



## How State Assessments Lead to Vacuous Thinking and Writing

GEORGE HILLOCKS, JR.

*The University of Chicago*

Most states, in their statements about goals for education in the schools, talk about the need for students to learn to be critical thinkers. As part of their writing assessments, most states ask student writers to produce some sort of persuasive writing which the states usually see as involving critical thinking. If we believe that states truly are interested in students' thinking critically, then one would expect to find that the persuasive writing in the testing programs would reflect the concern with critical thinking. An analysis of the test formats (including prompt, test settings, and time available); the criteria for judging the results; and the benchmark papers that demonstrate what the criteria really mean indicates that states are not at all much concerned with critical thinking. Rather, most of the assessments examined appear designed to elicit and reward shoddy, vacuous thinking.

**A**t present, 48 of the 50 states administer some form of writing assessment. In most of them, students may be called on to write any one of three to five kinds of writing. The most commonly occurring of these is generally called *persuasive writing*. What the term *persuasive writing* means varies widely from state to state. The variation is the result of the combination of three variables: the

---

**George Hillocks, Jr.**, Professor, Department of English Language and Literature, University of Chicago, is author of *The Testing Trap: How State Writing Assessments Control Learning* (2002) and several other books including *Research on Written Composition: New Directions for Teaching* (1986), *Teaching Writing as Reflective Practice* (1995), which won the NCTE Russell award for distinguished research in 1997, and *Ways of Thinking, Ways of Teaching* (1999). He served as the Thomas R. Watson Distinguished Visiting Professor of Rhetoric and Composition at the University of Louisville in 2000. His new book on teaching narrative writing in urban schools is in progress.

---

Direct all correspondence to: George Hillocks, Jr., School of Education, University of Chicago, 5835 Kimbark Ave., Chicago, IL 60637-1684.

---

test format (including prompt, test setting, and time available); the criteria (or scoring rubrics) for judging the results; and the benchmark papers that demonstrate what the criteria really mean.

I concentrate on persuasive writing because it is the one place at which most curricula call for some level of critical thinking or reasoning, perhaps the most valuable process we can teach in the schools. It is arguably essential to promoting a democratic community committed to social justice and ethnic and racial harmony. As Martha Nussbaum (1977), the 2001 Grawemeyer Award winner, states in her award winning book, *Cultivating Humanity*, “Failure to think critically produces a democracy in which people talk at one another but never have a genuine dialogue. In such an atmosphere, bad arguments pass for good arguments, and prejudice can all too easily masquerade as reason. To unmask prejudice and to secure justice, we need argument, an essential tool of civic freedom” (p. 19).

State learning standards appear to recognize the importance of critical thinking and, at least, make some allusion to it. New York, for example, appears to place high priority on critical thinking:

Goal 1: Each student will master communication and computation skills as a foundation to:

1.1 Think logically and creatively

1.2 Apply reasoning skills to issues and problems

Goal 2: Each student will be able to apply methods of inquiry and knowledge learned through the following disciplines and use the methods and knowledge in interdisciplinary applications. [English language arts is the first listed of these disciplines.] (New York State Education Department, 1994, pp. 59-61)

Even the Illinois standards, which are less explicit about thinking, mention the importance of problem solving in a prefatory statement to the standards called “Applications of Learning.” It states that students should “Recognize and investigate problems; formulate and propose solutions supported by reason and evidence” (Illinois State Board of Education, 1996, pp. 1-2). The actual learning goals or standards make no allusion to reason and evidence, not even in the brief discussions of the late high school “benchmarks of learning” (pp. 4-23). The closest these statements come to explaining what is meant by “reasoning” is a benchmark statement for late high school writing listed as 5.B.5: “Evaluate the usefulness of information; synthesize information to support a thesis; and present information in a logical manner in oral and written forms as individuals and members of a group” (p. 21; online version omits the final seven words). The problem with such statements, however well intended, is that they do not explain what will count for “evaluation,” “support,” or “logical manner.” I argue here that knowing what counts as support, evidence, logic, and reason results in very important differences in what students learn.

