A Systems Approach to Education Policy & Administration

By

David Wheat
President, Wheat Resources Inc.

P.O. Box 19234
Roanoke, Virginia 24019
www.wheatresources.com

888-667-8850 (toll free)
540-966-5167 (fax)
dwheat@wheatresources.com

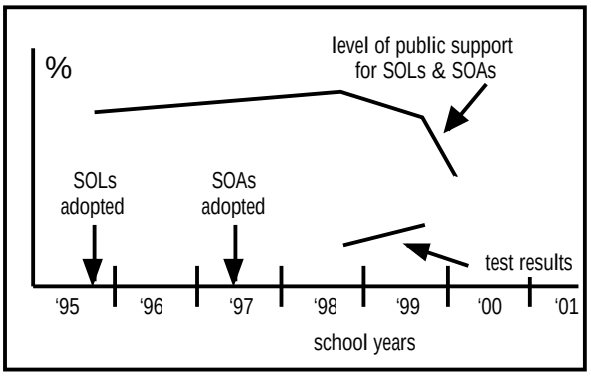

Figure 1. Change in SOL test results and public support for the SOLs and SOAs. Public support trends estimated by the author.

Think of this graph as a picture of a story unfolding rather than as a precise gauge. The direction of change is more important to the story than the quantitative precision, and here is the gist of the story: After the adoption of the SOLs in 1995, public support for the more rigorous standards rose gradually until falling in 1998, when the first test results were lower than expected. Support declined even more in 1999, despite test score improvement.

Those attentive to education policy in Virginia may disagree on the *slope* of the lines drawn in Figure 1 (i.e., the pace of change), but most will agree on the *direction* during the past five years. Between 1995 and 1998, the principle of holding students and schools to more rigorous academic standards seemed to be accepted gradually by more and more citizens of the Commonwealth.

The SOA 70 percent goal was established in 1997, but generated little concern at the time since no tests had been given. The first test results in 1998, however, brought a wake-up call, and public support declined a little. Despite improved test results the next year, the decline in support probably accelerated in 1999 because the dismal 1998 results had been discounted by many as a "start up" situation, while the still-lower-than-expected 1999 results received no comparable benefit of the doubt.[3]

Pressure for adjusting either the SOA goal or timetable has been building since the first round of testing in 1998. The Board of Education held its most recent hearings on these issues in several communities across the Commonwealth from November to January. Subsequent to the November hearings, Board President Kirk Schroder wrote seven public education advocacy organizations[4] that had jointly criticized the use of SOL tests for accountability purposes, inviting *them* to specify how *they* would design an accountability program.

Members of the Board of Education are justifiably concerned about the escalating attacks on the SOLs and SOAs, particularly when the raw data suggest that progress is being made in upgrading students' knowledge and skills.

Compared to 1998 results, 1999 statewide average scores improved on all 27 SOL tests administered to 3rd, 5th, and 8th graders and high school students. Even 60 percent of the schools improved their students' scores on the history tests, where performance has been the weakest. The number of schools passing all four tests tripled, from 39 to 116 (although the percentage of schools reaching that SOA goal was still extremely low).[5]

Those of us who were part of the team of teachers and Board members who framed the SOLs in 1994-95 believed at the time--and still believe today--that Virginia's public school students need to reach a higher level of academic achievement. Those who later developed the SOL testing program and

those who promulgated the SOA goal shared similar convictions.

Given the opportunity, Virginia's students should continue to make progress toward the learning goals of the 1995 *SOLs*, even if most schools don't actually reach the accreditation goal of the 1997 *SOAs*. Yet, concern over the accreditation standards has weakened public support for the SOL program and has jeopardized the future of the rigorous learning standards.

The SOLs and the SOAs are separate issues, however, and preserving the SOLs (standards that have been judged among the best in the nation in three[6] subject areas) requires aggressive remedial steps to modify the accreditation goal.

Using a system dynamics approach to policy analysis, this study evaluates the accreditation policy and develops an alternative: *value-added accountability*. The findings of the study are presented in three parts. First, the report clarifies the problem by contrasting the way the SOL system was *supposed* to work with the way it is actually operating today. Particular attention is given to the unintended, adverse side-effect of the SOA 70 percent goal.

Next, the 70 percent goal is shown to be an unrealistic target for most schools, which intensifies the adverse side-effect and threatens the future of the SOL program.

Finally, a new accreditation policy is recommended, one that is consistent with the desired performance of the SOL system, that insures accountability, and that supports rather than undermines rigorous learning standards.

THE PROBLEM

Dealing effectively with the accreditation policy problem requires identifying it explicitly, understanding how it developed, and perceiving its adverse effects on the standards of learning. All three requirements can be accomplished by contrasting the current SOL system with the one envisioned by those who developed the original SOL program in 1995.

The original SOL system was supposed to produce annual improvements in student learning of content-rich knowledge and valuable skills. The hoped-for pattern of cause-and-effect can be summarized as follows:

• The SOLs would have a positive impact on classroom learning productivity as teachers gained experience teaching the standards and students gained capacity to learn more due to prior SOL instruction.

• Greater learning productivity would raise students' annual learning rates (i.e., the total knowledge and skills learned each year).

• Higher learning rates would be reflected in higher standardized test scores.

• Better test results would raise teacher and student motivation and boost learning productivity, which would further enhance learning rates and test scores.

• Better test results would also raise public support for the learning standards.

• Greater support would minimize political pressures to change the standards or reduce their impact, and the SOLs would become a permanent feature in Virginia's education policy landscape.

This description of how the system was *supposed* to work is diagrammed in a model of the SOL system in Figure 2.

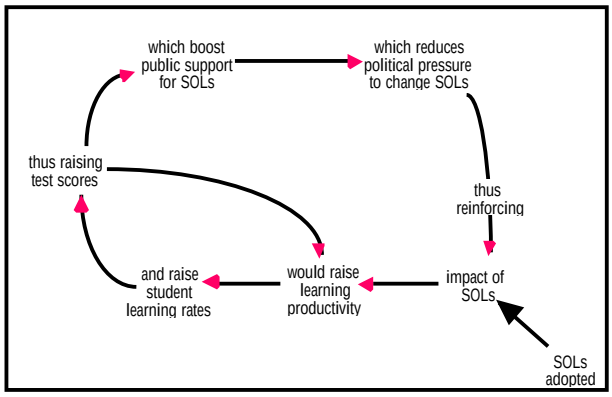

Figure 2. The continuous cause-and-effect relationships *intended* when the SOLs were adopted in 1995.

Understanding the difference between the 1995 SOL system in Figure 2 and the system operating today requires some familiarity with two conceptual building blocks: (a) cause-and-effect and (b) feedback.

While the general concept of cause-and-effect is familiar to public officials, the diagrams in this report may not be. Therefore, the full model of the SOL system is developed in stages, beginning with the simplest cause-and-effect relationships. Feedback, as the term is used in system dynamics, may be a less familiar concept. Again, simple illustrations precede more complex representations of dynamic system feedback.

