Advanced Teaching Methods for the Technology Classroom

Stephen Petrina
The University of British Columbia, Canada

Information Science Publishing
Hershey • London • Melbourne • Singapore

Chapter X

Assessment and Evaluation

Introduction

Some teachers view assessment as a necessary evil. Some view assessment as their only real tool of discipline and power. Still other teachers view assessment as an integral part of C&I, and the pivotal practice around which teaching methods and communication turns. Most teachers appreciate local, teacher-controlled assessment and loathe the high stakes assessment that produces anxiety, fear, and competitive tactics. For many administrators, parents and politicians, assessment has its justifications in accountability to standards. Indeed, it is difficult to navigate through the various forms of assessment and perspectives on assessment that teachers face on a daily basis. Everyday assessment entails hundreds of observations that teachers make of their students. This involves informal discussions, feedback and deliberate, staged activities and performances. Assessment involves volumes of documentary evidence, from daily assignments, quizzes, and tests to observations, projects, and digital artifacts. In its most stereotypical form, assessment in technology studies simply meant putting a mark on a completed project, much like a merchant places a price on a product. By current standards, this was inauthentic assessment. Since the late 1980s and early 1990s, authentic assessment has transformed the way we think about and carry out assessments in the schools. Technologies of assessment had similar effects.

Evaluation, which typically pertains to facility, program, or teacher evaluation, has conflicts and interpretations that are similar to those of assessment. With both assessment and evaluation, the goals are to provide feedback, to rank or sort and to provide a means of communication. However, in many cases there is a lot at stake for those who are being assessed and evaluated. It is no secret that, in light of these stakes, students can resort to desperate means to beat the assessment system. On the Web, an entire market for cheating has been generated in response to demands for devices to beat the system. The purpose of this chapter is to provide an overview of assessment and evaluation. We will focus on the types of assessments and evaluations that are complementary to practice in technology studies. We will also raise fundamental questions regarding the relevance of high stakes tests of technological literacy.

Assessment in Technology Studies

Currently, I administer a *Scale of Design Capability* each year to my new groups of students. This scale tells me, with some degree of accuracy, the varied levels of the students and exactly who will succeed as a design and technology teacher. The scale is tuned to a simple, particular performance. I have the entire group remove their shoes and, with my scale safely secured on my clipboard, monitor each student's

Table 1. Scale of design capability

Sorting process	**Student**	**Possible**
All shoes sorted		5
Small sizes sorted first		5
Boots sorted first		5
Trainers sorted second		5
Other shoe types sorted last		5
Shoes sorted within 6 minute limit		5
Sorting outcome	**Student**	**Possible**
Size presented from small to large		5
Shoes presented by colors (blues with blues, reds with reds, etc.)		5

Table 1. Scale of design capability

Shoes facing in same direction (toes in 1 dir.)	5	
Shoes placed accurately in four rows and three columns	5	
Shoe pairs placed 25-30mm from each other	5	
Shoes cleaned and dried	5	
Design capability =		**/ 60**
Attitude		
Shoes sorted with positive attitude		/ 5
Total		**/ 65**

design capability. Each student is given one trial to sort the shoes and I assess accordingly. Here is the scale that I have refined over the years:

After administering the scale, I tally up my marks and rank the students. I convert the marks to percentages. It is a good measuring scale because only about one-quarter of the students score higher than the mean, which is typically around 75%. Design is something that not everyone is good at and this scale basically proves it. I do *not* give them the scale prior to their performance, as this would remove the element of true design capability. Their identification with the criteria would spoil the performance and eliminate spontaneity. The entire process of assessment provides me with a pretty good judgment about who can design and who is weak. Of course there is always room for the students to improve. I have come a long way, as I am now rating process as well as product. The scale quantifies both process and product. I really like administering this scale because it allows me to model the way we want to assess students in the schools!

Pause: *Is there anything wrong with this scenario? Is the Scale of Design Capability valid? How authentic is this process of assessment?*

When I taught drafting and CAD, I merely assessed the students' drawings and plotted files. I had objective criteria that I used for each drawing, but the result was that

I was grading products and not students. I often got caught up in the assessment of artifacts rather than students. As mentioned in Chapter II, I wrote NEATNESS--2 across messy drawings to indicate my assessment of how the drawings looked. I deducted marks for ACCURACY as I assessed the solutions and for STANDARDIZATION as I assessed adherence to conventions.

Pause: *Is there anything wrong with this scenario? Was my assessment of drawings and files valid? How authentic is this process of assessment?*

For the CAD courses, I created a final exam that was comprehensive. It addressed all of the content covered in the course and I sampled from each unit and topic on the content outline. The exam consisted of 100 problems (true-false, multiple choice and matching) and was quite challenging for the students. The exam sorted the students fairly well but not quite according to a normal distribution (i.e., Bell Curve). The exam did not involve any problems that required the students to use CAD. The exam was worth 20% of the final grade. Drawing and modeling assignments were used for 65% of the final grade. A heavy emphasis was already placed on drawing and modeling so there was no need to require more of this in the exam. After all, there was more to the course than drawing and modeling.

Pause again: *Is there anything wrong with this scenario? Was my exam valid? How authentic is this process of assessment?*

Projects

Projects typically culminate in an artifact, medium, or performance that relates to the original purpose. As explained in the previous chapter, projects are not things, but may culminate in things. Projects or other forms of evidence for grading are not merely produced for the purpose of assessment. Projects and their artifacts ought to be produced in response to the larger aims, ends, or objectives of the course or program. As indicated in the last chapter, projects should not be seen as ends in themselves. Rather, the intended artifacts of C&I ought to disclose conditions of modern life. All too often, as Custer (1996) observed, teachers get caught in one of two traps. The first is the “Project Trap,” in which the artifacts of the project are selected as ends in themselves and the only good reason for their adoption. Here, the cart is placed before the horse (i.e., the course exists so the students can create this or that artifact). The second is the “Neat Activity Trap.” Here, teachers select projects because they are entertaining, rather than by identifying what specific content, emotions, and skills will be reinforced. Teachers often end up adapting their course to fit activities rather than adopting activities to fit the course.

Artifacts and media range from digital images and text to three-dimensional models, drawings, paintings, sculptures, songs, and useable products. As indicated in the previous chapter, some technology teachers argue that students must take a tangible artifact home. Teachers who tend to over-emphasize projects typically assess products for quality, to the neglect of processes—they assess projects rather than students. The product overshadows the value of the process. After the completion, the product is assessed for quality. In the woodworking courses of industrial education, a clock or table was assessed and given a mark of quality. Thirty products for thirty students were often lined up and given a mark. In information technology, spreadsheets were created and assessed. Of course, this still takes place and we are all guilty of this practice. But this kind of practice is not unique to technology studies. Math and science teachers assess problems; art teachers assess paintings. However, the trend in all subjects is to refocus on the process rather than the product.

In subjects such as art, business, home economics and technology where projects have dominated the school curriculum for well over a century, the transition from assessing projects to assessing processes is a special challenge. Learning about design and technology requires thatproducts be submitted to public critiques. Critiques or judgments of artifacts are a central facet of the design process (see, Figure 6). Critiques of the products of technology are also necessary modes of public feedback for engineers, designers and other technologists. Artifacts, including artistic artifacts, must never be above judgment. Designers and technologists tend to be quite pragmatic, assuming that solutions and "what works" are of interest and the process is secondary. Artists tend to look for aesthetic qualities in products. Some teachers react to the general decline of quality in technological practices and products and take a hard stand with the quality of their students' projects. So there are some very good reasons to assess the quality of projects.

In addition, the trend toward the assessment of processes is accompanied by a trend toward outcomes-based education (OBE) and norm-referenced assessment. OBE was manifested in the increases of standardized tests and standards throughout the 1990s. "Outcomes," for most administrators, parents and politicians, mean scores on standardized tests. These types of high-stakes assessments are norm referenced—comparisons are made with national and international averages and norms. In effect, technology educators, like art educators, find themselves reacting to two contradictory forces or trends. One trend is toward the assessment of processes and the other is toward the assessment of outcomes *via* standardized tests. As indicated in Chapter VIII, the content of technology studies is the primary justification for the subject. Hence, we cannot dismiss standardized tests of content. We have to respond to both forces at the same time.

In Chapter VIII, the arguments for a process-based curriculum were explained. The example of CAD was used, where instead of concentrating on the commands of AutoCAD and associated skills, teachers began to shift their efforts to the processes of communication, visualization, representation, detailing, documentation, presentation, and modeling. In general, process-based curriculum includes a shift toward intellectual processes such as observing, analyzing, computing, measuring, predicting, experimenting, modeling, creating and communicating. The challenge is *how* to assess the students. Projects provide *one* rather convenient form of *evidence* of these processes. Projects can be a powerful instrument of authentic assessment *if* we assess more than merely the artifacts. The project is *not* the artifact.

Authentic Assessment

A fair program of assessment demands a range of forms of documentation and evidence, such as experiments, images, innovations, interviews, quizzes, observations, presentations, problems, projects, portfolios, recordings, and rubrics. A fair program demands that we assess authentically. The techniques of authentic assessment help provide for a fair assessment of both products and processes. At this point, and to address these issues, it important that we have some working definitions of assessment.