## What Counts as Logic and Reason

Before turning to the arguments supporting that claim, it is useful to examine what counts for *logic* and *reason*. Nussbaum sees Socratic reasoning as the major tool for helping students learn to examine their concepts and assumptions. For Socrates and the Stoic philosophers, all people have the capacity for practical reason but tend to lead somnolent lives accepting traditions, norms, and beliefs learned from infancy without questioning, without taking charge of their own thinking. Socrates went about questioning everyone, about courage in the *Laches*, about friendship in the *Lysis*, about self-restraint in the *Charmides*, about piety in the *Euthyphro*. One point of these dialogues is that inadequate understanding of the concepts involved can lead to inappropriate judgments and actions. Socrates attempts to help the characters clarify the concepts so that they may make better, more appropriate decisions, but they frequently run from him after a bit of Socratic questioning.

In the *Euthyphro*, for example, Euthyphro tells Socrates that he is present in order to bring a suit against his own father for murder. Upon questioning, Euthyphro explains that his father discovered that “a hired workman” of his became drunk and then angry with one of the household slaves and “butchered him.”

So my father bound him hand and foot, threw him into a ditch, and sent a man here to Athens to ask the religious adviser what he ought to do. In the meantime he paid no attention to the man as he lay there bound, and neglected him, thinking that he was a murderer and that it did not matter if he were to die. And that is just what happened to him. For he died of hunger and cold and his bonds before the messenger came back from the adviser.

These statements of Euthyphro are never in dispute. Euthyphro explains that now his father and the rest of his relatives are angry with him for prosecuting his father for murder, which, he claims, “shows how little they know what the divine law is in regard to holiness and unholiness” (p. 15).

Socrates immediately asks Euthyphro,

do you think your knowledge about divine laws and holiness and unholiness is so exact that, when the facts are as you say, you are not afraid of doing something unholy yourself in prosecuting your father for murder? (p. 17)

From this point, the dialogue develops as an inquiry into the nature of holiness and unholiness with Socrates insisting that Euthyphro “tell the essential aspect by which all holy acts are holy.” Euthyphro makes several attempts, proposing first that “holiness is doing what I am now doing.” Socrates points out that his action is only one instance and not the essential aspect of holiness. Euthyphro claims that Holiness is what the gods love. Socrates makes the point that because the Gods quarrel, they do not all love the same things. It turns out, of course, that Euthyphro cannot define holiness at all. Finally, he exclaims, “But Socrates, I do not know how

to say what I mean. For whatever statement we advance, it moves about and won't stay where we put it," revealing, of course, the weakness of Euthyphro's case against his father.

Euthyphro's initial claim is that he is justified in prosecuting his father for murder. Socrates' original question to Euthyphro concerns how his action against his father can be justified. If it is to be justified as a pious act, then Euthyphro must show that the act he proposes is a pious one. To do that, he must be able to tell what holiness is and the sense in which his prosecution is holy. The definition of holiness would tie the facts of the case to Euthyphro's proposed action, the claim, by explaining how the action is justified. That is, the definition of piety or holiness would warrant the claim in light of the evidence. Socratic inquiry is always directed to such warrants and the definition of their underlying concepts or assumptions, their backing.

As we listen to the arguments about whether the United States should attack Iraq, we realize that there is more to critical inquiry than the analysis of underlying assumptions, important as that is. Aristotle's *Rhetoric* helps us here. In Aristotelian terms, the main argument about what Euthyphro should do is one of deliberation, an argument of policy the result of which no one can predict with any certainty. For Aristotle, it was important to bring to bear on policy as many perspectives as possible because of that unpredictability. That is why it is called argument of deliberation. Still, the argument must be justified (warranted) by some principle that all involved can accept.

Underlying such arguments are the arguments about whether any one course of proposed action is a good one, in this case, whether it is holy. (In our day, we might argue whether it is just.) The Aristotelian term for these arguments is *epideictic*, arguments of praise or blame, that may be called, more generally, arguments of judgment. These are always dependent on definitions for the warrants that tie evidence to the claims. That is, we cannot claim that an action is holy or just without careful definitions of what is holy or what is just. We cannot argue that a movie or book is a good one without some definition of a good movie or a good book.

The facts in Euthyphro's case, what happened to the slave, the workman, and what the father did are never in dispute. If they were, they would be involved in Aristotle's third type of argument, the forensic, arguments about the facts of a case. Although these are ignored by Plato, they are no less important to arguments of judgment and policy.