Cause and Effect

Implicit in any policy--whether in business or government--is the concept of cause-and-effect. A new policy (the "cause") is intended to produce some beneficial change (the "effect") in the *status quo*. The higher learning standards adopted in 1995, for example, were intended to have a positive impact on student acquisition of knowledge and skills.

Such cause-and-effect relationships can be diagrammed, albeit simplistically, as in Figure 3.

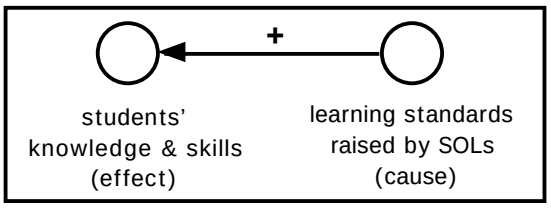

Figure 3. Implicit cause-and-effect relationship between higher learning standards and knowledge and skills gained by students. The plus (+) sign means that a change in the learning standards causes a change in student learning in the *same direction*. An increase would cause an increase, and a decrease would cause a decrease.

The arrow in the diagram indicates the direction of causality (i.e., "the learning standards" have an impact on "students' knowledge and skills."). The plus sign (+) indicates a positive correlation between the cause and the effect (i.e., *raising* the learning standards causes an *increase* in students' knowledge and skills, and a *decrease* in the former causes a *decrease* in the latter.) Had there been a minus (-) sign, it would have meant that *increasing* one factor caused a *decrease* in the other, and conversely.

Certainly, no one working on the SOLs in 1994-95 believed that mere adoption and publication of higher standards by the Board of Education would automatically lead to higher levels of learning by students. Thus, the model in Figure 3 simplifies the assumed cause-and-effect relationship by leaving out well known intermediate factors such as teacher ability and motivation to emphasize the new standards and develop effective lessons, student capacities and motivation to learn the SOL content, and resources necessary to enhance both teaching and learning.[7]

The interaction of these intermediate factors constitutes the "learning productivity" factor in the center of Figure 4, which should be interpreted as follows: The impact of higher standards causes learning activities to be more productive (quantitatively and qualitatively), which leads to more learning per year (i.e., a higher learning rate).[8]

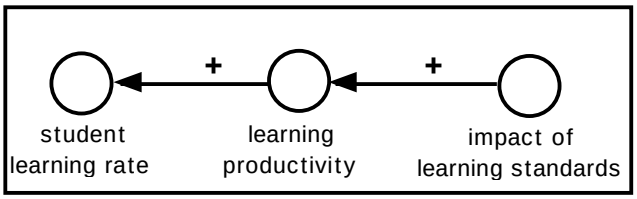

Figure 4. Learning productivity is the combined influence of teacher, student, and resources for instruction. The student learning rate is the annual gain in knowledge and skills.

Feedback

Even in Figure 4, however, the overall single-direction causality remains. That is to say, the model--and the implicit thinking behind it--does not reveal any feedback potential. It still implies a one-way system, meaning that the 1995 policy (i.e., the new SOLs) would always be a "cause" and would never be affected by the dynamics of the system.

The system dynamics approach to policy analysis, on the other hand, assumes that any policy is part of a system that, sooner or later, generates new issues due to perceptions of the policy's effectiveness or side-effects. The education policy that raised Virginia's learning standards is no exception to this normal systemic feedback process.

The diagram in Figure 5 displays such feedback effects. This model's simplicity stems from its omission of factors that explain how policy effects change political support levels and how government institutions generate new public policies. In this report, complexity will be added to the models a little at a time and only when necessary.[9]

In plain English, starting[10] with the impact of the 1995 policy (i.e., the SOLs), the Figure 5 model says:

- The impact of the SOLs makes learning activities more productive;
- Higher learning productivity raises the student learning rate;
- When learning rates rise, that boosts public support for the SOLs;
- Growing support reduces the political pressure for a policy change.
- A reduction in pressure for change increases the credibility and, therefore, the impact of the SOLs.[11]

The next step in the unveiling of the model of the 1995 SOL system is to display the impact of standardized test results. Ideally, the mere designation of the SOLs as mandatory would be sufficient to produce the desired performance in schools. However, the framers of the SOLs believed

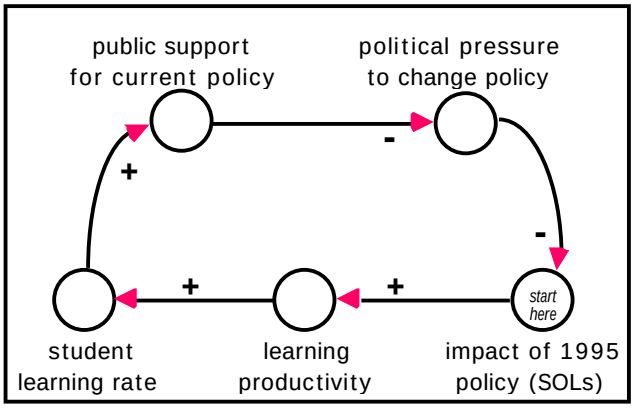

Figure 5. Feedback in the SOL policy system. Minus (-) signs mean that the cause and effect move in opposite directions. Thus, an increase in public support would cause political pressure to decline. A decline in political pressure would then increase the impact of the SOLs on learning productivity.

that accountability would be fostered by a testing program that would gauge learning progress and positively reinforce the incentives for schools and students (i.e., produce effects in the same direction as the SOL impact). Thus, from the beginning, a testing program was envisioned by those who developed the SOLs, even though the actual tests were designed by a different group and adopted at a later date.

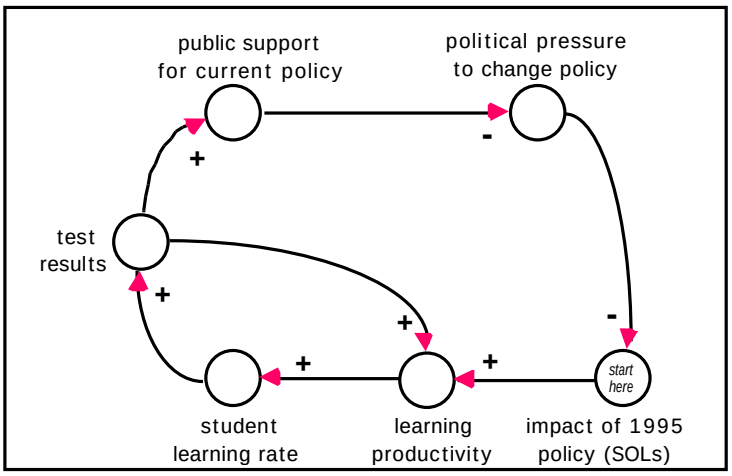

Figure 6. General SOL model showing the original conception of the test effects. Improvements in test results would cause an increase in public support for the SOLs, as well as an increase in learning productivity by boosting teacher and student morale.