- **Assessment:** Monitoring, documenting and communicating levels of quality and quantity of performance. Assessment is done in order to: (1) provide feedback for learning & growth; (2) rank or sort according to some characteristic; and (3) provide means of communication with parents, administrators, teachers, etc.
- **Formative assessment:** Assessment that is progressive in that the students'' progress is monitored and communicated at different periods in time throughout the course, unit, term, etc. "In-progress" assessment.
- **Summative assessment:** Assessment is final in that the students' performance is assessed at the end of a unit or course. "Final" assessment.
- **Authentic assessment:** Assess the genuine, "real," or actual thing (person, performance, etc.); Assess fairly; Use assessment to enhance learning.

Authentic assessment means that we assess the genuine, "real" or actual thing (person, performance, etc.). It means that we assess fairly and use assessment to enhance learning. Authentic assessment is that which has meaning in itself, has value beyond the classroom and is meaningful to the students. Assessment that directs and redirects learning is by necessity flexible and deals with a wide array of what students know, feel, and can do. Where quizzes and exams traditionally test for low-level cognitive processes (e.g., recall, recognition) and are primarily summative assessments, authentic assessments allow us to assess a wide expression of dispositions, knowledge, and skills and is primarily formative assessment. Assessment should be flexible enough to accommodate various learning styles and multiple intelligences. Authentic assessment has the potential to be an equitable and fair way of assessing and judging experience and expressions of competence.

Technology studies, with its experience-based nature and project or problem-based orientation, is well attuned to authentic assessments. Yet, as indicated in the section on projects, technology teachers have been slow to shift their assessments toward authentic techniques. Project assessment is not *ipso facto* authentic assessment. The range of activities in technology studies nevertheless lends itself well to authentic techniques. The most effective techniques include portfolios, performances and criterion-referenced assessment (rubrics) (BC MoE, 1994a).

Table 2. Techniques of authentic assessment

Portfolio assessment: Assess results and evidence of results over time.

A good portfolio...

- Has a clear purpose that was communicated clearly to all involved.
- Organizes level-appropriate activities that students are familiar with.
- Organizes evidence of process (as opposed to merely collecting products).
- Requires students to describe contents.
- Provides ample space to store contents.

Performance assessment: Assess the performance.

A good performance is...

- Congruent with the purposes of assessment.
- Interesting, challenging and fair for all students.
- Authentic; promotes transfer to other performances.
- Reflects intended outcomes and goals.
- Appropriate for the students' level of development.
- Directed by clear expectations of what is to be done and under what conditions.
- Directed by adequate information for successful completion.

Criterion-referenced assessment (Rubrics): Assess according to predetermined and communicated criteria.

A good criterion or rubric...

- Communicates essential standards of achievement.
- Operationalizes outcomes.
- Applies across contexts for similar behaviors.
- Focuses on current instruction, not prior learning.
- Is essential to judge the performance adequately.
- Communicates to all (students, teachers, parents) what is critical to successful levels of performance.

Table 3. Characteristics of authentic assessment (Kerka, 1995)

- Engaging, meaningful, worthy problems or tasks that match the content and outcomes of C&I.
- Real-life applicability.
- Multistaged demonstrations of knowing, knowing why and knowing how.
- Emphasis on process and product, conveying that both development and achievement matter.
- Rich, multidimensional, varied formats, both on-demand (in-class projects), and cumulative (portfolios).
- Opportunities for learner self-evaluation.
- Cognitive complexity requiring higher order thinking skills.
- Clear, concise, and openly communicated standards.
- Fairness in rating and scoring procedures and their application.

There are eight general questions that ought to guide assessment and help teachers link assessment with C&I: What should learners know and be able to do? What emotional, cognitive, and sensorimotor skills should they demonstrate? What types of activities, problems, or tasks involve those skills? What concepts or principles should be applied in performing those tasks? What are the reasons for this assessment? What use will be made of the results? By whom? What criteria should be used? The key to planning for authentic assessment is to plan for curriculum, instruction and assessment at the same time. The single most significant characteristic of assessing authentically is that assessment matches C&I.

If teachers use techniques of authentic assessment and use quizzes or tests that are fairly objective, the actual grading or marking of individual items should not be an issue. Nevertheless, anytime a teacher assigns a numerical or letter grade to a performance, s/he must be accurate, careful, consistent, and honest and systematic. Teachers must keep their biases in check as much as is humanly possible. Techniques of authentic assessment are of great assistance for this.

Portfolio Assessment

A portfolio is a *collection* of documents that attest to performances and proficiencies. A portfolio is an *attestation* of work—entries in the portfolio should have a brief description of what the selections attest to. The selections or entries should attest to particular aptitudes, knowledge, proficiencies or values. A portfolio provides evidence of dispositions, knowledge, and skills; it is a *collection* of evidence. This evidence represents a selection (typically the student's selection) of ideas and work, from various points and courses in time, regarding a student's performance and proficiency. Technology portfolios are typically ongoing projects, providing evidence of the students' progress in technology studies. Portfolios keep open the question of expression and may include combinations of artifacts, attestations, and productions (notes, drafts, journal entries, sketches, lesson plans, letters, drawings, programs, photographs, videos, audios, models, etc.). Portfolio assessment emphasizes the importance of student responsibility in their education by including them in the assessment process and by involving them in goal setting and criteria.

Portfolios serve as a catalyst for reflection on growth and development—students will have an organized collection of their work to review and think about as they transition from level to level. Portfolios also serve as a record for presenting oneself to potential employers or institutions of advanced education. With their various uses beyond the classroom, they are one of the most important means of authentic assessment.

Because of the varied artifacts created in subjects such as art, design, and technology, there is a tradition of portfolio use. But this tradition is limited to professional schools and practices. Professional animators, architects, artists, designers, and engineers are dependent on their portfolios to get them contracts, jobs, and advancements. Many maintain their portfolios over an entire career of 30-35 years of practice! Yet, for the most part, technology teachers were reluctant to use portfolios as a means of student assessment until the 1990s. This decade marked philosophical changes in the transitions from industrial to technology education and educational to information technology. Now, it is quite common to find the use of portfolios in labs and workshops. Digital technologies provided a catalyst for the adoption of e-portfolios in technology studies. In fact, entire districts and educational systems have turned toward e-portfolios. It is relatively easy to digitize 3D artifacts and place an entire e-portfolio online. Mark Sanders (2000) at Virginia Polytechnic and State University, began to place his students' e-portfolios online in the mid 1990s. Numerous e-portfolio applications, from proprietary to open source, are now readily accessible and convenient. However, most technology teachers find that constraints of e-portfolio applications standardize and limit design options for students. An e-portfolio is not a distinct type of portfolio, rather it is a mode of presentation.

There are three types of portfolios: Working Portfolio, Presentation Portfolio, and Cumulative Portfolio (Table 4). A working portfolio is "an ongoing collection of self-selected samples of work that are used to highlight the students' efforts, progress, achievement, and reflections." A presentation portfolio includes samples selected by the student and teacher. These samples are then *presented* to the teacher, potential employers or advanced educational institutions. A cumulative portfolio includes selections from working and presentation portfolios over long periods of time (i.e., years) (BC MOE, 1994b, p. 4-5).

Table 4. Types of portfolios (Adapted from BC MOE, 1994b)

Portfolio	Purpose	Content	Audience
Working	· To help students assess their work. · To help students observe patterns in their progress.	· Many samples of student work from many or one subject.	· Student and teacher.
Presentation	· To assist students in communication about their work and progress.	· Selected samples that represent a few chosen aspects of student work.	· Teachers. · Parents. · Future employers. · Advanced educational institutions.
Cumulative	· To help educators know where their students are in their education. · To assist in planning student programs. · To help students prepare for the real world.	· Selected samples of student work. · Student progress reports.	· Students. · Future teachers. · Administrators. · Future employers. · Advanced educational institutions.

Working portfolios help *describe* what students are doing and what they can do. The variety of samples included provides an indication of competencies. Presentation portfolios *showcase* the students' work. A select few samples are provided to highlight the best of what a student can do. Cumulative portfolios are *evaluative* in that they allow for a longitudinal judgment on the students' progress (Hoepfl, 1993).

My experience with digital portfolios suggests that it is best to provide a checklist of the artifacts students are expected to provide in their portfolio. They can select the artifacts for the various categories.

Performance Assessment

Performance assessment was a response to the limitations of standardized, norm-referenced testing. Performance assessment means that students demonstrate what they can do—we assess the actual performance of challenges and tasks. For example, if the students are expected to solve problems, then we must assess the way they perform in the process of problem-solving. We cannot merely assess the products of the problems they solve. The assessment of performance can be informal (observations of everyday progress) or formal (recording a structured event or situation) (BC MOE, 1994a).

Performance assessment derives from the world of work. If a manager wants to assess an employee, rarely will they give a job-specific test. They observe the employee's performance on the job—at the job site. However, managers are much less interested in authentic assessment than teachers. Hence, they often rely on secret assessments and assessors such as "mystery shoppers" hired to rate the performance of sales clerks or associates, and phone or computer monitoring software to monitor and assess receptionists. Authentic assessment requires that the students be informed when their performances are being assessed and the criteria used for assessment. Teachers are challenged to assess students "in the task" and communicate the means of assessment. Performances in technology include the creation of digital media, experiments, models, open-ended design and engineering problems, prototypes, research projects, technology fair projects, and an assessment of a new technology.

