

# **The Development of Standards for Minnesota Schools**

Richard Beach

Professor of English Education, Department of Curriculum and Instruction  
University of Minnesota

## **Summary**

The key difference between the Profile of Learning and the curriculum approach proposed by Governor Pawlenty and Commissioner Yecke can be seen by looking at standards statements, for example those related to high school literature instruction.

- The primary focus of the Profile standards is on students' demonstrating the ability to comprehend, interpret, and evaluate literary texts.
- In contrast, the Virginia SOL content-based standards focus on acquiring knowledge about literary history and authors.

## **Limitations of Content-Based Standards**

- They encourage teachers to adopt a transmission or lecture model of instruction that does not actively engage students in higher-order thinking.
- They represent an outmoded curriculum focus on coverage of information and concepts without providing students with the frameworks and problem-solving strategies necessary for success in an information-age economy.
- They represent a top-down standardization of curriculum content in which teachers are expected to cover the same content in order to prepare students for multiple-choice tests.

## **The Use of High-Stakes Multiple-Choice Testing**

- In a content-based standards approach, teachers teach for content knowledge recall consistent with material covered on multiple-choice tests.
- Multiple-choice test results provide a false sense of objectivity that does not measure the complexity or processes of learning, resulting in the trivialization of learning and teaching.
- Schools are labeled as failures based on test scores, misleading labels that are used to justify public policy such as the use of vouchers or privatization of public education.
- Multiple-choice tests measure a low level of thinking, as opposed to the ability to synthesize, evaluate, and critique information and ideas.
- Tests may be improperly used to determine teaching quality, uses that presuppose a

simplistic cause-effect relationship between instruction and test scores, which, as a gross measure of learning, may be influenced by a range of different factors.

- Development of a new battery of multiple-choice tests will be very costly, costs difficult to justify when the state faces massive budget deficits.

**Framing the Issue as an Either/Or Choice** erroneously assumes that nothing in the Profile, as developed over a ten-year period, is worth preserving. Some of the successful aspects of the Profile could be allowed to continue or revised at a fraction of the cost of developing a whole new system.

### **Positive Influences of the Profile on Student Learning**

Research evidence suggests that the Profile was serving to foster higher-order thinking and requiring students to take responsibility for their own learning.

### **Summary of Recommendations**

- Reject the “either/or” critique, in which everything developed over a ten-year period under the Profile is rejected, by requiring that the successful programs developed by some districts under the Profile be included in the new curriculum standards. Allow some districts and/or schools who have already demonstrated high levels of success under the Profile to continue to do so.
- Reject attempts to import off-the-shelf standards such as the Virginia standards that reflect an exclusively content-based focus, as well as outmoded, traditional curriculum inconsistent with the high quality of schooling in Minnesota. Stipulate that any standards developed for Minnesota need to be formulated by and for citizens and educators of Minnesota.
- Challenge the notion that any new standards can be developed and finalized by a group of subcommittees by March 31. This high-speed timetable deadline can only be achieved by importing of existing content-based standards from other states (probably Virginia) that may have little or no bearing on Minnesota schools and, for the reasons cited above, will turn back the clock on Minnesota’s education.
- Stipulate that any new set of standards developed for Minnesota need to include both content and higher-order thinking standards, as opposed to only content-based standards. In this balanced approach, some of the successful standards that were developed for the Profile could still be used, as well as revisions of these standards. New higher-order thinking standards could be introduced as consistent with new curriculum developments in the past few years.
- Provide teachers with assistance in formulating curriculum. One of the reasons that some teachers had difficulty implementing the Profile was that they never had a curriculum

framework or roadmap for assisting them in implementing the Profile. The CFL curriculum frameworks developed over the past two years for many subject matter areas could serve as the basis for assisting teachers.

- Recognize that funds for the purchase and administration of new tests (on top of the development of the new NCLB mandated tests for grades 4, 5, and 7) will divert funds away from the general education funding of schools, which will themselves be facing further cuts in teachers, resulting in increased class.
- Explore options to multiple-choice testing as the sole measure of student learning. One alternative to the narrow focus of multiple-choice tests is the use of state-wide performance-based assessments. The state must test students in grades 1-8 as mandated by the NCLB, as well as continue to administer the basic skills reading, math, and writing tests. Any other new subject matter tests for the upper grades should be phased in over time and developed as performance assessment measures consistent with teachers' classroom instruction.

# The Development of Standards for Minnesota Schools

Richard Beach

Professor of English Education, Department of Curriculum and Instruction  
University of Minnesota

## Introduction

Recent news accounts have reported that Governor Pawlenty and Commissioner of Education Yecke intend to recommend that the current state education standards, the Profile of Learning, be replaced by a new set of content-based standards.

As reported in the *StarTribune* (January 17, 2003):

It's clear that a Yecke priority is to jettison the Profile of Learning graduation rule and to replace it with academic standards that focus more on content -- that is, on specific things students must know. That would shift the focus on learning more toward pure knowledge and away from the Profile's emphasis on more active ways to learn.