Figure 6 illustrates the expected impact of test results when the SOL program was adopted in 1995. It was envisioned that test results would show regular improvements, at least during the first decade of the program. Rising test scores, in turn, were expected to raise public support for the SOLs, plus contribute to an increase in learning productivity by boosting the morale of teachers and students.

On a parallel track, learning productivity would also experience annual growth as teachers improved their SOL instructional effectiveness with experience, and students became increasingly capable of learning more due to SOL instruction in prior years. Higher productivity would generate a new round of higher learning rates, which would be reflected in higher test scores.

Better test results would raise support in the communities and productivity in the schools, and the result would be another increase in the learning rate the following year. A "virtuous upward spiral" would have been set in motion.[12]

The diagram in Figure 6 may be thought of as the "general model" of the 1995 SOL system. It could lead to both desirable or undesirable performance. For example, if something were to cause learning productivity to fall, then learning rates would fall, test scores would fall, public support would fall, and that would increase political pressure to alter the SOL program.

On the other hand, the diagram in Figure 7 (a reproduction of Figure 2) illustrates the optimistic scenario envisioned by the developers of the SOLs (i.e., the way the system was *supposed* to work).

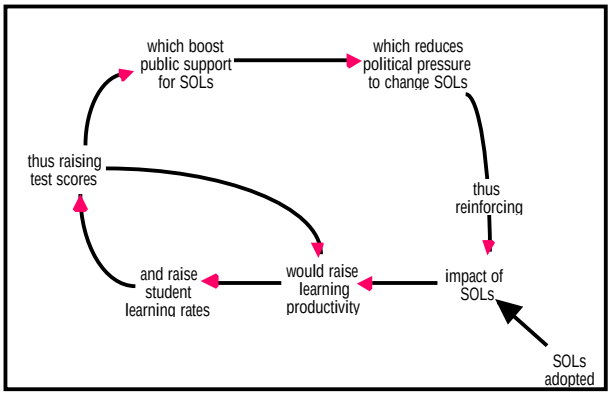

Figure 7. The scenario envisioned for the SOL system in 1995, a special "best case" version of the general model in Figure 6.

The models in Figures 6 and 7 can't be telling the whole story, however, because the dilemma confronting the Board of Education is *falling public support* (with its attendant political consequences) amidst

rising test scores that imply improvement in student learning. In short, the system modeled in Figures 6 and 7 cannot account for the real-world performance of the SOL system. Something critical is still missing.

The missing elements of the system model are those that were put in place with the adoption of the Standards of Accreditation 70 percent goal in 1997.

The Current System

The SOAs raised school performance standards by establishing a long-term goal for test results (70 percent of a school's students passing all four tests by 2007). *An unintended side-effect was a change in the performance pattern of the SOL system; a change that has undermined rather than supported the learning standards.*

The 70 percent goal raised expectations--among parents, the general public, and the media--regarding acceptable test results each year, locally and statewide. In effect, it created annual targets for "expected test results."

At first glance, expecting 70 percent of the students to pass all four tests seems reasonable and achievable. Of course, few if any parents anticipated that their local schools would reach the SOA goal overnight. Nevertheless, it is likely that the expected test results during the first two years have been much closer to the 70 percent mark than is realistic.

Moreover, those parental expectations are likely to rise each year, along with their impatience for more rapid progress toward the long-term goal. Indeed, the implicit local test result targets are likely to rise faster each year than actual improvements in test scores. Whenever expected test results exceed actual test results, a "Gap" exists.

The general version of the SOL system operating today is illustrated in Figure 8, and the diagram reveals how a Gap emerges when expected test results exceed actual test results.

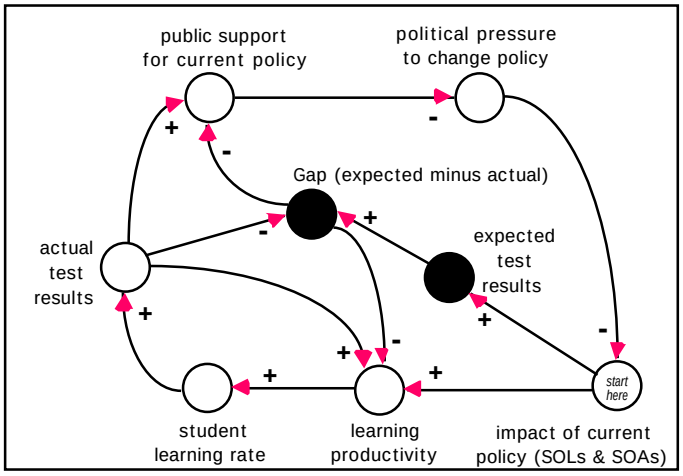

Figure 8. The general version of the current SOL system, showing unintended side-effect of the SOA 70 percent goal established in 1997. Note that public support and learning productivity are influenced by both positive and negative feedback, and the net effects depend on the relative strength of those feedback loops.

Note carefully the negative effects of the Gap, which could counteract the positive effects of rising test scores on both learning productivity and public support.

When the Gap grows, it has a depressing effect on teacher and student motivation, thereby reducing learning productivity and causing learning rates to be lower than they otherwise would be.

On the other hand, as long as actual test results continued to improve in absolute terms (i.e., without comparing them to expected test results), productivity would be receiving upward boosts, and learning rates would be propped up.

Likewise, the level of public support for the SOLs is affected positively by the

test results and negatively by the Gap between the expected and actual test results.

Improvement in test scores (without reference to any target) would bolster public support. A growing Gap, on the other hand, would continue to frustrate parents and lower public support.[13]

Declining public support affects the political climate and landscape, and increases the pressure for policy changes that would lower the standards--*both SOAs and SOLs as long as the two are inseparable in the public mind.* Such reactive policies would lower the standards and eventually cause learning rates to decline.

productive effects of the long-term goal established by the SOAs. Whether that potential is realized depends on the prospects for large differences between expected and actual test results (i.e., the size of the Gap), an issue examined in detail in the next section of this report.

Thus, the original SOL goal of students learning more at each grade level every year could be undermined in two ways as a result of the SOAs, *even when learning rates and test scores are on the rise*.

The negative effects of a Gap between expected and actual test results could overwhelm the positive effects of

The scenario depicted here occurs when:

- there is a Gap (i.e., expected test scores exceed actual test scores),
- the net effect on public support is negative (i.e., the impact of the Gap overwhelms the impact of rising test scores), while
- the net effect on learning productivity is positive (i.e., the impact of the SOLs and rising test scores outweigh the influence of the Gap) until political pressures dilute or rescind the SOLs.

Bottom Line Effect: Failing to distinguish between the SOLs and the SOAs, this scenario leads to dilution or elimination of the current learning standards program.