There are a few issues that teachers must be cognizant of when they use performance assessment techniques. Performances raise the question of developmental sensitivity, as younger students will often simply mimic adult capabilities and roles. There is also the problem of merely teaching students to perform, requiring cosmetic expertise at the expense of other important goals. Assessments structured around tasks, instead of constructs (such as student comprehension), risk a returning to simple behavioral assessments. Quite often, unintended consequences stemming from the specific tasks will unfold. For example, students mocking a required task may in fact be demonstrating a deep understanding of the task and in effect the irrelevance of it their lives. These unintended consequences need to be documented as they arise from the approach to assessment. Teachers using both outcomes-based testing and performance assessments are often left to deal with potentially conflicting teaching methods and educational goals. Their students are left to resolve the conflicts on their own. In summary, performance assessment requires careful attention to the setting, specification of criteria, multiple samples of student performance, and attention toward evidence and validity.

To be effective, performance assessment must be criterion-referenced assessment. This means that student performances are judged against pre-set criteria and performance standards rather than against each other. Performance assessment is impossible without clear criteria established well before the process begins. This is one of the most difficult challenges for teachers.

Criterion-Referenced Assessment and Rubrics

Assessment criteria are characteristics or guidelines with which we judge the performance of students. Assessment criteria are created and given to students prior to the event, situation, problem or project that will be assessed. Criterion-referenced assessment is based on the criteria created for different levels of performance (e.g., excellent, good, satisfactory, minimal, poor). The performance of students is compared to the criteria of different levels to provide feedback to their performance and determine their standings for reporting marks. In norm-referenced assessment, students are compared to the norm or group. In criterion-referenced assessment, students are compared with criteria rather than each other.

Some of the earliest forms of criterion-referenced assessment were established in England during the mid 1970s for the craft, design, and technology projects. The criteria represented an attempt to make the assessing of projects consistent and objective. Although not named as such, these were early attempts to make assessment authentic. One rather comprehensive approach took the form of a matrix or what we refer to as a rubric (Starmer, 1974). Criteria are developed for each concept at each level, titled Conception, Designing, Production, Valuation, Personal Qualities. The criteria for designing are provided in Table 5. The key concepts are listed in the left column of the rubric and levels of performance, or marks/points, along the top row. These criteria were given to the students prior to their performance and altered a bit to meet particular problems and projects.

From this tradition of rubrics for judging design and technological action, cognition and emotion, British technology educators have been quite progressive in their as-

Table 5. Expansion of design criteria in CDT Rubric (Adapted from Starmer, 1974)

	1	2	3	4	5
Designing · Formulation of design.	The project commenced with minimal thought of the progression of work	One solution considered but sketchily prepared	Only one solution but a good design of this prepared	A design prepared after alternative approaches have been mooted	A design formulated after careful consideration of several different approaches
· Testing of workability.	Only superficial testing	Reluctant to check any aspects	Tested and checked some aspects	Tested and checked major aspects where practically possible	Tested and checked all aspects where practically possible
· Suitability of design.	Specification badly adhered to	Specification adhered to only in a few aspects	Design fulfills specification in almost all aspects	Design fulfills specification but is too complex or not comprehensive enough	Design fulfills specification simply and comprehensively within limits of design brief

sessment techniques. Researchers adapted these original rubrics, entered the labs and workshops, and rated or scored the practice of the D&T students. National assessments were made in the late 1980s by researchers such as Richard Kimbell (1997) at Goldsmiths College in London. While rubrics were constructed in Canada and the U.S. during the mid 1980s, North Americans have been slower to accommodate criterion-referenced techniques.

Criterion-referenced assessment is nevertheless catching on with the growing popularity and utility of rubrics. Rubrics are the primary means for implementing criterion-referenced assessment. They are scoring or rating devices "designed to assist in the process of clarifying, communicating, and assessing expectations. Rubrics are grading tools which contain specific information about what is expected of students based on selected or defined criteria" (Custer, 1996, p. 29). One innovative aspect of rubrics is in the detailing of criteria or performance standards across multiple levels. Scales that merely left a number or letter to circle were transformed into much more detailed scales that communicate explicit criteria across multiple levels to students. In an earlier section, we inquired about the assessment of projects. We noted that technology teachers were in a (bad) habit of assessing projects by judiciously focusing on the artifact—on the tangible product of the project. They might have measured the artifact to determine how closely the finished sizes adhered to the blueprint dimensions in a production course. They might have measured the registration of screen prints superimposed on one another in a graphics course. They might have measured margins and the alignment of images and text in an information technology course. We noted that a formal assessment of the process was rarely made. A second innovative aspect of rubrics is the focus on process. Rubrics help us to focus on the process in an open manner—we create rubrics to give to students prior to our assessment. We create rubrics for assessing both processes and products. They help remove some of the subjectivity often associated with the assessment of processes. Generic and template rubrics help eliminate the need for constant adaptation to particular activities and projects. For example, David Romani (2002), a teacher in Vancouver uses a "generic" employability skills rubric as a complement to the assessment of activities and projects that the students complete in his secondary school (Figure 4).

The actual rubric has five levels of performance. Specific rubrics are created for the individual artifacts he assesses. Custer (1996) highlighted a generic rubric used by Jeanne Kirchoff (1996) in her school in Troy, Missouri (Figure 5). This rubric is used to assess the type of small group work and teamwork demanded by group design briefs and projects.

These rubrics can be customized and tailored for local courses and schools. A team of BC technology educators developed the following rubric for Web design in 1996 (Table 8) (BC MOE, 1996).

When general rubrics for employability and social skills are combined with rubrics for specific technology tasks, we have a powerful set of assessment tools. The

Table 6. Employability skills Rubric (Adapted from Romani, 2002)

Criteria	Level 1(A)	Level 2(B)	Level 3(C)
Communication	Successfully conveys and retrieves information in written, oral and sketch format.	Clearly articulates & supports information in two of the three formats	Evaluates the importance/ relevance of each format and its sources.
Autonomous learning	Models relevant techniques to assist the brain in acquiring new information	Successfully demonstrates the patience and perseverance needed to accept new information that is ongoing.	Works quietly on assignment, making good progress and asks questions when necessary
Innovation & creativity	Thinks laterally and derives new solutions from previously unrelated items.	Applies personal experiences to formulate new and different ways of attacking problems.	Makes new patterns of materials, words, or ideas
Technological literacy	Able to respond rationally to ethical dilemmas caused by technology	Able to value the benefits and assess the risks associated with technology	Understand how technological systems are designed, used and controlled
Critical thinking	High ability to think and reason	Inquiring and discovering information; then appraising the evidence	Identifies errors in information and processes
Common sense	Displays an upbeat—think before you do mentality	Considers possible outcomes for own actions	Polite & articulate individual who profits from learning.
Teamwork	Demonstrates an effective ability to listen, respect, & persuade others in a cooperative manner.	Questions and discusses approach with other team members	A keen willingness to assist others while respecting their work and personal space.
Attitude	Demonstrates an upbeat, polite and responsible disposition.	Assists others when called upon	Sees the positive in most tasks/activities
Decision making	Takes consistent action toward attaining preset goals by utilizing time	Varies action toward goals therefore lowering end results	Demonstrates action toward various goals.

Table 7. Character traits Rubric (Adapted from Kirchoff, 1996)

	3	2	1
Courtesy	Treated each member with complete courtesy and respect at all times	Treated each member with courtesy and respect most of the time	Courtesy and respect for others was lacking
Rules followed	Responded well to all rules	Responded well to rules most of the time with few lapses	Seldom stayed within the rules on her/his own

Table 7. continued

Performance of task	Worked diligently to do her/his part of the task	Needed some prompting to stay on or complete task	Caused confusion due to lack of staying on task
Cooperation with team members	Cooperated with team at all times to make completion of task smooth	Cooperated with team most of the time but needed reminding	Rarely cooperated with team without constant reminding

construction of rubrics takes time, but the wheel does not need to be reinvented for every task. There are effective rubrics for a wide range of technology processes and tasks that can be found on the internet and in curriculum documents. Rubrics help take the "guesswork" and mystery out of assessment by communicating clear performance criteria. They provide levels and standards of performance to guide both students and teachers. Rubrics provide criteria for teachers to back-up their marks and grades in an objective way. They also provide a clean way of shifting from letter grades to comments for communicating to parents. Rubrics are indispensable for assessing cognitive process such as problem-solving.

Table 8. Web page design Rubric (Adapted from BC MOE, 1996)

Outstanding	· The document incorporates the correct use of HTML-plus, enhanced HTML, or both; is free of structural and syntactical errors; and uses coding that is clearly and consistently formatted. · Links are among the best available on the selected theme, and all sites are listed clearly and concisely to facilitate their use. • The page is highly visually appealing.
Good	· The document is free of structural and syntactical errors, and coding is clearly and consistently formatted. · Links are useful and well documented. • The page is visually appealing.
Satisfactory	· The document is free of major structural and syntactical errors, and the coding is understandable. · Links are functional, and the page contains basic documentation. • The page is free of major visual formatting flaws.
Less than satisfactory	· The document contains major structural or syntactical errors, and coding is difficult to interpret the page is difficult to use or does not function. · Some or all links are not functional, or the page does not contain documentation on the links. • The page has major visual formatting flaws.

Problem-Solving

While problem-solving is one of the most heavily emphasized methods in technology studies, it is a challenge to authentically assess. Should we assess the intellectual processes used to solve problems, the creativity applied in resolving the problems or the solution to the problem itself? Should we assess the process, product or both? Should we assess qualitative issues (complexity and how it was solved) or quantitative issues (how fast or how many?), or both? Can we develop criteria to judge the quality of problem-solving?