And, the *Pioneer Press* (January 17, 2003):

Minnesota's Profile of Learning emphasizes students demonstrating what they have learned through reports and projects. Virginia's Standards of Learning list specific facts and skills that students should know for each grade, with schools measured on their performance based on student testing. Supporters of such a standards system say it is clear and specific so that students, parents and teachers know what is being taught and what is expected. Critics say it puts too much of an emphasis on rote memorization and does little to foster critical thinking skills.

The Profile was originally conceived of as a response to the content-based/basic skills approach of the 1970's/80's that was perceived by the framers of the Profile as not preparing students to engage in the kind of problem-solving thinking necessary for participation in an information-age economy. Research by the National Assessment for Educational Progress (NAEP) found a decline in higher-order thinking skills evident on their open-ended tests during the decade of the 1970's due to an excessive focus on a "basic-skills" curriculum during that period.

## Differences between the Profile and the Proposed Curriculum Model

The difference between the two approaches is evident in standards statements related to high school literature instruction. The Profile standards for that area are as follows:

### **Profile of Learning Standard: Reading, Listening, and Viewing Complex Information Grades: High Grades 9-12**

A student shall demonstrate the ability to comprehend and evaluate complex information in varied nonfiction by reading, listening, and viewing varied English language selections containing complex information and, in these selections:

- A. identify main ideas and supporting information;
- B. distinguish fact from opinion, fiction from nonfiction, or both;
- C. identify bias, point of view, and author's intent;
- D. identify relevant background information; and
- E. analyze and evaluate the credibility of evidence and source, the logic of reasoning, and how the type of communication shapes or limits information.

### **Profile of Learning Standard: Literature and Arts Analysis and Interpretation - Literature Grades: High Grades 9-12**

A student shall demonstrate the ability to interpret and evaluate complex works of literature by:

- A. describing the elements and structure of literature; the artistic intent; and the historical, cultural, and social background of the selected literature;
- B. applying specific critical criteria to interpret and analyze the selected literature;
- C. describing how particular effects are produced by the artist's use of the elements of literature; and
- D. communicating an informed interpretation using the vocabulary of literature.

The primary focus of these Profile standards is on students demonstrating the ability to comprehend, interpret, and evaluate literary texts. In contrast, the Virginia SOL standards in high school literature are as follows:

### **Virginia SOL Reading/Literature Standards**

**11.3** The student will read and analyze relationships among American literature, history, and culture.

- \* Describe contributions of different cultures to the development of American literature.
- \* Describe the development of American literature in the 17th, 18th, 19th, and 20th centuries.
- \* Contrast periods in American literature.
- \* Differentiate among archetypal characters in American literature.
- \* Describe the major themes in American literature.
- \* Describe how use of context and language structures conveys an author's point of view in contemporary and historical essays, speeches, and critical reviews.

**12.3** The student will analyze the development of British literature and literature of other cultures.

- \* Recognize major literary forms and techniques.
- \* Recognize the characteristics of major chronological eras.
- \* Relate literary works and authors to major themes and issues of their eras.

In contrast to acquiring knowledge about literary history and authors, the Profile standards focus more on the ability to engage in interpretation of literature itself as opposed to recall of

knowledge about literature. Similarly, the Profile standard for high school drama instruction is as follows:

**Profile of Learning Standard: Literature and Arts Analysis and Interpretation - Theater  
Grades: High Grades 9-12**

A student shall demonstrate the ability to interpret and evaluate complex works of theater by:

- A. describing the elements and structure of theater; the artistic intent; and the historical, cultural, and social background of the selected theatrical works;
- B. applying specific critical criteria to interpret and analyze the selected theatrical works;
- C. describing how particular effects are produced by the artist's use of the elements of theater; and
- D. communicating an informed interpretation using the vocabulary of theater.

The focus here is again on demonstrating understanding of drama through interpretation and evaluation. In contrast, the SOL standards for drama focus more on knowledge about formalist features of drama:

**Virginia SOL Drama Standards**

**11.6** The student will read a variety of dramatic selections.

- \* Describe the relationship between farce and characterization.
- \* Describe the dramatic conventions or devices used by playwrights to present selected plays.
- \* Explain the use of monologue and soliloquy.
- \* Explain the use of verbal and dramatic irony.

**12.6** The student will read and critique dramatic selections from a variety of authors.

- \* Describe the conflict, plot, climax, and setting.
- \* Compare and contrast ways in which dialogue and staging contribute to the theme.
- \* Identify the most effective elements of selected plays.
- \* Compare and contrast dramatic elements of plays from American, British, and other cultures.

All of this suggests that a primary difference in these approaches is that the SOL specifies the curriculum content that needs to be covered in all Virginia schools. This standards approach that focuses on specific content differs from standards articulated by national professional organizations such as the National Council of Teachers of English and the International Reading Association. The NCTE and IRA standards assume that standards function to foster and guide teachers' development of their own curriculum and instruction, as opposed to prescribing specific content to be covered. As they note;

These standards assume that literacy growth begins before children enter school as they experience and experiment with literacy activities—reading and writing, and associating spoken words with their graphic representations. Recognizing this fact, these standards encourage the development of curriculum and instruction that make productive use of the emerging literacy abilities that children bring to school.

Furthermore, the standards provide ample room for the innovation and creativity essential to teaching and learning. They are not prescriptions for particular curriculum or instruction. Although we present these standards as a list, we want to emphasize that they are not distinct and separable; they are, in fact, interrelated and should be considered as a whole.