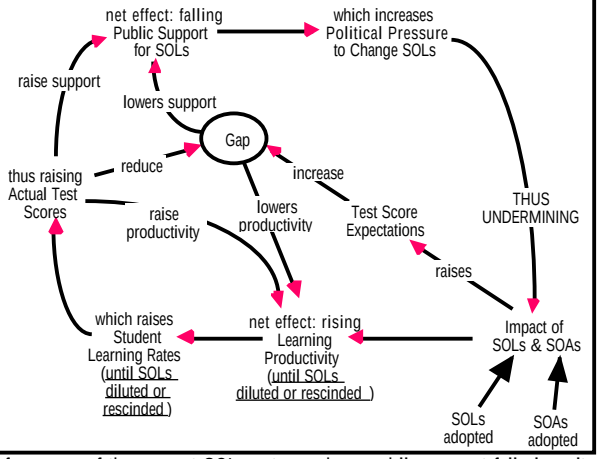

Figure 9. The Dismal Scenario: A "special case" of the performance of the current SOL system, where public support falls in spite of improvements in test scores, and political pressures undermine the impact of the SOLs and SOAs. The impact of the SOA 70 percent goal causes expected test results to be higher. Higher expectations increase the Gap, which lowers public support and triggers the adverse chain of events that undermine the SOLs.

This depressing scenario is illustrated in Figure 9, which is a special case of the general model in Figure 8. This scenario is not inevitable, of course. The actual performance of the current SOL system depends on the relative strength of the feedback loops influencing learning productivity and public support.[14]

For the moment, however, it is important to note the *potential* for counter-

actual improvements in test scores, with the result being lower learning productivity in the classrooms and lower public support in the communities.

Both developments eventually lead to lower learning rates, either in the short run by falling productivity or in the long run as political pressures mount to "do something about those unrealistic standards."

But is it the *learning* standards that are unrealistic, or is it the *accreditation* standard (i.e., the 70 percent goal) that is unrealistic? That distinction is too often overlooked in the public debate. The next section of the report provides an answer to that question by examining whether the SOA "70 percent goal" is a target within reach by a majority of Virginia's public schools.

The realism of that SOA "70 percent goal" has a direct bearing on the realism of parents' expectations for test results (and the resulting size of the Gap) and the relative strength of the critical feedback loops in the system. The realism of that goal determines which scenario is likely to emerge from the general model of the system in Figure 8:

> The "Up" Scenario: If the goal proves to be realistic, there will be steady increases in lerning rates as the SOLs gain public support and increase their classroom impact.
>
> The "Down" Scenario. If the goal is not realistic, learning rates will fall as SOLs lose support in the communities and credibility in the classroom, and are eventually diluted or rescinded.

HOW REALISTIC IS THE TARGET?

The high leverage point in today's SOL system is the Gap between actual test results and expected test results. That is where the system's dynamic relationships have the potential to reverse the virtuous upward spiral envisioned in 1995. From the Gap flows the potential negative forces that could reduce learning productivity in the schools and public support in the communities, and start a vicious downward spiral of the entire SOL program.

The positive impact of rising test scores on learning productivity and community support would probably offset the negative impact of a small Gap, and keep the SOL program on course. However, a large Gap, particularly one that increases yearly, spells trouble for the SOLs. To get a handle on the prospects for the Gap, we need to estimate the realistic prospects for reaching the SOA goal. That has a direct bearing on the realism of the expectations for test results in local communities. More realistic expectations would produce smaller Gaps.

When the SOL assessment tests were first administered in the spring of 1998, only 2.2 percent of Virginia's schools achieved the SOA goal (i.e., had 70 percent of their students passing SOL tests in each of the four subject areas). In the 1999 round of testing, scores improved considerably, but the number of schools "passing" was still only 6.5 percent of the total.

As this analysis progresses on the next few pages, it will be important to remain clear on the distinction between (a) the percentage of *schools* in which at least 70 percent of the students passed the tests in the four subject areas, (b) the percentage of *students* who passed specific tests.

The 1998 and 1999 statewide "pass rates" for *students* on all tests in various grades are included in Appendix B. However, the table in Figure 10 provides a

composite summary of those pass rates in the four subject areas:

Test	yr	Percentage of Students Passing at Each Grade Level 3	5	8	HS[15]
English	98	53	67	66	72
English	99	61	75	69	78
Science	98	63	59	71	61
Science	99	68	67	78	70
Math	98	63	47	53	41
Math	99	68	51	60	56
History	98	49	33	35	44
History	99	62	46	40	49

Figure 10. Percentage of students passing each category of tests, by grade level, in 1998 & 1999. Source: VDOE press release, July 1999

Compared to 1998, a higher percentage of students passed in every category of testing in 1999. When percentages across grade levels are averaged, here is the approximate percentage of students passing each set of tests in 1999:

English	70%
Science	70%
Math	60%
History	50%

Figure 11. Average Pass Rates

What is the outlook for most schools reaching the 70 percent target, either in the short run or by the end of the 2006-07 school year? How realistic is that goal?

Those who are most optimistic tend to describe the challenge this way:

> "Achieve passing levels on math and history by the 10-20 percent who failed those tests (while also maintaining English and science pass rates at 70 percent or above and raising performance by those students who are still unlikely to pass)."

*However, even if 70 percent of the **students** statewide eventually passed the four tests, there is troubling evidence that only a small percentage of the **schools** would be reaching the SOA goal at the same time.*

This is due to the unequal geographic distribution of students who are most likely to pass the tests. In a vivid example of the "80/20 Rule", a majority of the students passing all four tests are disproportionally located in a few high achievement schools.[16]

Sixty-six percent of the schools that achieved the SOA goal in both 1998 and 1999 were located in just three school divisions having 16 percent of Virginia's schools. In 1999, 26 percent of the schools in Fairfax, Henrico, or Chesterfield county met the SOA goal, in contrast with 2.6 percent of the rest of the schools in the state.[17] A school in those three counties was 10 times more likely to achieve the 70 percent goal than a school elsewhere in the Commonwealth.

Of course, the 10 to 1 ratio is not carved in stone, nor is the disparity between the statewide student pass rate and school pass rate. Both ratios would decline as more students passed the four tests statewide, but not at a constant rate, due to the

disproportionate number of "passing students" in a few school divisions.

In Figure 12, the horizontal axis represents the average of student pass rates on the four SOL tests. The vertical axis is the percentage of schools that reach the SOA 70 percent goal. The curve on the graph shows the relationship between the two percentages.

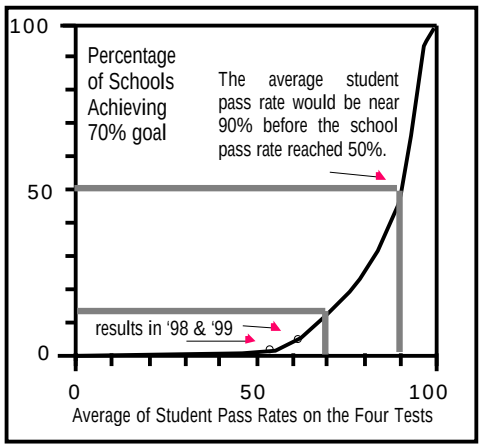

Figure 12. Projected relationship between average student pass rates and school pass rates, based on statistical regression analysis of 1998 and 1999 data for all elementary & middle schools (which account for over 80 percent of all schools and 90 percent of the schools achieving the SOA goal thus far.