There are generally three types of problems. Simple problems are highly structured and usually have a correct solution. They can be represented in a straightforward way. Applied problems are structured but require the drawing together of diverse procedures and background information. The form of the solution is defined or the sense of the form is implicit. Complex problems are loosely structured and open ended. They may require the development of new processes and strategies to solve. The process required may be ambiguous to use, there may not be an established or correct answer and the solution may be difficult to represent. How can we assess applied and complex problems?

One helpful way of authentically assessing problem-solving is to reduce it to four performance aspects. Engagement refers to the extent to which the student identifies something as a problem and becomes engaged in solving it. Background knowledge refers to the extent to which the student accesses and uses appropriate information. Process refers to the extent to which the student knows and can use appropriate problem-solving strategies. Representation refers to how effectively the student can communicate his or her solution and the thinking and processes behind it.

For the most part, the students' interests in solving a problem define whether a problem exists. When the teacher sets the problems, low levels of engagement suggest that the students do not identify the problem as a problem. It would be inappropriate to assess problem-solving behavior for simple problems. When the students identify something as a problem, then it is appropriate to assess problem-solving behavior. As problem solvers go about the business of resolving problems, they access and analyze prior knowledge to bridge the gaps between what they know and want to find out. The background knowledge necessary to solve a problem is extremely important and students will demonstrate the degree to which they are accessing the information needed. Effective problem-solving involves recognizing what to do, when to do it and how to do it. Problem-solving competence and maturity reflects a growing repertoire of strategies. This may mean that students draw on certain problem-solving methods (Chapter V) or particular reasoning strategies (Chapter II). Teachers typically introduce strategies for their students to apply and may intervene to support their students' development in the course of a problem. The communication of problems and solutions need not be written. There are various means of representation that can be used (e.g., digital animation, drama, images, structures). Students may arrive at the same solution to problems and choose to represent the solutions in different ways (BC MOE, 1995a). Once technology teachers understand these four aspects of problem-solving, they can develop rubrics for authentic assessment (Table 9).

Different levels for each aspect or concept can be developed to differentiate novice from mature and advanced problem solvers. The key to the authentic assessment of problem-solving lies in its reduction to significant aspects and their elaboration within a rubric, such as in Table 7. If we assess engagement, background knowl-

Table 9. Problem-solving Rubric (Adapted from BC MOE, 1995a)

Problem Aspect / Level	Low	Average	Advanced
Engagement · Interest in problem. · Involvement in problem. · Defines problem.	· Little. · Off-track. · With difficulty.	· Wants to solve. · Seeks/needs reinforcement. · With some difficulty.	· Active/thoughtful. · Independent. · Clarifies/copes with ambiguity.
Background knowledge · Content knowledge. · Focus range. · Applies techniques (rules, methods, plans, algorithms). · Transfers knowledge.	· Many gaps. · Narrow. · Seldom. · None.	· Some gaps. · Narrow/ some new information. · May apply some. · Makes generalizations.	· Complete. · Finds missing info. · Applies techniques. · Uses knowledge from many situations.
Process · Recognizes what to do. · Applies strategies. · Uses alternatives. · Monitors progress.	· Unsure/loses sight of problem. · Uncertain. · Resistant. · No.	· Uncertain of approach. · Yes, can't explain why. · Seeks suggestions/ gets frustrated. · Seeks help.	· Capable/changes when necessary. · Clarifies ideas. · Explores unique procedures. · Functions independently.
Representation · Restates the problem. · Communicates about process. · Organizes solution.	· With difficulty. · With difficulty. · Partial/disorganized /incorrect.	· Restates some features. · Reflects on some processes. · Complete, but not thorough.	· Communicates details. · Describes thinking processes. · Thorough & organized.

edge, process, and representation, we can develop a comprehensive approach to the assessment of problem-solving.

Tests and Measurements

Tests refer to a broad group of instruments and practices for assessing and measuring action, cognition and emotion. There are tests of educational knowledge, dexterity, fitness, intelligence, racial prejudice and religion preference. There are medical examinations and tests, and psychological tests of emotional and mental health. On the Web, we can find tests for just about anything imaginable. Educational tests refer to a wide range including exams and quizzes, scales that deal with affective issues and feelings, as well as a range of tests of dexterity, speed, strength and skill. These may be administered in an oral format, written paper and pencil formats, computer automated format, or they may require physical manipulation and movement. A psychologist, teacher, or expert of some sort is needed to administer some of these while others can be self-administered. *Measurement* means the quantification and qualification of traits of action, cognition, and emotion as well as to the methodological and statistical techniques used in quantitative and qualitative assessment. Measurement may mean simple measures of central tendency (mean, median, mode), measures of item discrimination in an exam, or quite complex statistics that allow for confident, diagnostic and prognostic predictions of failure and success, criminality and recidivism, or disease and wellness. This section is limited to simple aspects of tests and measurements in educational practice with the focus on teacher-made tests.

Within our context of authentic assessment, tests and measurements are effective tools to supplement, rather than dominate, an assessment system. Authentic assessment does not mean that tests are inauthentic. Rather, authentic techniques establish a context and role for quizzes and tests that differs from their role outside of a system of authentic assessment. Testing was traditionally used for quality control such as maintaining rigor and standards of the discipline or of achievement, for sorting students according to test scores, and for the sake of preparing students for the testing processes of higher education. This last use is reductionism, where C&I are reduced from the practices and entrance requirements of universities. Other than this, the traditional uses of tests are important. Again, tests and measurements play a complementary rather than dominant role in authentic assessment. Only secondarily ought they serve the administrative purposes that dominate traditional uses. Remember, our first criterion for assessing our assessment and measurement techniques is that they primarily serve the learning process.

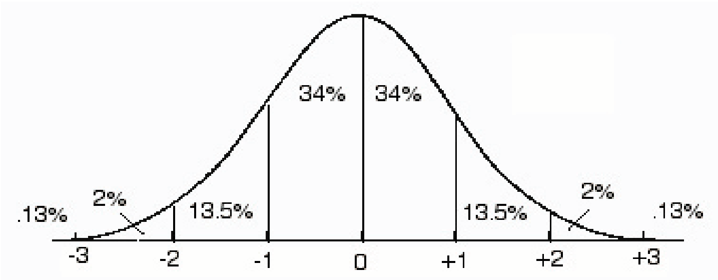
Tests and measurements play powerful roles in the lives of our students. They have the power to make or break students for life. Students who are diagnosed with a learning disability and treated with some therapeutic program of drugs and remediation will always carry this trauma. Students who are consistently belittled in the face of tests and measurements will internalize the frustrations of failure. Tests and measurements also have the power to help us monitor and direct the learning of our students—to detect deficiencies and proficiencies and be a positive force in their lives. With this type of power come responsibilities. Teachers have the responsibility to produce tests that are professional and of the highest quality. The National Council on Measurement in Education (1995) advocates, among others, the following responsibilities in a code of ethics:

- Ensure that assessment products and services are developed to meet applicable professional, technical, and legal standards.
- Develop assessment products and services that are as free as possible from bias due to characteristics irrelevant to the construct being measured, such as gender, ethnicity, race, socioeconomic status, disability, religion, age, sexuality, or national origin.
- Plan accommodations for groups of test takers with disabilities and other special needs when developing assessments.
- Use copyrighted materials in assessment products and services in accordance with the law.
- Protect the rights to privacy of those who are assessed as part of the assessment process.
- Develop reports and support materials that promote the understanding of assessment results.

One of the greatest challenges in tests and measurements is dealing with variability and diversity. Given the wide range of learning styles that students bring to education, we are challenged to assess in ways that respond to this wide range. In other words, if we are providing a wide range of activities that respond to different learning styles, then we necessarily owe it to our students to use assessments that allow for multiple intelligences and learning styles. Tests generally tend to force all students into a single learning style. Hence, as indicated, we should situate tests in a larger context of authentic assessment. This helps us to respond fairly to different learning styles but we are still confronted by the challenge of variance and diversity. The model of variability and diversity in tests and measurement is the Bell Curve or Normal Curve (Figure 1).

The Normal Curve was developed during the 19th century to account for the dispersion and distribution of certain biological and psychological traits. It is a form of regularity based on probability and random variation. If we measure the same trait of many cases that have differences caused by random variation, the frequency of

Figure 1. The normal curve



similarities and differences in this trait will take the shape of a normal distribution. To simplify, the concept is that if we take a measurement of some trait (e.g., height, weight) of a large number of people in a population, about 68% would be more or less average, about 13.5% would deviate one standard deviation above and below the norm, about 2% would deviate two standard deviations above and below and .13% three standard deviations. A standard deviation is a measure of distance from the mean. The point is that in populations and other social phenomena, about two-thirds would be around the average and the rest would be some distance higher or lower.

The normal curve presented a model for comparing students to a norm or mean. In the 1910s and 1920s, psychologists suggested that intelligence test scores were distributed on a normal curve—most students were average, some were sub-average or moron and feeble-minded, some were above average or gifted and the well-above average were genius. They also noted that the Stanford Achievement Test (SAT) sorted students according to the normal curve. A score of 500 on the SAT was and still is the average. From this practice of norm-referenced assessment, some psychometricians argued that the normal curve was invaluable for comparing students to national and international norms. The problem was that administrators eventually wanted their districts and schools to represent normal distributions and teachers adopted the mentality that their classes were mere samples of the normal distribution. The problem was that they overlooked one important detail of the bell curve: To approach a normal distribution we need the scores of thousands or hundreds of thousands of students taking the same test. Otherwise, we are dealing with small samples (i.e., 30 students in a class) and have to settle for so-called abnormal distributions. In fact, the "well curve," high on the ends and low in the middle, is a common distribution for classes as well as the size of business organizations. So we are left with the question of how do we know whether we have created a good test? If we can no longer use the normal curve as a model for our tests, where about 68% of our students would receive C's, 13% would get B's and D's, and about 3% would fail and 3% would get A's, how do we know if we have a good test? What determines an effective test?