Toulmin (2001) argued that now, at the beginning of the 21st century, we are beginning to see a return to reason, a return to argument. It is no accident that he entitled his 1958 book, *The Uses of Argument*, for he recognized that what he was proposing was far different from traditional or even modern logic. He proposes that when we deal in probabilities, we make a claim, based on grounds or evidence that might be drawn from many diverse sources. Although Toulmin does not deal explicitly with the nature of grounds, it is obvious from his examples that evidence, to be useful, must be relevant and verifiable. In different disciplines and fields of work, special procedures must be followed so that evidence will not be impeached (e.g., the collecting of blood evidence in the Simpson case or the qualitative analysis of unknown samples in chemistry). But basic to any kind of argument is the verifiability of the evidence. A literary critic must cite the works discussed and quote

from the texts to prove a claim. An historian must carefully note the artifactual or documentary evidence basic to the argument being made. A scientist must explain the nature of observations or experiments, the collection of data, and so on, so that the study can be replicated.

Additionally, in many cases, an argument requires stating explicitly how or why the evidence does indeed support the claim. Toulmin calls this part of the argument the *warrant*. A warrant is generally a rule that the arguer can count on being acceptable to the audience. As an activity for teaching Chicago high school students to write argument, one of my students at The University of Chicago invented a problem scenario including a sketch of a victim hanging from a chandelier, his feet dangling a distance above a stool on which he had presumably stood to do the deed. The sketch is accompanied by a note explaining reasons for committing suicide. Students were encouraged to examine the evidence of the sketch and the note to determine what had occurred. I have simplified the scenario a bit, but when students began their discussion in small groups, they attended to the note and seemed to examine the picture only cursorily. Before long, however, one young man proclaimed to his group that it could not be a case of suicide. "Look where his feet are," he explained. "If he hung himself, his feet would've been below the top of the stool. They're not. They're way above it." The student had hit on an important warrant. He explained it somewhat as follows after a few tries, "When a person hangs himself, he has to drop from some height so that the noose will tighten and strangle him." He followed up with stating the evidence, "See, look where his feet are, a couple of feet above the stool. He could not have jumped up, fastened the rope, put the noose around his neck, and hung himself." These were statements of evidence and a warrant that the class could accept.

When Euthyphro argues that his prosecution of his father is justified, his warrant is that what he does is holy because "holiness is doing what I am now doing, prosecuting the wrongdoer who commits murder or steals from the temples . . . whether he be your father or mother or anyone else, and not prosecuting him is unholy" (Plato, p. 19). Perhaps others in Euthyphro's community might accept that as appropriate. Indeed, it is reminiscent of the sort of warrants we hear in modern American political speeches, statements about the American way, our love of freedom, and a host of other glittering generalities. Unlike most Americans, Socrates does not accept Euthyphro's cliché. If Euthyphro were able to supply the definitions of the terms that Socrates seeks, they would constitute what Toulmin calls the *backing* for the warrant. Backing for warrants might include extensive definitions, references to laws, explanations and justifications of rules, arguments, in short, in support of the warrants. In explaining the impossibility of self-hanging from a low stool, the young man was providing backing for his warrant.

Toulmin's view of the nature of argument seems to me to be a useful tool for examining the nature of logic and reasoning we find in state learning standards and assessments. Although no one that I know makes use of classical logic beyond an outline of logical fallacies, a number of textbooks for college-level composition have adopted the Toulmin system. For example, a text by Ramage and Bean (1992) seems to be used extensively. The system of argument of logic works for all three kinds of arguments as defined by Aristotle in the *Rhetoric*, all of which Aristotle

clearly recognizes as arguments of probability, not susceptible to syllogistic reasoning. Furthermore, the kinds of persuasive assignments that are given on state examinations all involve arguments of probability and never any that are amenable to syllogistic reasoning. I turn then to the state assessments themselves.