The graph in Figure 12 indicates that the change would be slow at first and then accelerate if statewide student pass rates ever approached 100 percent. At that limit, obviously, with every student in the state passing all four tests, then every school would have reached the SOA goals. (See Appendix A.)

The findings in this study indicate that about 27 and 35 percent of Virginia's students reached their individual goal of passing all four tests in 1998 and 1999, respectively. The average of the pass rates on the four tests was 54 and 62 percent in those years. That *student* performance in 1998 and 1999, however, translated into just 2.2 and 6.5 percent of the *schools* reaching their 70 percent goal in those two years, as illustrated in Figure 12.[18]

From the 70 percent mark on the horizontal axis in Figure 12, a shaded vertical line has been drawn. At the curve, it meets another shaded line from the left axis.

Less than 15 percent of Virginia's schools are expected to be reaching the SOA goal when the average of student pass rates on the four tests reaches the 70 percent level statewide.

Higher on the curve, a shaded line from the 50 percent school pass rate mark intersects a shaded line from the 90 percent student pass rate mark.

This suggests that the statewide average of student pass rates on the four tests could approach the 90 percent level before even half of the schools in the state reach their SOA goal.

The graph in Figure 12 says nothing about the likelihood that student pass rates will actually reach particular levels. It merely projects what the corresponding school pass rate would be at various levels of student performance.

However, it should be clear that the accreditation of many (and perhaps more than half) of Virginia's *schools* will be in jeopardy even if Virginia's *students* statewide do even better than the goal established by the Board of Education!

Given the current average student pass rate of 62 percent and the improvement required to take it to the 80-90 percent range, an important conclusion also becomes clear:

The SOA goal for each school, while admirable for all schools and possibly attainable by many in the long run, is not a realistic goal for the majority of Virginia's public schools in the time frame contemplated under current policy.

Because the SOA 70 percent goal is unrealistic in the short run, test result expectations tied to that goal will also be unrealistic in the short run. Consequently, in most localities, parental expectations for their school's test results will almost certainly be higher than actual test results year after year. That means the Gap between actual and expected test results will be persistent and large, even when actual test scores are on the rise. Learning productivity in the schools and public support will be depressed, political pressures to lower or rescind the standards will increase, and the SOLs will face extinction. The last section of the report contains policy recommendations aimed at avoiding this dismal scenario.

POLICY RECOMMENDATIONS

When policies are not self-executing, those responsible for implementation should be accountable. An accountability process should satisfy at least two criteria. First, it should provide an unambiguous means for measuring the effectiveness of the policy. Second, it should provide incentives for effective implementation without producing side-effects that undercut the purpose of the policy.

Since "student learning rates" cannot be measured directly, the effectiveness of the SOL policy cannot be evaluated without an indirect indicator--namely, tests that purport to reflect the amount of learning that has taken place.[19]

The original SOLs were focused on *student* achievement, while the 1997 SOAs have shifted excessive attention to *school* performance. The current accreditation policy, with its 70 percent goal for schools, has generated confusion about which test results should be measured and evaluated.

Schools are responsible for implementation, and the accountability principle requires tracking their performance. However, that should not obscure the fact that school performance is merely a proxy for student performance. Unless carefully designed, school performance measures often mirror the students a school *has* rather than what a school *does* with the students it has. In other words, the current SOA policy fails to satisfy the criterion of providing unambiguous means for assessing the effectiveness of the SOL policy.

Even more critical, however, is the failure of the current SOA policy to provide incentives for effective implementation without adverse side-effects. The 70 percent goal creates unrealistic parental expectations for annual test results. The decline in public support resulting from unrealized expectations creates political pressure to dilute or rescind both the learning and accreditation standards.

School Accountability Policy

A new accreditation policy is needed to promote SOL accountability while supporting the learning standards, rather than undermining them as the current SOA policy does. The strategy suggested by this study is to repair the current SOL system by eliminating the adverse side effects of the 70 percent goal shown in Figure 13.

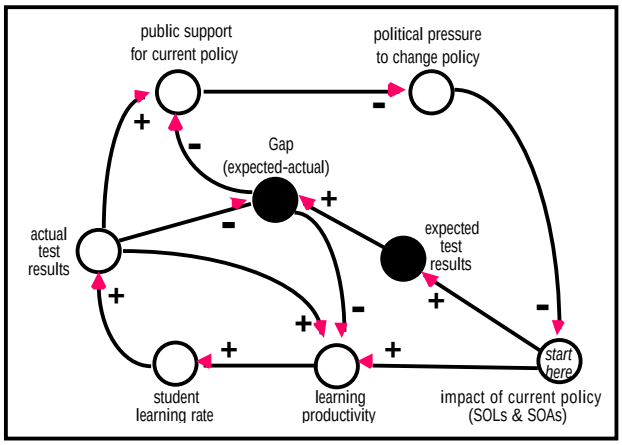

Figure 13. Today's SOL system, with adverse side-effects resulting from the SOA 70 percent goal established in 1997, needs systemic surgery.

That is, systemic surgery is needed to remove the Gap in Figure 13, so that learning productivity and public support will be affected only by changes in *actual* test results, rather than by comparisons with unrealistic *expected* test results. Such changes would restore the originally intended SOL system, displayed in Figure 14 (which is a reproduction of Figure 6 on page 7).

Also, an accreditation standard should focus on the quality of *instruction* in schools rather than the quality of *students* in schools. The current SOA policy, however, rewards (or punishes) schools on the basis of reaching (or failing to reach) an arbitrary finish line without regard to the starting line. Instructional quality may actually be superior in a school that makes substantial progress toward the SOA goal, compared to a school reaching the goal with large numbers of bright, self-motivated students requiring less innovative instructional methods to succeed. Which school deserves to lose accreditation? Neither.

There is something wrong with an accreditation policy that says 66 percent of the schools in Virginia currently deserving accreditation are located in just three counties having 16 percent of the state's schools. Any accreditation standard worthy of the label should not produce outcomes leading to such a distorted conclusion.

Accreditation should not depend on the students a school has. Rather, it should depend on what a school ***does*** *with the students it has.*

Consistent with this principle is an accreditation policy based on *value-added accountability*, which means holding schools responsible for performance improvements in each class cohort[20] that rises from one level to the next in a school division.