Table 10. Principles of test design

1.	**Validity:** Does the assessment process or scale really measure what it purports to measure (i.e., design capability by sorting shoes)? Does it look authentic or seem appropriate to the students (**face validity**)? Does it address and cover what was taught (**content validity**)? Does it discriminate to assess the students' actual levels of achievement and performance (**concurrent validity**)? How well does the assessment or scale predict how a student will perform at a certain task in the future (**predictive validity**)?
2.	**Reliability:** Will I get the same results if I assessed the students again with the same scale, or if someone else assessed the students? To have validity, the assessment or scale must have reliability.
3.	**Objectivity:** Do the items of assessment processes and scales offer a clear interpretation for the students?
4.	**Discrimination:** Does the assessment process or scale lend itself to the challenge of identifying different levels of comprehension or performance?
5.	**Comprehensiveness:** Does the assessment or scale sample the full range of content within the specified unit or course?
6.	**Usability:** Is the assessment process or scale designed so that it can be administered and scored with relative ease?

Constructing Effective Tests

Simply put, a good quiz or test is one that measures what it is supposed to measure. In our context of authentic assessment, a good test is one that is complementary to learning and instruction. Good tests are interdependent with C&I. Good tests are those that discriminate (between those who know their stuff and those that do not) but whose value is not dependent on discriminating to the point of a normal distribution. Good tests are moderately difficult and practical to administer. We judge teacher-made tests by how well they enhance learning and instruction. Teacher-made tests that adhere to general principles of test design will always be the best tests. Tests are extremely valuable for monitoring student progress, but must be carefully constructed with the principles of test design in mind to provide adequate feedback to students and teachers. There are seven general principles with which we judge tests: validity, reliability, objectivity, discrimination, comprehensiveness, and usability. Definitions are provided in Table 10.

The first step in constructing a good test is organization. Just as we need a blueprint in construction, manufacturing, and design, we need a test blueprint for constructing a good test. The test blueprint allows us to see the big picture (scope and sequence) while focusing on the logistics and pragmatics of the test itself. Test designers recommend that we create a grid or matrix with performance levels of our affective, cognitive, or psychomotor domain in columns and content or objectives addressed in our unit or course distributed among the rows (Table 11). If we include our entire outline of content or all our objectives in the rows—effectively our scope and sequence of content—we can then sample from the outline and tally the test items with levels that we want to test. Some content or objectives will be assessed through performances and portfolios. The blueprint provides a way of sampling what will be quizzed or tested. In Table 11, the content organizers from the ITEA's technology standards and the cognitive domain are used.

Once the blueprint is completed, the next step is determining the types of items that will be used in the test. The blueprint will serve two purposes at this stage.

Table 11. Test blueprint

Technology standards test	Performance Levels							
	Knowledge	Comprehension	Application	Analysis	Synthesis	Evaluation	Total	Percentage
Content Outline	Number of Items							
Characteristics and scope of technology	2				1		3	
Core concepts of technology			2				2	
Relationships among technology and other fields	2						2	
Cultural, social, economic and political effects of technology						2	2	
Effects of technology on the environment				2			2	
Etc.								

Table 12. Types of test items

Selected response	Constructed response
• Multiple-choice.	• Completion.
• True-false.	• Fill-in-the-blank.
• Analogies.	• Forced-choice.
• Sequences.	• Vignettes.
• Matching.	• Rearrangement or continuity items.
	• Essay.

One, it provides a quick reference for constructing the test. Two, the content topics sampled in the blueprint will serve as the subtitles in the test. Rather than a random organization of test items, the items will be grouped by content. There are generally two types of items: selected response and constructed response.

Both types of items are challenging to write, and both have their place in assessment. The primary principles of test design tilt toward the student. In other words, we ought to construct tests with the students' best interest in mind. Most of us have taken tests where these principles were clearly not operative. Quite often, tests are constructed in the interest of the teacher. These types of tests were created in the last minute, tested for things that the teacher thought were important but seemed unimportant to the students, and were technically inadequate in terms of clarity, readability and subjectivity. These tests were invalid. This is precisely what we want to avoid. We want to create tests that the students feel are valid. Certainly teachers need to keep their own interests in mind, but there are constructive ways of doing this. Recall that usability, which works in the teachers' favor, is our sixth principle of test design. Techniques and guidelines of test design are guided by our six principles.

After constructing the test blueprint then next step is constructing the test. Good tests consist of four separate sections (Table 13). The first section is the title block, the second section is the answer column on the left of the page, the third section is for item directions and the fourth is the test item section. The title block provides the test title and the spaces for student information. Place directions for answering the type of items in the section at the top of each test item section. Group items by content and, on comprehensive tests, these sections ought to have a content heading. Reserve a column on the left side of the test for the student answers. This allows for ease of scoring on paper and pencil tests. English and Romance languages are read from left to right, and it is easier to read from your answer key to the students' answers from left to right. The example format from a cumulative, final exam for technology teachers elaborates these sections with details.

Keeping the overall design format of quizzes and tests in mind, we have to choose the types of test items to correspond with the levels and objectives to be sampled. Will we be using selected response or constructed response items, or a combination?

Table 13. Test sections

Test title & title block (Space for Student Name, Date, etc.)	
Answer Column	Item type directions True-false items Multiple choice items Matching items Completion items

Final Exam
Curriculum and instruction for technology teachers

Name:_______________________________ Date:________________________

General Directions: You will have 2 hours to complete this final exam. Please DO NOT use notes or handouts for the exam. **GOOD LUCK!**
100 Points Total

True-False: Choose the best answer. Circle "T" if the statement is true and "F" if it is false.

History of technology studies

T F 1. The British Arts and Crafts movement emphasized design for manufacturing.
T F 2. The audio-visual movement pre-dates the educational technology movement.

Demonstrating

T F 16. There is a single "best" sequence for demonstrating an application, tool, machine, or process.

Assessment

T F 59. In criterion assessment, good criteria are essential to judging the performance adequately.

Matching: Choose the concept that best matches the statement. Place the letter of the concept on the right next to the numbered statement on the left. Concepts CAN be used more than once.

Gender and technology studies

For Questions 64-77
____ 64. Approach such as "powder-puff" technology. A. Holistic

B. Equal Treatment
____ 65. Approach where girls are given extra help. C. Equal Output

D. Equal Access
____ 66. Approach where a teacher examines her or his practices

and disposes of projects that have questionable gender content.

Table 13. continued

____ 67. Approach where counselors remove barriers so girls

can take shop courses.

Multiple Choice: Choose the best answer. Circle the letter to the left of the option

Classroom management

97. When we say that teachers must model appropriate behavior and language, we mean that they ought to model

A) respect for students and guests in the workshop.
B) gender and racial equity.
C) skills without reinforcing gender roles.
D) ecological practice without appearing preachy.
E) All of the above

Writing good test items is a challenge, but if we have prepared the module, unit or course according to instructional design principles, the process of test construction can be quite smooth. If we have written an ample amount of objectives for the module, unit or course, stated them in assessable terms, and devised a test blueprint, the process of writing test items is straightforward. However, we have to attend to the techniques of item design. Guidelines and techniques for item construction are especially helpful (Sparzo, 1990). The following techniques for true-false, multiple choice, and matching items were provided by Davis and Spencer (2002) as shown in Tables 14, 15, and 16.

True-False Examples:

T	F	1. The Pelton water wheel is more efficient than the turbine.
T	F	2. A wheelbarrow is an example of a second-class lever.
T	F	3. A solar cell converts mechanical energy to electrical energy.
T	F	4. The steam engine is an external combustion engine.
T	F	5. A flashlight battery is a type of wet cell.

Table 14. True-false item techniques (Davis & Spencer, 2002)

Guideline	Explanation
• Use vocabulary that corresponds to the test material and is appropriate for the targeted age level.	True/False items are very straightforward items. These items serve to measure recognition and recollection of facts.
• Test one idea at a time. Divide compound ideas into two or more items.	
• Avoid specific determiners such as: *only, exactly, precisely, absolutely.*	These terms indicate circumstances that have no exceptions (these circumstances are very uncommon). Thus, the correct response given this pragmatic principle is to mark the item false.
• Avoid using terms that suggest an indefinite amount or degree, such as: *small, large, a long time ago, often, seldom, high, low, sometimes, usually, typically, and generally*.	These terms lead to challenges of their meaning which reduces the consistency of responses to the item. This lowers the validity of the test item.
• Avoid stating the test item in a negative sense by using *no* or *not.*	The negative phrasing complicates the logical structure of statements making the item unnecessarily difficult. If the statement cannot be worded positively, emphasize the negative terms by underlining or bolding.
• Use popular misconceptions as false statements.	
• Construct true-false items that require the use of introductory materials such as maps, graphs, or readings.	
• Have the students correct the false statements by changing them into true statements.	