The three NCTE/IRA standards related to literature instruction, in contrast to the SOL standards, do not specify certain content, but rather, adopt a more balanced approach, focusing on the need

to provide students with not only certain kinds of desired literary experiences, but also ways of interpreting literary texts:

- Students read a wide range of print and non-print texts to build an understanding of texts, of themselves, and of the cultures of the United States and the world; to acquire new information; to respond to the needs and demands of society and the workplace; and for personal fulfillment. Among these texts are fiction and nonfiction, classic and contemporary works.
- Students read a wide range of literature from many periods in many genres to build an understanding of the many dimensions (e.g., philosophical, ethical, aesthetic) of human experience.
- Students apply a wide range of strategies to comprehend, interpret, evaluate, and appreciate texts. They draw on their prior experience, their interactions with other readers and writers, their knowledge of word meaning and of other texts, their word identification strategies, and their understanding of textual features (e.g., sound-letter correspondence, sentence structure, context, graphics).

Other states have successfully demonstrated that performance assessment and portfolios of student classroom work can be used to provide teachers and parents with valid profiles of student progress at a fraction of the cost paid to testing corporations. The state of Kentucky has a strong track record of using performance-based assessments to improve student learning and teaching. One advantage of the Kentucky portfolio-assessment model is that the teachers' instruction is directly aligned with how students are assessed. The Kentucky program should serve as a model for devising assessments that are consistent with higher-order standards. Other states such as Connecticut have developed effective performance assessment standards.

<<http://www.state.ct.us/sde/dtl/curriculum/currkey3.htm>>

The New Standards Project operated by National Center on Education and the Economy in Washington, D.C., and the Learning Research and Development Center at the University of Pittsburgh has developed a set of performance assessment measures that have served as an effective tool for evaluating student learning. The New Standards Project combines both content and higher-order thinking standards. The Project believes that if teachers are to be held accountable for their students' performance, then they should have some input on how that performance is assessed. In some cases, traditional tests could be used over a 3-4 year period as a baseline measure of change, but then use the tests only to determine overall school or state performance.

The New Standards formulate specific tasks and assessment criteria for evaluation of task performance. For example, the following are two tasks related to 10<sup>th</sup> grade reading:

The student reads and comprehends at least four books (or book equivalents) about one issue or subject, or four books by a single writer, or four books in one genre, and produces evidence of reading that:

- Makes and supports warranted and responsible assertions about the texts;
- Supports assertions with elaborated and convincing evidence;
- Draws the texts together to compare and contrast themes, characters, and ideas;
- Makes perceptive and well developed connections;
- Evaluates writing strategies and elements of the author's craft.

Examples of activities through which students might produce evidence of reading comprehension include:

*A saturation report (a report that recounts substantial information on a topic gathered by a student over a*

*period of time)*

*Construct a review of several works by a single author*

*Produce a literary response paper*

*Produce a research report*

*Participate in formal or informal book talks*

*Create an annotated book list organized according to author, theme, or genre*

The student reads and comprehends informational materials to develop understanding and expertise and produces written or oral work that:

Restates or summarizes information

Relates new information to prior knowledge and experience

Extends ideas

Makes connections to related topics or information

Examples of activities through which students might produce evidence of reading informational materials include:

*Use information to support or enhance a project*

*Write a report of information that draws from at least four sources*

*Incorporate expert opinions into a speech or position paper*

*Develop a proposal based on data obtained from reading informational texts*

*Use informational materials to reach a conclusion regarding a controversial topic*

*Develop a portfolio of materials regarding a particular career choice*

*Write exhibit notes for historical or artistic exhibits.*

In contrast to the content-based SOL standards previously noted, these standards specify important comprehension strategies necessary for successful reading. For state-wide assessments, evidence for students' demonstration of these strategies could be readily collected and scored by teams of trained judges.

State-wide performance assessments are typically questioned as lacking reliability and generalizability and are therefore perceived as not useful for determining accountability. However, results of the Vermont and Kentucky portfolio scoring reliability indicate that decent reliabilities can be achieved. The limitations of performance assessment need to be weighed against the positive influences of assessment on curriculum and instruction in serving, as Hillocks's research indicates, to promote higher-order thinking and literacy development, and to influence instruction in a positive manner.

The School Redesign Network at Stanford University has sponsored examples of consortiums of schools who successfully use standards-based performance assessment tools <http://www.stanford.edu/dept/SUSE/csrn/features/standards/>. Examples of other performance assessment programs and rubrics can be found at the following site: [http://score.rims.k12.ca.us/sub\\_standards/alt\\_assessment\\_res\\_on\\_web.html](http://score.rims.k12.ca.us/sub_standards/alt_assessment_res_on_web.html)

## **Limitations of Content-Based Standards**

There are a number of limitations to organizing the curriculum solely on the basis of content-based standards.

***“Transmission” versus “hands-on” models of instruction.*** One of the major limitations of content-based standards are that they encourage teachers to adopt a transmission model of

instruction in which the teacher's role is to simply provide students with information or concepts and then test students on whether they have acquired that information or particular concepts.