### The Nature of State Writing Assessments

Several years ago, I began a study, sponsored by the Spencer Foundation, of state mandatory writing assessments that examines the assessments in five states selected for differences in their assessment formats, procedures, and stakes for students, teachers, and schools. My research assistants and I made in-depth studies of assessments in Illinois, Kentucky, New York, Oregon, and Texas. Our study involved more than 350, including teachers and administrators in six districts in each state and state-level officials responsible for the assessments including state superintendents in two of the states. We examined teacher materials for teaching writing, published materials purporting to be appropriate preparation for the assessments, and relevant state documents, which were bountifully present in most of our states. Additionally, we examined assessments for most of the states that have them, at this writing, 48 of 50,

### Variation in State Assessments

The five state assessments that constitute the focus of the study differ widely. The tests themselves differ in the kind of prompt for writing, the actual writing situation, and the criteria and benchmark papers for judging the writing. In Illinois, students have 40 minutes to write in response to a prompt that provides no information about the topic. Similar prompts appear in Texas and Oregon, although the kinds of writing differ. In New York, students at the 11th grade write four essays, and the prompts differ from those in other states in that, for three of the essays, the test provides texts for the students to write about. The fourth essay requires that students recall two literary works that they have read. In Kentucky, the assessment requires a portfolio of several pieces of writing of different types, some of which must have their genesis in subject areas other than language arts. Students select their own topics and have the year or more to develop their several pieces.

The criteria indicating what counts as a successful piece of writing also differ. For persuasive writing, Illinois and Texas call for elaboration and support, for example, whereas the other three states call for elaboration, but for persuasive writing and for writing about literature, New York stipulates that students include evidence. A careful analysis of the benchmark papers in Texas and Illinois, on the other hand, indicates that no evidence is necessary to achieve the highest scores. In fact, few of the benchmark papers at the highest levels contain any evidence at all. These differences have serious consequences for how teachers teach and what students learn about writing and thinking from state to state. In New York and Kentucky, students are likely to learn that evidence is important to making a case, whereas in Illinois and Texas, they are likely to learn how to avoid the inclusion of evidence,

indeed, to learn that it is unnecessary. The problem is that most states do not have assessments similar to those in New York or Kentucky. Most are similar to those in Illinois and Texas.

### Writing Prompts

In states with assessments like those in Illinois and Texas, for persuasive writing, students receive a prompt focused on an issue about which they may or may not be informed. Neither teachers nor students know in advance what the prompt will be. Usually, the prompt sets forth a problem and asks students to address the problem by taking a position and convincing some audience that the position is a good one. Sometimes, the problem is one of supporting a judgment, an epideictic argument. For example, the Illinois writing guide, *Write On, Illinois!* presents two prompts concerning who should receive the relative-of-the-year award and what is the best place to live. More often, the prompt presents an issue related to some policy: required community service for high school graduates, school uniforms, a tutorial program serviced by high school students to help elementary students, or an extension of the school year into the summer months. These prompts present no information about the issues, no data, and no other help. Here is an example from New Jersey:

#### WRITING SITUATION

This year your school's soccer team won the state championship. When your team won, students ripped up pieces of the soccer field and cut the goal nets to keep as souvenirs of the game. As a result, the school principal has announced that any money raised this year by all the school clubs will go toward repairing the playing field and replacing the soccer equipment.

The principal's decision has created a controversy in your school. You decide to write to the editor of your school newspaper about this decision.

#### WRITING TASK

Write a letter to the editor of your school newspaper. Explain your views on the principal's decision requiring that money raised by school clubs be used to repair the field and replace soccer equipment. Use examples, facts, and other evidence to support your point of view. (New Jersey Department of Education, 2000, p. 22).

The following example is from Texas:

Some people believe that all teenagers should be required to perform one year of unpaid service for their community right after they graduate from high school. This community service might include helping to clean up parks, delivering food to the elderly, or working in a hospital.

What is your position concerning this issue? Write a letter to your senator in which you state your position and support it with convincing reasons. Be sure to explain your reasons fully. (Texas Education Agency [TEA], 1993, p. G5)

In both of these examples and in many others from other states, students are required to support their positions but are given no facts or opinions related to the problem. Such prompts appear to indicate that the legislatures or administrators who put them in place prize some ability to make a case, to present evidence in support of a position. This is an admirable goal.

