Kappan*, 80:7, 488-499.
- Wineburg, S. S. (2005). "What does NCATE have to say to future history teachers? Not much." *Phi Delta Kappan*, 86:9, 658-665.

In addition to the resources already mentioned in the manual, we'd like to suggest the following.

Resources for Teaching History

The *History Matters* website contains an extensive collection of primary source documents on U.S. history, resources for teaching history, and links to other websites. Their primary sources tend to focus on the experiences of everyday Americans at different stages in history, although they will provide links to the classical texts. If you're teaching U.S. History, this is probably one of the most useful sites on the web: www.historymatters.gmu.edu.

World History Matters is similar but not as extensive and is primarily designed for high school and college teachers: <http://chnm.gmu.edu/worldhistorymatters/>.

The American Memory collection at the Library of Congress contains over 7 million primary sources related to U.S. History. Needless to say, it can be a bit difficult to find the perfect resource but their learning page for teachers tries to make things easier: <http://memory.loc.gov/learn/>

The Avalon Project at Yale has an extensive collection of online documents related to law, history, and diplomacy around the world. This is the place to go for treaties, constitutions, military documents, and the like. Many of the classic historical texts can also be found here: <http://www.yale.edu/lawweb/avalon/avalon.htm>

History World includes timelines and basic factual history for countries around the world. This would be good for providing an overview of history for a particular time period or a particular country: www.historyworld.net.

The National Humanities Center is developing toolboxes for teachers on periods in U.S. History. These toolboxes include primary sources, some background information, and guiding questions for discussion: <http://www.nhc.rtp.nc.us/>

The National Center for History in the Schools is the author of the National History Standards. Their website also contains sample units in world and U.S. history: <http://www.sscnet.ucla.edu/nchs/>

Resources on Teaching Writing

Urquhart, V. and McIver, M. (2005). *Teaching Writing in the Content Areas*. Alexandria, VA: Association for Supervision and Curriculum Development. This recently published book comes with a study guide for teachers to use.

Caswell, R. and Mahler, B. (2004). *Strategies for Teaching Writing: An ASCD Action Tool*. Alexandria, VA: Association for Supervision and Curriculum Development. In addition to writing strategies, this three-ring binder includes resources such as templates and rubrics.

Northwest Regional Education Laboratory connects assessment and writing in its 6+1Trait Writing professional development and resources. More information can be found at their website: <http://www.nwrel.org/assessment/departement.php?d=1>

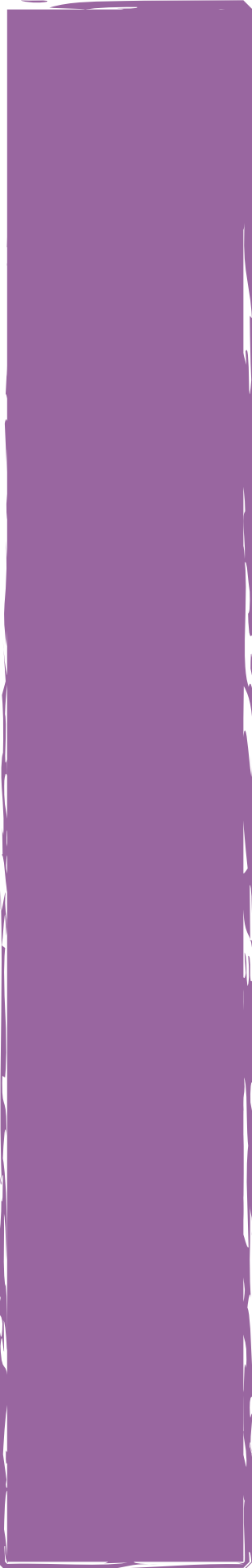
ReadWriteThink is a website sponsored by the National Council of Teachers of English and the International Reading Association. It contains resources related to improving reading and writing: www.readwritethink.org

The National Writing Project is centered on improving the teaching of writing. Information on its books and professional development can be found on its website: www.writingproject.org.

Resources on Dialogue and Questioning:

The National Paideia Center (www.paideia.org) provides free online curriculum resources for planning and implementing seminars in your classroom. They also provide professional development.

The Critical Thinking Community promotes critical thinking in schools. They include resources on questioning: www.criticalthinking.org.



The SERVE Center at UNCG, under the leadership of Dr. Ludwig David van Broekhuizen, is an education organization with the mission to promote and support the continuous improvement of educational opportunities for all learners in the Southeast. The organization's commitment to continuous improvement is manifest in an applied research-to-practice model that drives all of its work. Building on research, professional wisdom, and craft knowledge, SERVE staff members develop tools, processes, and interventions designed to assist practitioners and policymakers with their work. SERVE's ultimate goal is to raise the level of student achievement in the region. Evaluation of the impact of these activities combined with input from stakeholders expands SERVE's knowledge base and informs future research.

This rigorous and practical approach to research and development is supported by an experienced staff strategically located throughout the region. This staff is highly skilled in providing needs assessment services, conducting applied research in schools, and developing processes, products, and programs that support educational improvement and increase student achievement. In the last three years, in addition to its basic research and development work with over 170 southeastern schools, SERVE staff provided technical assistance and training to more than 18,000 teachers and administrators across the region.

The SERVE Center is governed by a board of directors that includes the governors, chief state school officers, educators, legislators, and private sector leaders from Alabama, Florida, Georgia, Mississippi, North Carolina, and South Carolina.

SERVE's operational core is the Regional Educational Laboratory. Funded by the U.S. Department of Education's Institute of Education Sciences, the Regional Educational Laboratory for the Southeast is one of ten Laboratories providing research-based information and services to all 50 states and territories. These Laboratories form a nationwide education knowledge network, building a bank of information and resources shared and disseminated nationally and regionally to improve student achievement. SERVE's National Leadership Area, Expanded Learning Opportunities, focuses on improving student outcomes through the use of exemplary pre-K and extended-day programs.

SERVE Center

at the University of North Carolina

at Greensboro

P.O. Box 5367 • Greensboro, NC 27435
800-755-3277 • 336-315-7400 • Fax 336-315-7457

Ludwig David van Broekhuizen, Ph.D., Executive Director

WWW.SERVE.ORG