Commissioner Yecke argues that the Profile dictated how teachers should teach, with the implication that content-based standards do not dictate how teachers should teach. However, if it is the case that certain content is going to be covered on the tests, teachers will have to adopt more of a transmission/ lecture approach to providing students with the relevant information and concepts, because they are rightfully concerned about their students' performance on content-based multiple-choice tests. What is problematic about this transmission model is that it does not provide students with the frameworks or strategies for them to learn how to apply or use information or concepts to address real-world, complex problems. While facts and information are readily available on-line in a digital world, people need to know how to use facts and information to address everyday, complex problems and issues.

One advantage of the hands-on approach involved in completing projects associated with the Profile is that through participating in these projects, students learned how to pose questions, formulate problems, collect information, and then synthesize and analyze that information to address these questions or problems. It is these question-asking/problem-solving practices that are essential for participating in a 21<sup>st</sup> century economy in which people are constructing new knowledge rather than simply regurgitating information. As John Dewey noted, "learning is a process, not a static object; it must grow out of reflection on experience, not simply be imported from some pre-existing structure."

***Imposition of an outmoded curriculum.*** As an English educator, it is my own judgment and the judgment of other literacy educators with whom I have spoken that the Virginia standards for English represent an outmoded curriculum. The Virginia standards imply the need to teach traditional American and British literature survey courses, the traditional expository essay, and grammar, all components of an English language arts curriculum that was prevalent some forty years ago. Since that time, the English curriculum has changed markedly, focusing much more on critical response to literature, multicultural literature, composing processes, writing for social purposes, inquiry/thematic approaches, studying language use in context, digital/media literacies, and a redefinition of middle-school language arts. The fact that students devote much of their non-school time to acquiring a range of digital/media literacies is very much marginalized in the Virginia curriculum.

A draft of a new K-12 language arts framework has been formulated by a CFL committee over the past two years, a committee on which I served along with teachers from throughout the state. This framework is designed to help language arts teachers implement the Profile. It focuses on teaching specific strategies involved in understanding and producing texts: engaging, activating prior knowledge, decoding/encoding, interpreting, describing/organizing, connecting, developing vocabulary, monitoring, and judging. These strategies serve as a clearly defined, rigorous focus for teaching language arts.

Curriculum frameworks were also developed for a range of other subject-matter areas. Unfortunately, because these frameworks were linked to the Profile standards, they may be

perceived as inconsistent with a content-based approach, and two years of the CFL committee work will be ignored.

***Top-down standardization of curriculum content.*** A content-based standards curriculum represents a marked shift in the role of state government in dictating specific content that is unprecedented in Minnesota's history. In the past, the State Department of Education played a minimal role in specifying a state-wide curriculum. It provided only general curriculum guidelines, allowing local districts to define their curriculum content and teachers to formulate their curriculum to suit the needs, interests, and abilities of their students. The proposed content-based standards approach will prescribe specific content to be covered by all teachers in Minnesota. Given a high-stakes accountability approach, failure to cover that content may result in not only students, but also their schools being labeled as "failures." (In Virginia, schools that fail to achieve a 70% pass rate over a certain number of years receive sanctions.) This top-down imposition of specific content affords teachers, administrators, and districts little or no autonomy in terms of deciding on their curriculum content.

During this past summer, I worked with Virginia social studies and English teachers in a two-day workshop in Roanoke. These teachers consistently expressed resentment about the top-down nature of the SOL and the fact that it left them with little opportunity to develop their own curriculum and instruction. They also noted that they were unable to use a lot of current curriculum strategies because these strategies were inconsistent with the traditional SOL social studies/English curriculum standards. Virginia teachers were also critical of attempts to impose an external assessment system on their teaching.

All of this suggests that if these teachers were concerned about the imposition of the Profile, they will be equally, if not more concerned about the imposition of specific content curriculum. While it could be argued that teachers should certainly be expected to teach certain content consistent with a school or district curriculum, it is usually the case that the teachers themselves have formulated that school's or district's curriculum. However, on a state-wide level, teachers may have little or no opportunity to influence decisions regarding a uniform state-wide curriculum. This serves to undermine teachers' sense of their own professional autonomy and academic freedom to teach certain content, particularly if that content is perceived to be inconsistent with "what's on the test."

### **The Use Of High-Stakes Multiple-Choice Testing**

A second major problem with a content-based model is the potential for exclusive reliance on multiple-choice testing as the sole measure of student learning. This use of high-stakes testing places heavy emphasis on objective measures such as test scores that are then used to determine the success or failure of schools in an impersonal mode that avoids the complexity and values operating in those schools. As Professor Norman Fairclough of the University of Lancaster noted, "so-called 'objective' standards of assessment are abstractions, a very long way away from lived reality."

There are a number of problems with the exclusive use of multiple-choice testing.

***Testing and teaching for only content knowledge recall and application.*** The proposed content-based standards, if they are at all consistent with the Virginia model, will be driven heavily by multiple-choice tests of students' knowledge of the content. For example, the SOL Mathematics test consists of 60 to 70 single-answer, multiple-choice questions. These multiple-choice questions cover the subject areas of number & number sense, computation & estimation, measurement & geometry, probability & statistics, and patterns, functions & algebra. The Reading assessment consists of 35 to 55 single-answer, multiple choice questions covering the following subject areas: using word analysis strategies, understanding a variety of printed materials & resources, and understanding elements of literature. The Writing assessment consists of a direct writing component and 30-35 multiple-choice questions.