The question is whether that goal is adequate to what future citizens need to know. Certainly, we can all agree with Nussbaum's (1997) cogent arguments that the most troublesome issues facing all Americans now and in the foreseeable future are those of human dignity, equity, and justice. In an increasingly diverse society, our judgments of others are often stated as truths and acted on as though they were. The assumptions underlying those judgments, like those of Euthyphro, often remain unexamined. Making informed judgments and choices requires a profound understanding of the human condition and the principles by which we operate within our democratic systems. That profound understanding entails empathy and compassion, but in addition, it entails the ability to identify, explicate, evaluate, and to reject or modify the assumptions underlying issues, judgements, and policies. That, in turn, requires an ability to think through a case, to question evidence, to determine what suffices as evidence for a given case, and how that evidence is tied to the claim. In short, it requires the abilities to construct and deconstruct arguments. Although many states appear to hope for the development of such thinking capacities in their schools, their writing prompts work against it.

Consider this. Could any of you reading this article write a cogent argument for or against all teens performing mandatory community service following high school? Could you write a convincing argument supporting or opposing the use of school uniforms, or the extension of the school year? Could you do it in 40 minutes as required in Illinois? Or in 1 day, as in Texas? For most of you, the answer would be no, not because you are unaware of how to develop a convincing case, but because without relevant data, you would have little or nothing to say besides general opinions, which may or may not be convincing, but which would certainly be open to attack.

### Criteria for Evaluating Writing

In order to examine the tests, we have to look at more than the prompts. The criteria for judging the writing and the benchmark papers that illustrate the various score levels are of crucial importance. In Texas, for example, teachers told us in many interviews that elaboration was the single most important criterion for passing the test at all grade levels. One supervisor told me that she commonly walked about in a classroom while students were writing during the exam and encouraged students to write more, an illegal practice in Texas. The most interesting thing is that she did not call for high-quality additions, just additions. Research by Bereiter and Scardamalia (1982) concluded that students wrote in conversational bursts, about as much as one would say in a conversational turn, that they did not exhaust the information available to them in one of these bursts, and that if encouraged to say more, they would add another burst of about the same length. Depending on what elaboration means, as is seen here, simply adding more stuff is likely to push up the score assigned to the paper.



In Texas and Illinois and in several other states, elaboration is the key criterion. Look for a moment at the Texas criterion for elaboration. The Texas scoring guides devote the most space to discussions of elaboration. They tell us the following:

An essential part of successful writing is the effective use of support or elaboration, which requires the student to develop clearly, logically, and completely those ideas that lead the reader toward an understanding of the writer's purpose in the response. The degree to which support or elaboration is achieved, therefore, is dependent not only on the student's ability to generate ideas, but also upon the extent to which the student provides the reader with a detailed explanation of those ideas. . . . The more thoroughly and precisely each idea is developed with specific details, the stronger and more complete the support or elaboration will likely be. (TEA, 1993, p. 3)

Although we can hardly disagree with this statement, it is necessary to see what actually counts for the “detailed explanation of . . . ideas.” Fortunately, the guide presents examples of reasons given at different levels of elaboration. The examples are based on a response to the “exit-level” prompt about community service. The wording of the prompt is important and is reflected in the analysis of what counts for elaboration: “convincing reasons” and “explanations of the reasons.”

The guide provides a series of reasons from “bare reason” to what the guide calls “fully elaborated.” Let us examine what the Texas guide calls the “fully elaborated reason.”

I am against the idea of making all high school graduates serve the community because it will disrupt and interfere with our plans to go to college or work. *Students who plan to go to college will forget most of what they have learned if they spend a year away from school. Their chances to achieve a high GPA will be hamstrung. As far as plans to work go, some graduates, like me, plan to work for necessities. I have to pay \$350 a month car and insurance payments on my Nissan 280 ZX. Making your payments on your first loan is important so you can establish good credit. Also, some people have to help their parents with rent and groceries, and medical bills.* (TEA, 1993, p. G6)

To analyze this, let us return to the Toulmin (1958) model of reasoning and argument. When we talk about an argument that something should or should not be done, we may have a major claim about what to do and several reasons for doing it or not doing it. The reasons may be called subclaims, which themselves require some kind of support. If, for example, a teenage girl is arguing to a parent the need for buying a used car, she might claim that she should buy a particular 1996 Camry because it is in good condition, it is priced appropriately for the budget available, the Camry has a reputation for low rates of mechanical failure, and she can expect to get 5 or more years of use from it. Each of the four reasons is, itself, a claim that requires support. When we provide support for the subclaims in an argument, we usually provide evidence, hard data about the “facts of a case” that are purported to be true. Evidence that our used Camry is in good condition might include the

depth of tread on the tires, the mileage on the car, the sound of the engine as it starts and idles, an inspection of the body for past accidents, and the results of an inspection by a trained mechanic.