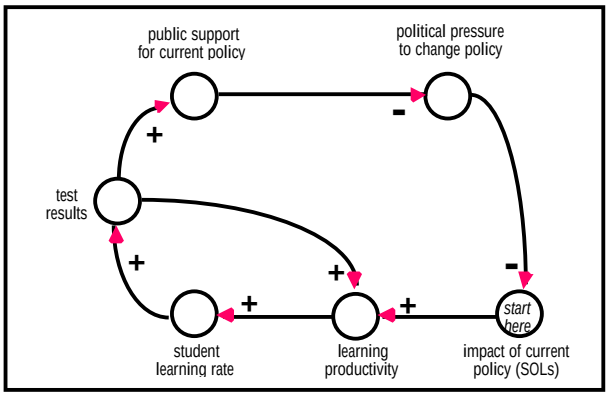

Figure 14. After systemic surgery: restoration of the continuous cause-and-effect relationships of the system intended by the 1995 SOLs.

To illustrate, the cohort of students who are now in the 5th grade will take the 5th grade SOL tests this spring. The effectiveness of an elementary school should be evaluated by comparing this year's 5th grade pass rates in each category of tests (e.g., math) with the pass rates achieved by the same cohort of students when they took the 3rd grade SOL tests in 1998.[21] Looking ahead to 2003, the middle school should be evaluated by comparing its 8th grade SOL pass rates that year with this year's 5th grade SOL pass rates for the same cohort.

Since high school SOL tests are end-of-course tests, not all students taking tests will be in the same grade. Therefore, to measure value-added at that level, each test-taking student's performance (pass or fail) should be compared to that student's 8th grade performance, and the aggregate improvement rate for the school could then be calculated.

Schools that fail to add value consistently (according to Board of Education definition) should be sanctioned, and be subject to losing accreditation.[22]

RECOMMENDATION 1. Amend the SOAs so that schools will be accredited on the basis of improvement in class cohort pass rates on individual SOL tests.

Student Accountability Policy

Current SOA policy also requires secondary students to pass end-of-course SOL tests to earn certain graduation credits, beginning with the Class of 2004. That policy has not been critiqued thus far in this report, primarily because it can be addressed without complex analysis. It is not unreasonable to expect students to pass end-of-course SOL tests in order to receive graduation credits for courses in which they have enrolled. If properly constructed to assess what students were supposed to gain from the SOLs for a particular course, such tests should be no more difficult than a teacher-made final exam. However, the timetable for holding students accountable should be amended to reflect the fact that the seniors in 2007 will be the first class with twelve years of SOL instruction.

RECOMMENDATION 2. Amend the SOAs so that end-of-course SOL test requirements apply initially to students in the graduating class of 2007, and to each class thereafter.

CONCLUSION

Accountability is imperative when a policy is not self-executing. Designing an accountability standard, however, requires utmost attention to the incentives it sets in motion. It must highlight the right indicators, and it must enhance the effectiveness of policy implementation.

A value-added approach to accountability would restore the SOL principle of cumulative learning growth.

Moreover, with the focus on strengthening class cohorts, the key "expectation" would be for test score *improvements*. There would be none of the adverse side effects resulting from comparisons with unrealistic targets. Teachers and students could tackle the SOLs in a spirit of discovery rather than an aura of anxiety.

Those committed to the purpose of the 1995 SOL program--rigorous *learning* standards leading to higher levels of academic achievement by Virginia's students--should endorse amending a 1997 *accreditation* standard that threatens to undermine the fundamental reform it was intended to support.

EPILOGUE

Upon release of this study in February by the Thomas Jefferson Institute for Public Policy, the initial reaction of the Governor and the President of Virginia's Board of Education was to treat the recommendations as proposals for watering down Virginia's education standards. In a February 9, 2000 *Washington Post* article, both reaffirmed their support of the 70 percent standard.

The Board president said that he would only consider “improvement” as an accreditation criterion if it could be incorporated into a yardstick that maintains the 70 percent standard. A week later, he told the Hampton Roads *Virginian-Pilot* that an “improvement” from 7 percent to 10 percent, for example, would not be an achievement worthy of praise or accreditation.

In that example, of course, he is right. However, under current policy, a school could improve its pass rate from 20 percent to 60 percent and still be *denied* accreditation. Such a disincentive structure can only be counterproductive for educators in many schools. In addition, it undermines public support for standards that label such schools "losers."

In late April, two months after this study was released, Virginia's Board of Education published proposed changes to the Accreditation Standards. The Board will vote in in June on amendments that, while continuing to give lip service to the 70 percent standard, will result in more schools "passing*" even if test scores don't improve*.[23]

The real "watering down" predicted by our system dynamics model has now begun.

This back-door departure from the 70 percent standard by redefining what constitutes "passing" is a sad chapter in a story that, like all policy processes, never ends. Rather than admit the weakness in the current accreditation standard and its corrosive effect on the much more important learning standards, the Administration seems to be maintaining just the facade of a "tough" policy while, in fact, retreating.

Such an approach may relieve some of the local community pressure in the short run if more schools do, in fact, retain their accreditation. At first, few parents will know that more schools are "passing" due to grade inflation rather performance improvement.

It won't remain a secret, however. During Virginia's state legislative and gubernatorial campaigns in 2001, the minority political party is certain to spotlight both the unpopular accreditation issue and the Administration's awkward handling of it.

Thus, pressure continues to build for dilution or outright recision of Virginia's education standards, including not only the discredited accreditation policy but also the nationally recognized learning standards. The system is never static. Today's policy is tomorrow's issue.

Appendix A

Pass Rate Growth Methodology

Derivation of the pass rate growth curve in Figure 12 (page 11) required successive comparisons of curves based on different assumptions about the variance in student test performance from one school to another, plus a non-linear regression analysis and computer simulation of actual 1998 and 1999 results.

Alternative Assumptions Regarding Variation in Student Test Performance

Case A1. No Variation. This assumption is that students are distributed randomly across the state and among schools, with no difference in their ability to pass the SOL tests. This leads to a projection that all schools will reach the 70% goal in the same year. Thus the "curve" in Figure A1 is a straight line along the horizontal axis (i.e., "zero schools passing") that turns vertical at the point where 70% of the students pass the four tests statewide. The vertical line goes to 100%, meaning that all schools pass at the same time.

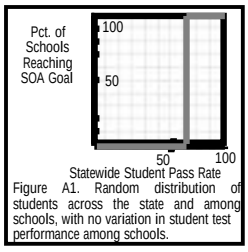

Figure A1. Random distribution of students across the state and among schools, with no variation in student test performance among schools.

Case A2. Only Random Variation. Although still assuming random distribution of students, this assumption admits that there will be random fluctuations in test scores from school to school even as all schools move toward the 70% goal. Some schools will get there sooner and others later than the average school, which will get there when statewide student pass rates are 70%. This produces a non-linear curve similar in shape to Figure 12 in the report, but it slopes upward much sooner. It predicts that 70% of the schools will achieve the SOA goal when 70% of the students statewide are passing the four tests.