Because these are high-stakes tests, teachers invariably teach to the test. A study of actual math teacher practices in response to high-stakes testing (Firestone & Mayrowetz, 2000) found that 88% of teachers in Maryland believed that they were under "undue pressure" to improve test scores. Seventy-five % of the Maryland math teachers reported making changes in the content taught in order to teach to the test; teachers also devoted more instructional time to helping students learn to take multiple-choice tests.

Teaching to the test does not necessarily have negative consequences when the tests represent the same rich, robust learning involved in the classroom. In a study of writing instruction in five states, George Hillocks (2002) found that the four states employing a traditional writing test assessment similar to that employed in both Virginia and Minnesota, teachers were teaching the five-paragraph format, with little attention to the composing/thinking processes, audience analysis, inquiry-strategies, or writing across the curriculum. In contrast, in Kentucky, which employs a portfolio writing assessment, teachers were focusing more on teaching process writing, rhetorical strategies, revision skills, and writing in different contexts. Hillocks's research demonstrates that the type of assessment employed determines the focus and type of instruction employed.

Employing more performance-based assessments may actually have positive influences on instruction by encouraging uses of the higher-order thinking involved in such assessments. The proposed use of multiple-choice tests as the driving engine behind the content-based standards will limit teachers to reviewing subject-matter content and test-taking skills.

***Labeling schools as "failing" based on test scores.*** One of the basic assumptions behind the uses of high-stakes tests is the need to enforce sanctions on "failing schools" and to no longer let students fall through the cracks. The further assumption is that by adopting high-stakes testing, that schools will improve, i.e., that more learning will occur. If states' high-stakes testing actually improves student learning, then one would expect that test scores on other national test would also be increasing. However, a recent University of Arizona study of the impact of high-stakes testing in 18 states found no increase in SAT, ACT, NAEP, or AP test scores over a four-year period (Amrein & Berliner, 2002).

Analysis of the Virginia students' tests found that while there was no basis for comparison with NAEP reading and math scores, there was no change in NAEP science scores; on the Stanford 9

scores, while there were increases in 4<sup>th</sup> grade scores, there was no increase in upper grades' scores. There was no increase in the percentage of students taking the SAT tests; there was a decline in the percentage taking the SAT II tests. The study also reports that the percentage of Virginia students taking AP tests rose slightly in 2002 but at a lower rate than nationally. For Virginia public school students, the percentage of students scoring 3 or better on all AP tests has declined from 65% in 1997 to 62% in 2002. In particular on AP Biology, English Composition, English Literature, and U.S. History, the percentage of public school students scoring 3 or better has declined since 1997. And, the percentages of students receiving diplomas declined from 75% in 1997 to 71.8% in 2002. (For further information on this data see: <http://www.SOLreform.com/achievement.htm>)

In 1998, the first year of the implementation of the Virginia testing, only 39 of the state's 1,800 schools, or 2.2 percent, met the performance goals the first time out. Another 128 would have met the requirements with better performance on just one test (Olson, *Education Weekly*, Feb. 3, 1999). The percentage of students who passed Virginia's tests on the first go-around ranged from a low of 32.8 percent in 5th grade history to a high of 71.1 percent in 8th grade science. The average passing rate across all the tests was 41 percent.

One reaction to the potential sanctions of being labeled as a "failing school" is to put pressure on state officials to lower the cut scores for failure. In Virginia, given the fact that the pass rates had not been changing, the Department of Education lowered the cut scores for many of the history and social studies' tests.

Further analysis of the pass rates in Virginia schools finds that scores were "adjusted" depending on whether special education/special population students are included or excluded from the total scores ("How Virginia Plays the Accreditation Numbers Games <http://www.SOLreform.com/hardapples.htm>)

There are a number of problems with this inevitable accounting fix of shifting the cut scores. It creates a high level of cynicism regarding the entire process of labeling schools as failures based on shifting, inconsistent, unreliable standards. It also penalizes students who may have scored lower than the previous cut score, resulting in being labeled as "not proficient," but who would now possibly be judged as "proficient" simply because they took the tests with a higher cut score.

*Cautions from test experts.* Test experts argue strongly against the use of aggregated test scores to label schools. They also caution against the use of test scores to label students. A Rand Corporation analysis of then-Governor Bush's so-called "Texas Miracle" revealed that state test scores rose, not because of increased learning, but because of increased drop-out rates among students who were already struggling in school. <<http://www.rand.org/publications/IP/IP202/>>. In other words, the students who most needed better instruction were given high stakes tests instead, and they dropped out. Test scores are often used to label students as failures, undermining their motivation to want to improve, so they give up on the system. The American Educational Research Association (AERA) position statement on this matter is as follows:

Reporting of test results may also be beneficial in directing public attention to gross achievement disparities among schools or among student groups. However, if high-stakes testing programs are implemented in

circumstances where educational resources are inadequate or where tests lack sufficient reliability and validity for their intended purposes, there is potential for serious harm. Policy makers and the public may be misled by spurious test score increases unrelated to any fundamental educational improvement; students may be placed at increased risk of educational failure and dropping out; teachers may be blamed or punished for inequitable resources over which they have no control; and curriculum and instruction may be severely distorted if high test scores per se, rather than learning, become the overriding goal of classroom instruction.