Sometimes, such an argument also includes statements that serve to explain why data support a claim. These are general rules that parties in a discussion can accept as true or at least are willing to agree to tentatively. One indirect piece of evidence that the Camry is in good condition is that the current owner is a 67-year-old grandmother. With such evidence we almost always need to explain why it supports the claim. In this case, it is the general rule, our belief that older women tend to drive carefully and take good care of their cars. Such a statement is called a warrant because it explains why data are appropriate in support of a claim. That is, it warrants the data as support for the claim.

If we use the framework of major claims, subclaims, evidence, and warrants, we can analyze what Texas calls a “fully elaborated reason” as follows:

Main Proposition: I am against the idea of making all high school graduates serve the community because it will disrupt and interfere with our plans to go to college or work.

Subclaim 1: Students who plan to go to college will forget most of what they have learned if they spend a year away from school. (Therefore, the writer implies)

Subclaim 2: Their chances to achieve a high GPA will be hamstrung.

Subclaim 3: As far as plans to work go, some graduates, like me, plan to work for necessities.

Evidence: I have to pay \$350 a month car and insurance payments on my Nissan 280 ZX.

Subclaim 4: Making your payments on your first loan is important so you can establish good credit. (If this statement tied the evidence to subclaim 3 about having to pay for necessities, it would function as a warrant. However, subclaim 3 does not deal with establishing credit.)

Subclaim 5: Also, some people have to help their parents with rent and groceries, and medical bills.

This analysis indicates that the subclaims remain unsupported, which in turn leaves the major claim open to serious question. The strongest part of the argument is subclaim 3, which is followed by evidence of a sort and a warrant that helps to establish the need to pay bills for what the writer calls necessities. (It is, of course, problematic that the Nissan 280 ZX is a necessity.) But that is clearly the closest that the elaboration comes to presenting evidence and a warrant. Subclaims 1, 2, and 5 remain totally without support. This scoring guide example supports teacher beliefs that “adding more stuff” will result in a higher score.

What is disturbing about this example, provided so that teachers will understand what is meant by elaboration, is that it is likely to become the standard toward which teachers push. The example says, in effect, that evidence is not necessary. When such examples are presented as exemplary, they effectively lower the stan-

dard for persuasive writing, even, I suspect, when the persuasive writing in classrooms is not for the writing tests.

Here is a top-scoring paper from Texas written in response to the same prompt: I present it with my analysis.

Main proposition: A full year of community service would . . . have more of a negative than positive effect on the individual and community.

Subclaim 1: After graduation, many ambitious individuals plan to attend college or university.

(Subclaims 1a to 1c prepare for subclaim 2.)

Subclaim 1a: One's first semester would begin a short 3 months away, ample time to get all the necessary supplies for college.

Subclaim 1b: In this 3-month time period, one has just enough time to rest from his or her senior year of high school,

Subclaim 1c: but it is not so long that he or she is absent from education for such a length of time that he or she loses some of the valuable material retained in high school.

Subclaim 2: A full year is a long span of time that is precious to one's education.

Evidence: I, from personal experience, know how easy it is to forget material over the summer, and that is only a 3-month period. (Note that this evidence contradicts subclaim 1c.)

Subclaim 2 (reiterated): A year would be entirely too long to be absent from schooling.

Subclaim 3: The community service would also be devastating to an athlete planning to participate in college.

Subclaim 3 (reiterated): The time would greatly effect one's status as an athlete and may hurt one's chances of competing on a collegiate level.

The *scoring guide* comment on this paper reads, in part, "This fluent, well-organized response provides convincing evidence that the required community service would be detrimental to those students with educational or career plans." Just where is all the convincing evidence? And this, remember, is a top-rated paper.

In Texas, the top-rated papers are two levels above passing. This is important to consider when President Bush brags about excellence in Texas. His comments are about schools and the percent of students reaching passing levels on the tests. Just how excellent are compositions judged to be at the level of passing, a score of 2? Let us examine an example.