Multiple choice example:

1. AutoCAD drawing files are typically saved with what type of file extension?

A. dwg

B. gif

C. mov

D. cad

E. jpeg

Table 15. Multiple choice item techniques (Davis & Spencer, 2002)

Guideline	Explanation
• The multiple choice stem needs to present one problem to the student before the options are considered.	A student should be able to formulate the answer before reading the options. This can be achieved by clearly indicating the topic of the test item using clear, simple, and direct language. The stem should be in the form of a specific incomplete statement or a direct question.
• Avoid writing stems that contain extra information as an introduction to the question. The stem should only include the question or the incomplete statement.	
• Restrict the use of negative terms in the stem.	Students may not notice the use of the negative term in the question. If the statement cannot be worded positively, emphasize the negative terms by underlining or bolding them.
• Construct stems that require the selection of the *best* answer when each of the options contains elements of correctness.	Multiple choice tests should measure a level of comprehension beyond pure memorization. Construct stems that require the student to use higher order thinking, not just a simple recall of the facts, to answer the item correctly. These questions are more difficult and discriminating than questions that ask for the recall of a single fact.
• Each distracter should have about the same number of words as the correct option.	One option that is longer than the rest often indicates the correct answer to the student.
• Make all distracters plausible.	
• Do not repeat wording or common elements from the stem in the correct option.	
• Make sure the stem is grammatically consistent with all of the options.	Students may reject options which are grammatically incorrect with the stem without truly knowing the content.
• Avoid using overlapping distracters.	
• Avoid the use of indefinite terms such as *usually* and *generally* in the options and distracters.	These terms indicate circumstances that have many exceptions (these circumstances are very common). Thus, the student may be cued into selecting the option with this term in it as the correct answer without truly knowing the content.
• Avoid the use of absolute terms such as *never* and *always* in the options and distracters.	These terms indicate circumstances that are without exceptions (these circumstances are very rare). Thus, the student may be cued into ruling out this option as the correct answer without truly knowing the content.

Table 15. Multiple choice item techniques (Davis & Spencer, 2002)

• Avoid using "all of the above."	Recognition of one wrong option eliminates "all of the above." Recognition of two right options identifies "all of the above" as the answer, even if the other options are completely unknown to the student.
• To increase item difficulty, include "none of the above" as a final option.	
• After the options are written, vary the location of the answer.	

Matching example:

_____1. *Match Parts of a DC Motor with Diagram (Below)*

_____2. A. Brush

B. Commutator

_____3. C. Power Supply

D. Armature

_____4. E. Field Magnet

_____5.

_____6.

Table 16. Matching item construction techniques (Davis & Spencer, 2002)

Guideline	Explanation
• Include specific, clear directions for the students.	
• Use only items that share the same foundation of information.	Unrelated topics included in the same matching item may allow for obvious matches and mismatches.
• Avoid using matching items that require sentence completion.	This technique provides the student with grammatical clues, which enable them to complete the sentence correctly without needing any knowledge of the topic.
• Write more responses than stimuli for each matching item.	This will help prevent the students answering the items by the process of elimination.

Table 16. continued

• The stimuli should be numbered and listed in a column on the left, while the responses should be lettered and laid out in a column on the right.	
• The column of stimuli on the left should set the question clearly.	
• The items for a matching exercise should be listed on one page.	This prevents unnecessary confusion created by flipping back and forth between pages.
• Limit the list of stimuli to fewer than 8 in order to keep the number of matching items brief.	

Grading, Marking and Reporting

Administrators, school boards, and teachers all have responsibilities to make sure that guidelines for student reporting are followed. Schools have very formal processes for reporting student progress. Most require three written reports (report cards) during the year, including one at the year's end, and at least two informal reports. Teachers are responsible for reporting for a number of reasons, including: (1) provincial or state legislation and policy for reporting on student progress; (2) accurate assessments for parents to comprehend their children's performance; (3) support of classroom learning; and (4) policy related to students with special needs. Additional reasons are commonly used for reporting at various grade levels. For instance, percentages and letter grades are required at the junior and senior levels and detailed literacy reports at the primary grades.

Match Parts of a DC Motor

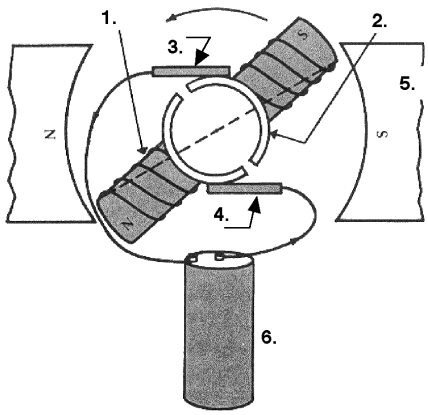


Table 17. Letter grade percentages and interpretations (Adapted from BC MOE, 1995b, p. 8)

A	B	C+	C	C-	F
86 - 100	73 - 85	67 - 72	60 - 66	50 - 59	0 - 49

A The student demonstrates excellent or outstanding performance in relation to the expected learning outcomes for the course or subject and grade.

B The student demonstrates very good performance in relation to the expected learning outcomes for the course or subject and grade.

C+ The student demonstrates good performance in relation to the expected learning outcomes for the course or subject and grade.

C The student demonstrates satisfactory performance in relation to the expected learning outcomes for the course or subject and grade.

C- The student demonstrates minimally acceptable performance in relation to the expected learning outcomes for the course or subject and grade.

F Failed or Failing. The student has not demonstrated, or is not demonstrating, minimally acceptable performance in relation to the expected learning outcomes for the course or subject and grade.

I In Progress or Incomplete. The student, for a variety of reasons, is not demonstrating minimally acceptable performance in relation to the expected learning outcomes.

Structured written reports to parents or guardians must be direct and use plain language. Written reports must follow the specific requirements for reporting student progress at each grade level. Comments in a student progress report describe, in relation to the curriculum:

- What the student is able to do.
- Areas of learning that require further attention or development.
- Ways the teacher is supporting the student's learning needs (and, where appropriate, ways the student or the parents might support the learning).

Each school year, teachers typically provide parents with a minimum of two informal reports. Informal reports usually describe the same things a formal report describes. However, the informal reports are commonly oral. They provide an important link between home and school and can be accomplished in a variety of ways, such as telephone calls, interim reports (written or oral) or conferences (parent-teacher, three-way, student-led, etc.). Most schools ensure that parents have the opportunity to meet with teachers for a conference at least once each school year. Teachers normally keep a record of each informal report noting the date of report, type of report and topics(s) of discussion.

In most school districts, letter grades are assigned to courses in grade 4 and higher. The successful completion of courses in grades 11 or 12 usually requires a minimum of a C-. Letter grades should be criterion-referenced throughout the students' courses in grades 4-12. In other words, teachers ought to create rubrics to clearly inform the students what the criteria are and mean for an A or B or etcetera. The numerical percentages and interpretations associated with letter grades are generally as follows in Table 17.

Summative assessments that are written in formal reports to parents or guardians must be, and usually are, based on a series of formative assessments. These formative and summative assessment data must be qualitatively and quantitatively analyzed with any number of means. For summative marks that must be in a quantitative format (percentages), formative assessments must be quantified throughout the term. Teachers assign a weight, indicating importance, to various performances (e.g., artifacts, cooperation, images, participation, portfolios, presentations, problem-solving, projects, quizzes, reports) to use for assessment. Some of these may account for as little as 5% or 10% of the final grade. Others, with a weight indicating importance, will account for as much as 40% or 50% of a grade. If the teacher is providing individual points or marks, then s/he may allot 10 marks for the minor performances and 35 marks for the major performances. The total marks allotted for all the performances in this case would tally to 100 marks. This is one way of dealing with weights. The assessment component of the teacher's outline for the course would look like this:

Module #1- 5 marks	Participation- 15 marks
Module #2- 5 marks	Project #1- 15 marks
Quiz #1- 10 marks	Project #2- 15 marks
Quiz #2- 10 marks	Portfolio- 25 marks
	Total= 100 marks

Here, a student who got 10 marks for the modules, 15 out of 20 for the quizzes, 13 for participation, 20 on the projects, and 20 for the portfolio would receive a summative or final mark of 78 or 78%. This is a B on our typical grade point and letter scale. Or, the teacher may choose to mark all assignments or performances on a 10 mark or 100 mark scale. The rubrics for this teacher would divide the levels of criteria into divisions of 10 or 100. The same teacher's assessment component on the course outline would look like this (weights are noted by percentages):

Module #1- 5%	Participation- 15%
Module #2- 5%	Project #1- 15%
Quiz #1- 10%	Project #2- 15%
Quiz #2- 10%	Portfolio- 25%
	Total= 100%

To calculate final marks, the teacher would have to write an equation or formula for a calculator or spreadsheet. The equation would be written as follows:

Total= (Item1 x .05) + (Item2 x .05) + (Item3 x .10) + (Item 4 x .10) + (Item5 x .15) + (Item6 x .15) + (Item7 x .15) + (Item8 x .25)

On the scale of 10 or 100 for each item, the same students above could have received 10 out of 10 for each of the modules, 9 out of 10 for each of the quizzes and for participation, and 8 of 10 (80%) for each the projects and the portfolio. The final mark would be 7.75 or 78%. Some psychometric experts suggest that in order to properly account for the variance within each of the assessment items, the individual items for each student should be transformed into z scores (normalized). Then, they say, the weights carried by each individual item will be more accurately calculated into the equation. This can easily be done with a spreadsheet but is overkill for most teaching situations.

Rubrics made the scoring of performances (e.g., artifacts, cooperation, images, participation, portfolios, presentations, problem-solving, projects, reports) easier for teachers. Marking, grading and scoring remains a very tedious process, nevertheless. Teachers who mark portfolios know all too well the numbers of hours spent reviewing and deliberating on entries within the individual portfolios. The marking of the artifacts from art or technology assignments can take days or weeks. The same goes for the marking of reports. In the late 1990s, software programs with artificial intelligence appeared on the market for marking essay questions and reports. Teachers who automate their quizzes and tests by placing them online, use courseware quiz scripts, or use Scantron answer forms that can be scanned and automatically corrected, benefit from the ease with which items can be marked and graded. Essay grading software is also readily available. There is great value in using computers for the testing process. New teachers should take every opportunity that they can to ease the grading process for themselves.