Decisions that affect individual students' life chances or educational opportunities should not be made on the basis of test scores alone. Other relevant information should be taken into account to enhance the overall validity of such decisions. As a minimum assurance of fairness, when tests are used as part of making high-stakes decisions for individual students such as promotion to the next grade or high school graduation, students must be afforded multiple opportunities to pass the test. More importantly, when there is credible evidence that a test score may not adequately reflect a student's true proficiency, alternative acceptable means should be provided by which to demonstrate attainment of the tested standards.

Moreover, AERA statement raises concern about the alignment of the curriculum to the test, noting that students should not be tested on material that has not been taught:

When content standards and associated tests are introduced as a reform to change and thereby improve current practice, opportunities to access appropriate materials and retraining consistent with the intended changes should be provided before schools, teachers, or students are sanctioned for failing to meet the new standards. In particular, when testing is used for individual student accountability or certification, students must have had a meaningful opportunity to learn the tested content and cognitive processes. Thus, it must be shown that the tested content has been incorporated into the curriculum, materials, and instruction students are provided before high-stakes consequences are imposed for failing examination.

Labeling schools as failures needs to be placed within the larger conservative agenda related to public education—the promotion of school vouchers and the privatization of public education. Labeling schools as failing serves to encourage parents and schools to perceive vouchers, particularly for private school enrollment, as a viable option. As noted by Norman Draper in the *StarTribune* (December 15, 2002):

Some even wonder if there's a hidden agenda to discredit public schools and pave the way for privatizing the public school system. Even educators who stop short of making such predictions project a high failure rate among the state's schools...

Two years of substandard test scores places a school on the "needs improvement" list. The school has to make that fact public, and the state has to submit a list of such schools to the U.S. Department of Education. If the school remains on the list another year, it must offer students to the option of attending other, better performing schools, and has to set aside some of its federal funds to bus such students to their schools of choice. A school on the list for a third year has to offer additional educational service to its students, and must continue to offer the option for students to attend other schools.

And, labeling schools as failures—often because they contain a high percentage of low-income students—undermines the larger sense of equity in society. As Paul Wellstone noted:

Making students accountable for test scores works well on a bumper sticker and it allows many politicians to look good by saying that they will not tolerate failure. But it represents a hollow promise. Far from improving education, high-stakes testing marks a major retreat from fairness, from accuracy, from quality, and from equity.

***The low level of thinking measured by multiple-choice tests.*** Multiple-choice items do not serve as valid measures of the ability to engage in critical or higher-order thinking, for example, the ability to synthesize, integrate, or evaluate information. In response to the question regarding the Virginia SOL tests on their Web page: "Do the SOL tests measure higher order or complex behavior?" the answer notes that synthesis and evaluation have been eliminated because they cannot be measured:

Most teachers are familiar with the cognitive taxonomy attributed to Benjamin Bloom. The taxonomy includes six levels of behavior; knowledge, comprehension, application, analysis, synthesis, and evaluation. The new SOL include a simpler hierarchy across the content areas and include four levels: knowledge, comprehension, application, and reasoning.

Most teachers will want their students to be learning at all levels of the SOL hierarchy; however, the state tests will seldom go beyond application. This is because test items that measure reasoning and beyond do not lend themselves to the limitations imposed by most standardized tests and the multiple choice format.

In arguing that testing actually lowers standards, Donald Graves (2002), Professor Emeritus at the University of New Hampshire argues that testing does not require "long thinking:"

Currently, we are testing what we value, quick thinking. But what about long thinking? Can we discern thinkers like Thomas Jefferson, Albert Einstein, and Charles Darwin, who were self-professed long, slow thinkers. Can we identify and encourage the children who can formulate a question, find the information, structure an evaluation design, and know if they have answered their original question? The problems of a democracy are not solved through single answers but by tough-minded thinkers who sustain thought on one problem for days, months, or years.

The massive amounts of time spent either preparing for tests or taking them have displaced writing and original, long thinking, and have dulled the thinking edges of our students. Indeed, current approaches to assessment are lowering standards, and America is in danger of losing an entire generation of its future citizens whose problems may be even greater than our own.

Performance assessment measures, such as having to write an analysis of a problem or issue are more likely to serve as valid measures of higher-order thinking involved in synthesizing or evaluating, as well as critiquing, than do multiple-choice tests. And, while all assessments contain some bias, multiple-choice tests are more likely to reflect prior knowledge of the content contained in texts included in tests, knowledge acquired in certain cultural contexts, thus privileging students from certain cultural backgrounds over others.

***The use of tests to determine teaching quality.*** Equally problematic is the proposal to use test scores as the basis for judging teacher quality, as Governor Pawlenty has proposed. While districts are experimenting with the use of a range of different outcome measures to assess teacher quality, limiting that assessment to student test scores assumes a direct cause-effect relationship between student test performance and their instruction. Student performance is shaped by a wide range of different factors, only one of which may be the instruction they receive. While this issue may seem tangential to the large issue of the curriculum, adopting such an approach will only further encourage teachers to "teach to the test." Moreover, such a policy will encourage teachers to gravitate to schools, classes, or students perceived as those with high test scores.