As you will recall, the prompt contains no information about community service, the legality of such a program, or data about real or fictional community service projects. Any student responding to this prompt must pull all reasons and explanations from his or her own experience. The prompt, however, calls only for reasons and explanations, not evidence, even though the scoring guide comments state that the pieces include evidence.

Dear Senator,

As a teenager about to graduate from high school, I think it is rather unfair to do these services without being paid for it. Therefore, I believe we shouldn't have to do these services right when we get out of high school.

First of all, when people graduate from high school, a majority of the people will either go to a junior college or college. During the summer break, most of them will get jobs to help pay for college. Another reason is, it takes money to drive around town and do these services.

Personally, I think you all should use all of the unemployed people that receive unemployment checks because they're the ones that have nothing to do.

These are the reasons why I think we shouldn't have to do these services.

Sincerely, (TEA, 1993, p. G14a)

The first reason given for opposing the community service proposal is that "it is rather unfair to do these services without being paid for it." Focusing on pay evades the problem of requiring service of teenagers, which given required school attendance and the military drafts of the past is not automatically dismissible. Additionally, there is no argument as to the unfairness of not being paid. The writer must assume that his assertion of unfairness is true *prima facie*. This is probably the strongest reason that the writer presents, but it remains unsupported and unexplained. The second reason presented is that "when people graduate from high school, a majority of the people will either go to a junior college or college." If there were a law requiring that high school graduates had to provide full-time community service during their first year out of high school, it would be challenged not because "the people will either go to a junior college or college," but because it is considered an infringement of human rights. There will be arguments on the other side as well because it is already the case that some schools require community service for graduation. The final reason is that "it takes money to drive around town and do these services." This reason assumes that the service will require personal transportation and that expenses will not be paid. However, it simply fails to address the central problem. The response is shallow, totally devoid of evidence, and certainly fails to persuade anyone of anything.

The Texas scoring guide commentary on the paper was as follows:

This controlled, organized response takes a clear position against requiring community service. The section discussing the necessity of working for pay is somewhat elaborated while the solution adds elaboration by offering a ready alternative (the unemployed) to employing high school graduates. In total, a minimally sufficient amount of evidence is provided, and the response demonstrates minimal success with the persuasive task. (p. G14a)

The scoring guide does not contain any explanation of what counts as evidence. So there can be no explication of the phrase, "minimally sufficient amount of evidence." Although the commentary praises the "ready alternative (the unemployed)," the suggestion simply evades the central question. Is it just or appropriate to select any defined group of people in the community and require that they do

anything? Suggesting the use of the unemployed implies that the idea of conscription for unpaid work is just and appropriate, the very idea that the writer opposes. My Scottish grandmother called this kind of talk “blether,” talk that goes on and on to no purpose.

This sample of student writing and its commentary indicate that Texas accepts a very low level of performance as passing. What we really have is a badly stated proposition supported with three generalizations that are themselves poorly explained and unsupported, along with the statement of an alternative that serves to undercut the writer’s main contention. But it meets Mr. Bush’s standards for excellence.

Although I have drawn all of these examples from Texas, it would be an easy matter to draw many others from Illinois and many other states. The fact is that most states use prompts and criteria that are comparable to those in Illinois and Texas. Asking students to write about prompts that present issues about which they have little knowledge is almost destined to result in highly generalized writing that amounts to no more than “blether.” Furthermore, when teachers are under pressure to raise school scores, whether the real stakes are high as in Texas or low as in Illinois, teachers prepare their students to pass the tests. They can see that no real thinking is required by examining the scoring guides. They also see that it must be organized. Our interviews with teachers and administrators in Texas and Illinois revealed that very high percentages concentrated on formulaic writing, primarily the five-paragraph theme (nearly 55% in Texas and 72% in Illinois). Furthermore, interviews with administrators indicate that such formulaic writing has become district-wide policy. One elementary school principal spoke of the faculty’s consideration of bringing the rudiments of the five-paragraph theme into the kindergarten curriculum by teaching youngsters to say they have three reasons for so and so. A third-grade teacher in the same district told me that, for her class, the formula involves the use of *first, next, last* as a way of ensuring three-body paragraphs, and she said, “we pound it, pound it, pound it!” And no teachers interviewed in this district talked about teaching students how to develop the substance of an argument. Not surprisingly, in Kentucky and New York, with quite different writing assessments, few teachers (6% in Kentucky and 5% in New York) spoke about teaching the five-paragraph theme or any other formulaic structure. It seems clear that this kind of writing assessment has an immense impact on what is taught about writing and what students are allowed to learn. What students appear to be learning in states with assessments like those in Illinois and Texas is that thinking is a matter of spouting opinions, examined or not, and that any opinion is as sound as any other. Clearly, this practice works against the careful examination that Nussbaum called for as a necessity for enhancing justice and social equity in a democracy.