During the 1980s, a number of applications for grading, or automated grade books, appeared on the software market. By the late 1990s, many schools required their teachers to use a digital grade book adopted by their district. This made the submission and maintenance of student records much easier in many ways. On the other hand, it has forced teachers into somewhat uniform grading systems—they had to alter their grading practices to work with the digital grade book. The adoption of school-wide and district-wide digital grade books has also opened up the teachers' assessment practices to parents in ways never dreamed of thirty years ago. Parents and students now have access to browse the grade books via the web at anytime during the year.

Software grade book programs can range from glorified spreadsheets with simple functions to large databases with AI and plug-ins for a variety of other applications. The simplest grade books allow teachers to take care of the basics such as setting up classes, entering assignments, selecting grading methods used for grade calculation, and printing simple progress reports and reports. The more sophisticated programs offer the ability to create graphs of student progress, monitor attendance, record notes about student work, track parent contacts, print reports and link into school and district databases. Some allow for the use of a Personal Digital Assistant. Available digital grade books include Class Mates Grading Tools for Windows, 1st Class Gradebook, Grade Machine and Grade View, Gradebook 2, Gradechecker, Integrade Pro, MicroGrade, Parent Internet Viewer, Teachers Assistant Professional and ThinkWave Educator.

Grade Inflation

Veteran teachers who have been in the schools for thirty or forty years often note the changes in grading patterns over the course of their careers. "Courses were more rigorous," they note, and their grading practices were tougher thirty years ago. Everyone is easier on the students. Grading is one symptom in the larger pattern of softness. Work that would have received a "C" thirty years ago receives a "B+" or "A" in schools today, they note. By definition, grade inflation is the rise in the average marks of students over time. It is the skewing of the normal curve toward the upper end of the scale. It is the overly generous awarding of marks for under achievement. It results in the increase of student grade point averages (GPA) over time. Is grade inflation really an issue?

Researchers interested in grade inflation usually compare SAT test scores over time with GPAs over time. In recent studies, the records of 2.6 million students were examined. The researchers often compare affluent districts with poverty-stricken districts and schools. They consistently find that the advantaged schools rank higher in standardized test (ACT, SAT) scores and but, on average, have lower GPAs than the disadvantaged schools. The students in the disadvantaged schools get higher marks and grades than the advantaged students but score lower on the international and national tests. This irks some analysts. "Grade inflation is particularly extensive in high schools with a high percentage of disadvantaged students," M. Donald Thomas, an education advisor, reported to a national audience of school administrators in the U.S. He drew from this the conclusion that "this indicates clearly that expectations for students are very low, and standards do not match those of testing agencies." Does this indicate low standards? Is it fair to make these types of comparisons? Do we want disadvantaged students to *appear* to be low achievers and, in effect, dumber on all educational indicators? Could it be that students who score lower on the ACT and SAT tests, which deal only with English, math and science, actually excel in the subjects that are not tested on these tests (e.g., art, home economics, technology)?

How do we know if our students' grades are inflated next to our colleagues' students? How do we know if we are inflating our students' grades? Some analysts insist on the normal curve. They suggest that in any course, about 68% of students ought to get C's, about 14% ought to get D's and 14% ought to get B's and about 2% should get A's while another 2% should fail. Other analysts suggest that there is no magic bullet with which to curb grade inflation or keep it in check. And still others emphasize the importance of authentic assessment. If we authentically assess our students, they note, then grade inflation is not an issue. Grading, they insist, ought to inform instruction and actually help students improve their performance. Can we have high expectations and standards, and at the same time award the majority of our students with A's and B's?

Questionnaires and Scales of Technological Literacy

The Holy Grail for researchers in technology studies is a reliable, valid and standardized scale of technological literacy. In Chapter VII, we defined technological literacy to include action, cognition and emotion. The challenge is to create a scale that can be used nationally and internationally for comparative research and policy. In Chapter VIII we asked what *should* all students know about technology? Scales for research and policy ask what *do* students know about technology? Both questions are significant for researchers and teachers. While a number of scales of technological literacy have been constructed, none have been universally accepted. The difficulty of defining technological literacy, the changing nature of technology, and the lack of funds all contribute to the eventual success and obsolescence of a single, universal scale. With the ITEA's and ISTE's standards projects, a standardized scale of technological literacy is immanent. To date however, researchers have been more likely to use scales that measure attitudes than those that measure literacy (Hoepfl & Lindstrom, Forthcoming).

For example, in the mid 1980s, Marc deVries of Eindhoven University along with E. Allen Bame and William E. Dugger of Virginia Tech created a scale to measure grades 8-12 students' attitudes and values toward technology. This became known as the *Pupil's Attitudes Toward Technology (PATT) Scale*. The PATT Scale continues to be used and remains one of our most reliable tools for comparative measures of students' attitudes and values toward technology. The most common version of the PATT Scale consists of 100 Likert type items. The first 11 items are used to collect demographic information for the individual students. The remaining items deal with issues that force students to form an opinion (although there is a "neutral" option). Sample items include the following:

Agree	Tend to agree	Neutral	Tend to disagree	Disagree

12. When something new is discovered, I want to know more about it immediately.
18. I would like to know more about computers.
24. A girl can become an auto mechanic.
40. I think visiting a factory is boring.
43. To study technology you have to be talented.
69. With a technological job your future is promised.
76. In my opinion, technology is not very old.
97. Technology has little to do with daily life.

The PATT Scale is similar to scales of environmental values, militarism-pacifism, and the politics of technology. However, these latter scales are inherently more political. For instance, a popular scale on environmental values includes the following items:

Slightly agree	Agree	Strongly agree slightly disagree	Disagree	Strongly disagree

15. We're advancing so fast and are so out of control that we should just shut down and go back the way it was in colonial times.
23. People should pay the environmental costs of the things they buy. Products should be taxed depending on their effect on the environment.
37. We don't have to reduce our standard of living to solve global climate change or other environmental problems.

117. As new technologies become available that are less environmentally damaging, companies will naturally want to adopt and use them.

These scales are effective tools for social science research. But they are not just for university researchers. More and more teachers are involving themselves in research, mainly through action research programs. Action research is research that is directed at resolving immediate problems or policies that teachers face. Teachers with minimal capabilities in tests and measurements are feeling empowered to construct questionnaires and scales to investigate their students' attitudes and knowledge of technology. Some are adopting the validated scales such as PATT while others are constructing their own.

The key to constructing questionnaires and scales is to write items that provoke students to form an opinion and make a decision on something of value. Items that push students toward the extremes on a Likert scale are generally the best items. Some researchers recommend removing the neutral option in the middle. However, we must remain sensitive to cultural factors related to our students. Aboriginal peoples, for instance, tend to be reluctant to publicly express extreme opinions.

Likert type items are distinguished by the fact that they are assertions rather than questions. The respondent's task is to indicate the degree to which s/he agrees or disagrees with the assertion. Instead of having students explain their positions on issues, Likert scales force the students (or research participants) to respond to an issue or positions already formulated as an assertion. The use of Likert items allows for a quantification of responses and comparisons among groups. The key is to word the Likert items in very simple terms. Both negative and positive assertions are made to encourage respondents to deal with the content of the assertions rather than falling into an automatic response pattern. Typically, responses to the Likert items form "response sets" that allow teachers and researchers to assess where students stand on a range of issues. The response set helps researchers in determining that if respondents take a certain position with one issue they will take a related position on another issue.

Evaluation

The connotation of evaluation is that it involves inquiry that explores a characteristic, event, program or system in order to make a judgment on its merit or worthiness. However, there are actually four possible orientations to evaluation: (1) goal-attainment, (2) judgmental, (3) decision facilitation, and (4) illuminative. Goal-attainment evaluation is objectives-driven and the goal is to determine the extent to which intended outcomes are achieved. Judgmental evaluation means that a judgment is made on the value, or worth, of an endeavor or personnel based on external criteria. The judgmental orientation may focus on the professional judgment of the evaluator, as in a formal professional review system such as accreditation. Or a group of evaluators may judge the character of a leader by using prepared leadership evaluation scales. Decision facilitation evaluation typically means that evaluators do not personally assess merit or worth. Instead, they limit their role to gathering information for a decision-maker who will determine merit or worth. Illuminative or naturalistic evaluation is participant-oriented and focuses on the issues identified as important by stakeholders such as administrators, program staff, and students. The goal is to document the realities of individuals who experience the program first-hand. Adversarial approaches attempt to provide a balanced view by investigating different sides of issues, as represented by different participants. This goal is to generate opposing points of view within the overall evaluation process (Ruhe, 2003).

Neither assessment nor evaluation is limited to students. Student teachers are formally evaluated throughout their student-teaching practicum at least eight times and also go through a final or summative evaluation. New teachers go through a series of evaluations prior to receiving a form of tenure or a full-time appointment within their district. The evaluations may occur twice per year for the first three years of practice. The key to evaluation is preparation. Do not get caught off guard. Surprise evaluations are rare, so there is always time to prepare for an evaluation. Teacher evaluations are often high stakes, as they may determine career stability or salary raises. When you have time to prepare, you have time to prepare to look your best. Choose lessons and demonstrations for your evaluations that place attention on your strengths. Use lessons that you are familiar with and have practiced.