***The high costs of testing in a period of budget cuts.*** The new program will entail the development of a whole new battery of tests for different subject matter areas and different grade levels. The overall development costs over a five-year period for the Virginia SOL was approximately \$100 million, much of which was devoted to developing, administering, and scoring tests.

The least expensive option is for the state to purchase norm-based off-the-shelf tests from testing companies. However, using such tests is a problem because norm-based tests are designed to measure a range of students' performance, as opposed to determining their performance relative to standards. This means that the state will need to contract with a testing company (four large companies hold a monopoly in what has become a \$700 million business) to create criterion-based tests aligned to the new state standards. (In some cases, states are pooling efforts and purchasing tests across states, but if the state standards vary, then this option is problematic.)

There are high costs associated with purchasing and scoring these tests. These costs will be incurred on top of the need to develop three new grade level reading and math tests for grades 4, 5, and 7 as mandated by the federal No Child Left Behind (NCLB) annual testing program. These costs will divert scarce funds away from school-focused funds; the tests will themselves not have any direct benefits to education. In an article on testing by Stephen Metcalf in *The Nation*, Wisconsin's State Superintendent of Public Instruction, Elizabeth Burmaster, indicated that the NCLB testing mandate will be counterproductive and redundant with their state's existing accountability system:

"The money we have for public education is going to lowering class size . . . but the federal legislation basically says, 'Nope, you have to go back in and redo your state assessment system.' To what purpose?"

In addition to paying for the development of the grade 4, 5, and 7 tests to meet the federal NCLB mandate, having to develop new state standards-based multiple-choice content tests for each subject matter for each of the upper grades will cost millions of dollars. It is difficult to justify diverting funds away from support of basic education needs when Minnesota faces a budget crisis, particularly when those funds will not produce any positive educational result.

### **Framing the Issue as an Either/Or Choice**

Unfortunately, Profile critics have framed the debate as an either/or option. Profile opponents posit that the Profile's higher-order thinking standards need to be totally jettisoned in favor of the content-based standards. Framing the debate as an either/or option assumes that nothing in the Profile, as developed over a ten-year period, is worth preserving. Furthermore, framing the issue in this manner assumes that teachers can simply focus on teaching content without considering the strategies and frameworks students need in order to use or apply this content. Framing the issue as an either/or choice fails to recognize that, despite all of the problems with its implementation, students and teachers did benefit from certain aspects of the Profile. In a financial crisis in which there is limited funding to mount an entirely new curriculum program, the notion of eliminating the positive aspects of the Profile simply because it is not consistent with a content-based approach rests on this false distinction between the Profile and the proposed alternative.

## Positive Influences of the Profile on Student Learning

There is evidence that the Profile was serving to foster higher-order thinking. In a study conducted by Professor Pat Avery, JoDiann Coler, and myself of 171 English teachers and 487 Social Studies teachers' perceptions of the Profile--published in a refereed journal, *Education Policy Analysis Archives* <http://epaa.asu.edu/epaa/v11n7/>

we found that, despite their concerns about the amount of time the Profile was taking to implement and the high level of frustration with CFL's leadership, a fair number of teachers were seeing positive outcomes in terms of student learning. Fifty-one% of the teachers perceived an increase in their students' higher order thinking as a result of the Profile. Also, those teachers who rated their training and/or their access to resources as "good" or "excellent" were much more likely to report positive student outcomes associated with the Profile.

The teachers noted that the Profile requires students to take responsibility for their own learning, apply their own knowledge, and work together collaboratively. They also noted the value of having students demonstrate proficiency through "hands-on" learning associated with higher levels of student involvement in the classroom. One teacher cited a specific example from a class simulation of the Treaty of Versailles she created to meet a content standard:

I can remember one young man two years ago who got into being part of the Turkish delegation at the Treaty of Versailles, and came up with original pieces of documentation that he just loved. And it turned him onto history. That's the payback. When you see the light bulb turn on and history becomes more than a textbook or a dry set of facts.

Teachers reported that rather than focus on their own instruction, they had to focus on student learning because the *Profile* required them to explain their learning expectations to students and parents, clarify criteria for evaluation, share these criteria with colleagues, and display student work. Most educational reformers would view these developments as positive.

Unfortunately, because CFL and the media provided the public with little specific information about the Profile over the past ten years, more ideologically-based critiques have prevailed in terms of shaping public policy towards the Profile. For example, Commissioner Yecke and others consistently cite the negative evaluation afforded the Profile by groups such as the Fordham Foundation. In contrast to our empirical data regarding the Profile, those evaluations represent the opinions of highly conservative consultants who reviewed the Profile based on their own ideological perspective. For example, the Thomas B. Fordham Foundation states that its mission is to articulate "a solid core curriculum taught by knowledgeable, expert instructors." It is also committed to the use of testing as a tool for accountability and the need for "competition:" "The Foundation also welcomes the "reinvention" of K-12 education to include such alternatives as charter schools, contract-management, student scholarships and other strategies for stimulating more education choices, greater competition and real consumer empowerment."

The external reviewers for the Profile included some of the most conservative educators in the country. In reviewing the Profile, they awarded the Profile a D- grade because of the lack of specific content and on the fact that the standards were not "measurable."