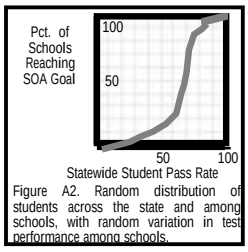

Figure A2. Random distribution of students across the state and among schools, with random variation in test performance among schools.

Case A3. Test Score Variation that is Not Random. This view recognizes that students are not randomly distributed among schools. In both 1998 and 1999, 66% of the schools reaching the SOA goal were located in three school divisions having just 16% of the schools in the state. Given that those three school divisions are accounting for a big fraction of the increase in "accredited schools" until just about all of their schools are accredited, the increase in student pass rates will not generate as many newly accredited schools as quickly as Figure A2 predicts, since many of the schools in those divisions will exceed the 70% mark. This makes the curve extend further to the right before it curves upward steeply, as shown by the solid line in Figure A3.

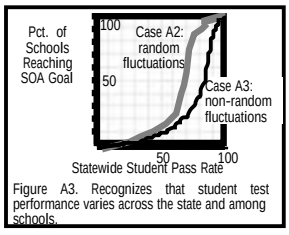

Figure A3. Recognizes that student test performance varies across the state and among schools.

Derivation of Figure 12. In Figure A4, each dot represents a school division's elementary and middle schools (which account for nearly 85% of all schools in Virginia, and 90% of the schools reaching the SOA goal). At each extreme point (0% and 100%), the dot is weighted by the total number of school divisions (i.e., all school divisions would have 100% of their schools passing when their average student pass rate reached 100%, and conversely with 0%.) Between the extremes, the dots show the actual 1998 and 1999 results of SOL testing at the school division level, and the two hollow circles (not part of the regression model) represent the actual statewide results. What appears to be a solid straight line ranging from about 10% to 45% on the horizontal axis is actually dozens of school division dots bunched together. The solid upward sloping curve is the "best fit" regression line, based on a non-linear (exponential) model, weighted by the number of schools in each division. Thus, dots representing divisions with larger numbers of schools have more influence on the shape of the curve, an important point to remember since many of the dots above the curve "that are easy to see" contain just one school and exert little influence on the shape of the regression line (see notes on the graphs). The dashed curves on both sides are the 95% confidence limits.

Derivation of Figure A5, which is the same as Figure 12 on page 11, requires staying within the 95% confidence limits below. The dashed vertical lines were calculated as upper and lower limits. And, of course, the curve should begin at 0%, end at 100%, intersect the actual 1998 and 1999 results, while maintaining a smooth exponential growth until leveling off as it approaches the 100% mark. Given these constraints, the resulting curve in Figure A5 was drawn, which is the same as Figure 12.

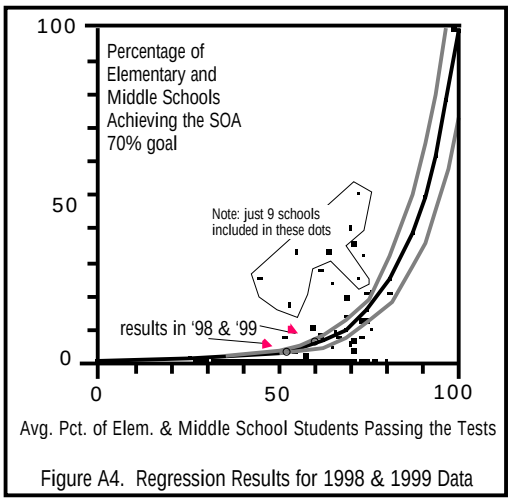

Figure A4. Regression Results for 1998 & 1999 Data

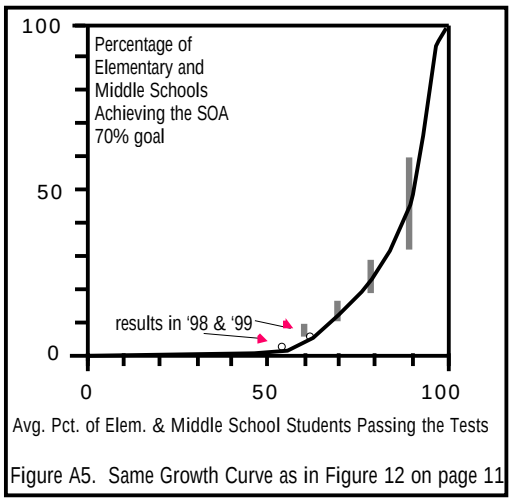

Figure A5. Same Growth Curve as in Figure 12 on page 11

Appendix B

Virginia

Standards of Learning Assessments

Statewide Passing Rates

Spring 1998 and Spring 1999

SOL Test	1998 Passing Rate (%)	1999 Passing Rate (%)	Change from 1998 to 1999
Grade 3			
English	53	61	+ 8
Mathematics	63	68	+ 5
History & Social Science	49	62	+13
Science	63	68	+ 5
Grade 5			
English: Reading, Literature, & Research	68	69	+ 1
English: Writing	65	81	+16
Mathematics	47	51	+ 4
History & Social Science	33	46	+13
Science	59	67	+ 8
Computer/Technology	72	81	+ 9
Grade 8			
English: Reading, Literature, & Research	64	67	+ 3
English: Writing	67	70	+ 3
Mathematics	53	60	+ 7
History & Social Science	35	40	+ 5
Science	71	78	+ 7
Computer/Technology	63	72	+ 9
High School			
English: Reading, Literature, & Research	72	75	+ 3
English: Writing	71	81	+10
Algebra I	40	56	+16
Algebra II	31	51	+20
Geometry	52	62	+10
Earth Science	58	65	+ 7
Biology	72	81	+ 9
Chemistry	54	64	+10
World History to 1000 A. D. + World Geography	62	68	+ 6
World History from 1000 A.D. to the Present +World Geography	41	47	+ 6
U. S. History	30	32	+ 2

Virginia Department of Education
Division of Assessment and Reporting

July 1999

About the Author

David Wheat is a consultant, educator, and writer. He is president of Wheat Resources Inc., a consulting firm established in 1981 that specializes in helping clients solve problems and choose strategies through the application of systems thinking and computer simulation. He also coaches teachers in the use of systemic learning skills in the classroom. His political science students practice those skills at Virginia Western Community College, where he is an adjunct faculty member.

His consulting work is enhanced by several years of nationally recognized classroom instruction experience in Virginia, during which time 70 percent of his Advanced Placement Government students scored high enough on national exams to earn college credit, compared to 30 percent nationwide. While a public school teacher, he served on the Governor's Commission on Champion Schools, where he participated in the upgrading of the history and government SOLs for Virginia's students.

He received his Master's Degree in Public Policy from Harvard University's Kennedy School of Government in 1972, and then served at the White House as Staff Assistant to the President, specializing in economic and energy issues. Later, he was an adjunct instructor in public policy at the University of Houston, where he also served as Director of Federal Relations. He is a member of the Association for Supervision & Curriculum Development and the System Dynamics Society.