Most evaluations of teacher practices involve judgmental and decision-making orientations. Advisors, peer teachers, and administrators make the formal evaluations. Evaluation forms may be anecdotal checklists that resemble rubrics or open-ended to allow for free-flowing narrative. For instance, the anecdotal evaluation scale at the University of British Columbia (UBC) involves forty criteria by which student teachers are judged. Evaluation, including the evaluation of student teaching, is an extremely serious, and usually political, process. Similar to the authentic assessment of students, where they are given the criteria well ahead of the assessment, it is in your best interest to acquire the forms or scale with which you will be evaluated.

Like assessment, evaluation should be fair and should inform the process of improvement. Evaluation requires that deficits be candidly and clearly communicated in a constructive and timely fashion so that they can be eventually overcome. Administrators may have the upper hand in the evaluation of programs and personnel, but is it also common for teachers to evaluate their administration. In fact, the evaluation of leadership is an extremely active practice within the discipline of leadership studies (Wenig, 1995). Many leadership evaluation scales are designed to determine whether individuals have the "right stuff" to lead organizations. Other scales are similar to teacher evaluation instruments and allow for a deep analysis of the issues of organizational leadership. For example, one popular scale begins with the following items (SyberVision, 1993):

(-)		(+)	Score
1. Weak sense of purpose	1 2 3 4 5 6 7	Strong sense of purpose	____
2. Gives up easily	1 2 3 4 5 6 7	Very persistent	____
6. Unable to attract others	1 2 3 4 5 6 7	Magnetic, draws others	____
10. Self-ambitious; focused on own wants	1 2 3 4 5 6 7	Seeks to serve needs of others	____

As you can imagine, the evaluation of leadership is extremely sensitive. Subordinates often fear retaliation and opt for forms that allow for anonymous evaluation. In many cases, external teams are assembled to help mediate the process and provide an arm's length evaluation.

Courses, facilities, and programs are evaluated as well as personnel. Students in post-secondary institutions are quite familiar with course evaluations. Typically, the students submit each course and instructor to a process of evaluation. The most common evaluation scales for courses are forms with item "bubbles" to fill in with a dark pencil or pen. Course evaluation scales at most institutions are similar to each other. The sample items below from a popular scale will be familiar to post-secondary students and pre-service teachers.

0	1	2	3	4	5	6	7
Not applicable	Disagree very strongly		Disagree somewhat		Agree somewhat		Agree very strongly

0	1	2	3	4	5	6	7	
o	o	o	o	o	o	o	o	1. My interest in the course has increased because the way it was taught.
o	o	o	o	o	o	o	o	6. Course objectives were made clear.
o	o	o	o	o	o	o	o	14. This was an interesting course.
o	o	o	o	o	o	o	o	20. Course requirements were unclear.

There is also an opportunity for students to provide written comments. Both the scale and written comments are anonymous. Anonymity gives the students confidence to submit the evaluations without fear of retribution. In most countries, programs within post-secondary institutions are evaluated through a process of accreditation. Post-secondary technology programs are accountable to a range of governing bodies that require periodic evaluations.

Accreditation is the process whereby an association or agency recognizes an institution or program as having met certain qualifications or standards. It is generally a voluntary, non-governmental process of peer evaluation. This process holds institutions or programs accountable to certain, defined standards or criteria. Accreditation is often confused with certification: institutions and programs are accredited, and individuals are certified. Specialized or professional accreditors evaluate specific educational programs. For instance, there are specialized accreditation organizations for architecture, education, engineering, design, law, medicine, and the sciences. The Accreditation Board for Engineering and Technology (ABET), National Association of Schools of Art and Design (NASAD) and the National Council for Accreditation of Teacher Education (NCATE) are professional accrediting organizations that accredit programs in their respective disciplines. The ISTE, ITEA, and the Council for

Technology Teacher Education (CTTE) work with NCATE to accredit technology teacher education programs. The work of ABET, NASAD, and NCATE accounts for the accreditation of most post-secondary technology programs in North America.

Accreditation is a particular process of evaluation with specific evaluation items and criteria. This evaluation process requires to institutions to document the ways in which their programs meet standards regarding a range of items. For example, the ITEA/CTTE- NCATE process requires evidence for the ways that their students address items such as the following:

2.0 Possess the necessary depth and breadth in mathematics, science, and related disciplines to be able to successfully teach technology education.

3.0 Master teaching and technical skills appropriate to successfully teach the study of technology.

3.1 Possess knowledge about the development of technology, its effects on people, the environment and culture; and industry, its organization, personnel systems, techniques resources and products; and their impact on society and culture.

This process is supposed to influence provincial and state certification boards. In other words, the certification of technology teachers should be aligned with the accreditation requirements of the technology post-secondary programs. Theoretically, the provinces and states are supposed to hold teachers accountable who in turn would embody the contemporary standards and establish programs in the schools that are in line with state and accreditation standards. However, provincial and state certification programs often lag behind the accreditation process in the post-secondary institutions (Wiens, 1990).

Evaluation is an extremely important and ubiquitous process in business, education, and industry. In business and industry, facilities, personnel, and products are submitted to a regular routine of evaluation. In the previous chapter, we addressed the evaluation of curriculum materials, focusing on instructional design. The evaluation of facilities is covered in the next chapter.

Projection and Reflective Practice

In Chapter IX, we dealt with curriculum and instructional design. The decisions we make about "what should be learned?" and "how should it be organized for teaching?" are directly linked to assessment. In fact, authentic assessment originates with the ways that we answer to these problems of C&I. Most importantly, assessment is authentic when it informs the learning process. We began this chapter by setting up some rather problematic practices in technology studies. One example we used was the grading of artifacts rather than students—the grading of products over processes. We introduced three techniques of authentic assessment (portfolio, performance, and criterion-referenced assessment) to help us contradict the problematic practices. Portfolios and rubrics, which are used to cross reference criteria of performances, are extremely applicable to technology studies. Testing and measurements are redefined in the context of authentic assessment to serve the processes of learning and instruction. Designing effective tests is a challenging endeavor and guidelines for test design were provided. The processes of grading and reporting are also challenging. New, automated, and online modes of grading and maintaining records offer essential techniques for teachers and allow for the manipulation of large databases of information. We differentiated between assessment and evaluation by noting that assessment was associated with student progress and evaluation was associated with judgments on personnel, courses and programs. Accreditation and credentialing are particular forms of evaluation. In the next chapter, we will deal with the nuances of classroom management, facilities design, and safety. We will also address the challenges of special needs students and follow-up with specialized assessment and evaluation techniques.

References

Bame, E. A., Dugger, W., de Vries, M., & McBee, J. (1993). Pupils' attitudes toward technology—PATT-USA. *Journal of Epsilon Pi Tau, 12*(1), 40-48.

British Columbia Ministry of Education. (1994a). *Performance assessment.* Victoria: Author.

British Columbia Ministry of Education. (1994b). *Portfolio assessment.* Victoria: Author.

British Columbia Ministry of Education. (1995a). *Evaluating problem-solving across the curriculum.* Victoria: Author.

British Columbia Ministry of Education. (1995b). *Guidelines for student reporting for the kindergarten to grade 12 education plan*. Victoria: BC.

British Columbia Ministry of Education. (1996). *Information technology 11 and 12.* Victoria: Author.

Custer, R. L. (1996). Rubrics. *The Technology Teacher, 55*(4), 27-37.

Custer, R. L. (2000). *Using authentic assessment in vocational education.* Columbus, OH: ERIC Clearinghouse on Adult, Career, and Vocational Education.

Davis, T., & Spencer, K. (2002). *Test item construction*. Retrieved July 5, 2004, from http://www.empowermentzone.tamu.edu

Hoepfl, M. (1993). Portfolio assessment. *The Technology Teacher, 53*(2), 28-29.

Hoepfl, M., & Lindstrom, M. (Forthcoming). *Assessment of technology education*. New York: Glencoe.

Kerka, S. (1995). *Techniques for authentic assessment: Practice application brief*. ERIC. Retrieved July 5, 2004, from http://eric.acve.org/pubs.html

Kimbell, R. (1997). *Assessing technology: International trends in curriculum and assessment*. Buckingham: Open University Press.

Kirchoff, J. (1996). Character traits rubric. In R. Custer (Ed.), *Rubrics: The technology teacher, 55*(4), 27-37.

Romani, D. (2002). *Employability skills rubric.* Coquitlam, BC: Author.

Ruhe, V. (2003). *Applying Messick's framework to the evaluation data of distance/distributed instructional programs*. Unpublished PhD dissertation. University of British Columbia).

Sanders, M. (2000). Web-based portfolios for technology education: A case study. *Journal of Technology Studies*, *26*(1), 11-20.

Sparzo, Frank J. (1990). *Preparing better teacher-made tests: A practical guide*. *Fastback 311*. Bloomington, IN: Phi Delta Kappa Educational Foundation.

Starmer, G. H. (1974). An assessment scheme for technology projects. In *School technology in action* (pp. 80-89). London: Open University Press.

SyberVision. (1993). *Leadership characteristic personal assessment form*. Washington, DC: Author.

Wiens, E. (1990). CTTE/ITEA NCATE. *Journal of Technology Education, 2*(1), 56-60.

Wenig, R. (1995). Leadership. In G. E. Martin (Ed.), *Foundations of technology education, 44th yearbook of the council on technology teacher education*. New York: Glencoe.