These external reviews reflected a bias towards a content-based approach to standards curriculum development. However, the public has been led to believe that because some conservative reviewers labeled the Profile a “failure,” then it needs to be replaced, a public-policy decision based on little or no empirical evidence. While the No Child Left Behind legislation demands that instruction be “results-based”—that decisions to adopt a certain approach be based on empirical evidence--the decision to drop the Profile appears to be based more on ideology than on results.

## **Recommendations**

Given the various problems with the proposed content-based standards curriculum, I recommend that the Senate:

- Reject the “either/or” critique, in which everything developed over a ten-year period under the Profile is rejected, by requiring that the successful programs developed by some districts under the Profile be included in the new curriculum. I recommend that those aspects of the Profile that have been successful, as indicated in our study and other research, be highlighted and employed in order to insure that teachers effectively integrate content and pedagogy consistent with achieving higher-order standards. Allow some districts and/or schools who have already demonstrated high levels of success under the Profile to continue to do so. Given the demands for choice in education, there is no reason why some schools who have been successful not be allowed to continue to use what has contributed to their success.
- Reject attempts to import off-the-shelf standards such as the Virginia standards that reflect an exclusively content-based focus, as well as outmoded, traditional curriculum inconsistent with the high quality of schooling in Minnesota. Stipulate that any standards developed for Minnesota need to be formulated by and for citizens and educators of Minnesota. This includes garnering input from the major subject-matter professional organizations in the state. This would not preclude study of high quality performance based standards developed in other states.
- Challenge the notion that any new standards can be developed and finalized by a group of subcommittees by March 31. This high-speed timetable deadline can only be achieved by importing existing content-based standards from other states (probably Virginia) that may have little or no bearing on Minnesota schools and, for the reasons cited above, will turn back the clock on Minnesota’s education.

Over a one-month time, the curriculum subcommittees will have little time to construct and evaluate standards in any thoughtful manner; it is more likely that they will simply be asked to rubber stamp imported standards. It is highly unlikely that the subcommittees, especially if they are primarily “people who have strong content knowledge,” will reflect a range of different educational philosophies and perspectives. These voices and

perspectives should include all of the major subject-matter professional organizations, many of whom had little initial input into the formulation of the Profile.

One of the lessons learned from the development of the Profile was that many teachers believed that they had little input into the formulation of the policy and standards. In our study (Avery, Beach, & Coler, in press), one teacher argued that teachers would have readily accepted the *Profile* “if it had been generated from the general teaching populace instead of the state ‘experts’.” In multiple interviews, teachers expressed frustration and resentment towards the state legislature and CFL for what they perceived as a top-down mandate. If teachers have little input in the formulation of these new standards, to which they will be held accountable, that frustration and resentment will only grow.

- Stipulate that any new set of standards developed for Minnesota need to include both content and higher-order thinking standards, as opposed to only content-based standards. In this balanced approach, some of the successful standards that were developed for the Profile could still be used, as well as revisions of these standards. New higher-order thinking standards could be introduced as consistent with new curriculum developments in the past few years.
- Provide teachers with assistance in formulating curriculum. One of the reasons that some teachers had difficulty implementing the Profile was that they never had a curriculum framework or roadmap for assisting them in implementing the Profile. The CFL curriculum frameworks developed over the past two years for many subject matter areas could serve as the basis for assisting teachers. I have worked with many teachers who cite examples of effective curriculum they have developed over the past several years. It is essential that teachers not be forced to suddenly change course in mid-stream and adopt a totally different approach to education.
- Recognize that funds for the purchase and administration of new tests (on top of the development of the new NCLB mandated tests for grades 4, 5, and 7) will divert funds away from the general education funding of schools, which will themselves be facing further cuts in teachers, resulting in increased class sizes. (Minneapolis has already laid off 290 teachers and may have to make further cuts with the new 2003-2004 budget).
- Explore options to multiple-choice testing as the sole measure of student learning. One alternative to the narrow focus of multiple-choice tests is the use of state-wide performance-based assessments employed in states such as Kentucky, Vermont, and Connecticut. While Minnesota must test students in grades 1-8 as mandated by the NCLB, as well as continue to administer the basic skills reading, math, and writing tests, any other new subject matter tests for the upper grades should be phased in over time and developed as performance assessment measures consistent with teachers’ classroom instruction.

## References

Amrein, A.L. & Berliner, D.C. (2002, March 28). High-stakes testing, uncertainty, and student

- learning *Education Policy Analysis Archives*, 10(18). Retrieved 2/2/03 from <http://epaa.asu.edu/epaa/v10n18/>.
- Avery, P., Beach, R., & Coler, J. (in press). The impact of Minnesota's "Profile of Learning" on teaching and learning in English and social studies classrooms. *Education Policy Analysis Archives*. <http://epaa.asu.edu/epaa/v11n7/>
- Graves, D. (2002, April/May). When testing lowers standards. *Reading Today*. [http://www.reading.org/publications/rty/archives/02apr\\_testing.html](http://www.reading.org/publications/rty/archives/02apr_testing.html)
- Hillocks, G. (2002). *The testing trap: How state writing assessments control learning*. New York: Teachers College Press.