Other reports prepared for the Jefferson Institute, where he is Senior Fellow, include:

- *Y2K Priorities for K-12* (October 1999)
- *Deficient Diplomas: Is It Time for A Graduate Warranty Program?* (September 1998)
- *Local Perspective in a State Office: The Legislator's Dilemma* (March 1998)
- *Raising Student Attendance: Some Low Cost Strategies* (March 1998)
- *Car Tax Cuts: How Should Localities be Reimbursed?* (February 1998)
- *2000 New Teachers: Where Are They Needed Most?* (February 1998)
- *Understanding Virginia's Report Card: Why Standardized Test Scores Vary from One Community to Another* (November 1997)

Copies of each are available from the Jefferson Institute (703-440-9447) or from the author (540-966-5939). Copies may be downloaded at **www.thomasjeffersoninst.org** and at **www.wheatresources.com**.

END NOTES

[1] This paper is adapted from *Value-Added Accountability: A Systems Solution to the School Accreditation Problem*, published by the Thomas Jefferson Institute for Public Policy (Feb 2000: Springfield, VA). The web site URL for the Jefferson Institute is **www.thomasjeffersoninst.org**.

[2] The author can be contacted by email (dwheat@wheatresources.com), telephone (toll free: 888-667-8850), or fax (540-966-5167). His firm's web site URL is **www.wheatresources.com**.

[3] On the contrary, *the 1999 improvements were probably discounted* by some who believed that the higher scores included one-time, quantum adjustments by teachers and students in terms of instructional strategies and test-taking skills, and that such adjustments would be incremental in later years.

[4] The Virginia Association of Elementary School Principals, the Virginia Association of School Superintendents, the Virginia Association of Secondary School Principals, the Virginia Congress of Parents and Teachers, the Virginia Education Association, and the Virginia School Counselor Association, and the Virginia Counselors Association.

[5] Virginia Department of Education (VDOE) press release, August 13, 1999.

[6] The Fordham Foundation analyzed learning standards in all fifty states, and assigned these "grades" to Virginia's SOLs: History--A, English--B, Math--B, Science--D, and Geography--D. For an extensive summary of the Fordham findings, see the appendix in *Y2K Priorities for K-12*, D. Wheat (Thomas Jefferson Institute, Oct. 1999).

[7] Based on the author's involvement in that process, the significance accorded these intermediate factors varied considerably among participants. When the SOLs were published, unaccompanied by either resource materials or implementation strategies, many teachers and administrators wondered whether we had given *any* consideration to the intermediate factors.

[8] Within the circle labeled "learning productivity" in Figure 4 is a complex sub-system of teacher and student capabilities and motivation and instructional resources. If it were necessary to specify the structural components of that sub-system, we could "zoom" down to that level and do so. However, while extremely important to student learning rates, the interaction of those factors is not something within the sphere of direct influence by the Board of Education's accreditation policies. Therefore, going to that level of detail is is not necessary for our purpose; nor is it desirable due to the complexity it would add without corresponding gains in understanding of the system dynamics of interest in this study.

[9] Computer simulation of the SOL model (downloadable at **www.wheatresources**) requires many more details than necessary for public officials to understand the basic dynamics of the system described here.

[10] The "starting point" in reading such diagrams can be puzzling at times because one presumes that a "correct" starting point exists. In principle, one may start at any point and follow the logic around the loops. In practice, however, it is a good idea to start where a change has taken place; in this case, raising of the standards by the SOLs.

[11] On the other hand, falling political support would increase the pressure for a policy change and reduce the impact of the SOLs, by undermining their classroom credibility, or diluting or rescinding them by regulation or statute..

[12] Productivity growth, however, cannot be expected to increase forever. Eventually, the learning capacities of students, the instructional skills of teachers, or the effectiveness of instructional resource materials--or all three--will grow at a slower rate or not at all. When any of these limits begins to curtail productivity and learning rates, the system will generate pressures for changes in policy. (In principle, there is also an absolute upper limit to test scores, although that is unlikely to become a bbinding constraint.)

[13] A persistent Gap--whether growing or not--may be just as damaging to public support. Social phenomena are often simplified as "win or lose" outcomes, and the persistent failure of large numbers of schools to measure up to expectations despite actual improvements in performance will undoubtedly affect public evaluation of the schools *or the standards that make schools appear to be losers*.

[14] The net effects cannot be measured directly , but can be inferred as the system operates in real life. Also, different assumptions about the relative strength of the feedback loops can be investigated by computer simulation.

[15] For high school end-of-course tests, the percentage for a subject area (e.g., math) is the unweighted average of the passing rates for tests in that area (e.g., algebra, algebra II, and

[16] In general terms, the "80/20 Rule" says that a large percentage (e.g., 80%) of significant events will occur in a small percentage (e.g., 20%) of the possible cases. For example, a large majority of a company's profits will come from a small minority of its customers. In schools, a large majority of the disciplinary problems will be caused by a small minority of the students. Widely acknowledged in business management literature for its relevance, the concept was developed by Italian economist Vilfredo Pareto in the late 19th century and *receives in-depth* treatment in The 80/20 Principle by Richard Koch (Doubleday: 1998).

[17] Individual school SOL test results were obtained from the August 13, 1999 VDOE press release, while the county school division data were obtained from their individual web sites on the internet.

[18] The 27 and 35 percent figures refer to the percentage of students reaching their own goal: passing all four tests. The 54 and 62 percent figures refer to the average of the student pass rates on the four tests (i.e, the average of the percentage passing math, the percentage passing English, the percentage passing history, and the percentage passing science). See Appendix A

[19] Another issue is whether current SOL tests are accurately assessing the knowledge and skills that students have gained from the SOLs. That topic will be the subject of a future Jefferson Institute study.

[20] A cohort consists of students at the same grade level in a particular school. In localities with high student turnover, the policy would have to operate in a way that minimized distortions resulting from mere changes in the composition of cohorts.

[21] To avoid incentives for "low" 3rd grade scores that would inflate "improvement" at the 5th grade level, it would be necessary to calculate a base period 3rd grade score, using 1998 and 1999 3rd grade scores.

[22] A school could still be "graded" according to annual SOL test performance, if so desired, but such a report card should be interpreted primarily as an indicator of the "academic ability" of the current student body and secondarily as an instructional quality measure (similar to the way athletic teams are *per*ceived each year). It should not be the sole accreditation criterion.

[23] Among the adjustments are (1) redefining the criteria in elementary schools so that scores on history and science tests don't count in the accreditation determination, and (2) creating a special category of high school students "who are unlikely to pass the SOL end-of-course tests" and replacing SOL test requirements with less rigorous *literacy test requirements, and passing the literacy tests will count toward the schools' pass rates on SOL tests!* See Pamela Stallsmith's article,"Combined SOL Scores Endorsed/VEA Cheers Board's Decision" in the Metro Section of the April 27, 2000 issue of the *Richmond Times-Dispatch*.