### Alternatives

**T**he kind of writing prompt and the attendant criteria that reward organized “blether” are not necessities of life, even in writing-on-demand testing. One alternative is the portfolio assessment, which has been very successful in

Kentucky. Another involves the presentation of data relevant to the issue, as part of the test, evidence that will allow the support of several different positions on the issue on which a prompt focuses. New York provides such material. In Texas, with virtually a full school day to write the response, such data could be a feasible part of the writing test.

Recently, at Maine West High School in a suburb of Chicago, two teachers, Seth McLowry and Sherri Koeppen, headed a program to improve writing across the curriculum. The program focused on the teaching of argument at the ninth-grade level (using the Toulmin theory outlined previously), with the key unit of instruction of 5 or 6 weeks in English classes and with in-service training of English faculty on how to understand the theory and to teach lessons that would enable students to generate real arguments with evidence, rebuttals, qualifications, and warrants. Faculty in biology and social studies were provided in-service training on the theory and how to adapt it to their writing assignments. As an advisor to the project, I recommended that they make use of a pretest and posttest that provided students with relevant material for making a case. Dr. Thomas McCann, of West Chicago High School, whose doctoral dissertation involved the teaching of argument, provided such tests. Each involved a supposed security problem, one at a high school and one at a mall. Both fictitious institutions had a security policy in place and both were recommending changes that would have an impact on the rights of teenagers. Students taking the exams were given copies of both along with tables of raw statistics, indicating changes in rates of accidents and thefts at the shopping mall over a 5-year period and changes in rates of theft, vandalism, sexual harassment, drug violations, and so on over a 5-year period for the high school. They also received a table indicating increase in mall patronage over the 5-year period and a table indicating growth of the student population over a 5-year period for the high school prompt. Finally, each prompt included a page of testimonials from a variety of concerned individuals. Care was taken to ensure that prompts and related materials were parallel in length and difficulty. For each test, students had one class period to read the materials and begin planning a response and a second day to write the response. The pretest was given at the beginning of the school year, the second at the end.

The scoring guide, which I provided, allowed raters to concentrate on the development of the arguments that writers made, focusing on the use and relevance of evidence, warrants, qualifications, rebuttals, and so forth. We used a blind rating procedure of a random sample of about half the students in each English class involved. The analysis indicated that the students had made highly significant gains (effect size over all classes was over 1.2 standard deviations). Here is an example of a posttest paper:

Dear Mr. Paul,

I am a citizen of Floodrock and I am concerned about Floodrock High School. I have heard about theft, battery, vandalism, sexual assault, and drug possession occurring in this school. I demand that some action be taken. The school's security measures are obviously not good enough if there has been a rise in crime. The school should install more cameras and hire more paraprofessionals and police officers.

I feel that there is a major problem in this school. Vandalism, theft, battery and drug possession rates have risen dramatically over the last 5 years. The rate of infraction for vandalism went from 17 to 48. The number of battery infractions has gone up from 40 to 107. Drug possession and theft have also risen over the last 5 years. These rates have gone up dramatically, but the student population has only gone up by about 150 students each year. Mr. Stan Leaver, an accounting teacher at the high school said this. "I hear kids threaten each other all the time. I've been teaching for over 26 years, & at no time have I seen conditions worse than they are now." Crime rates are going up faster than the student population rate, and teachers are even saying that the conditions at this school are worse than ever. This seems to me that there is a major problem.

Floodrock High School is also not taking enough security precautions within the school. They need to buy and install cameras in most hallways (not just in remote areas) and hire many more paraprofessionals and police officers. A drug/crime awareness group made up of students would also be helpful in cutting down on crime. More cameras and security personnel are needed to watch the children at most times. If the students know they are being watched, they are less likely to do something illegal.

Are there any problems with installing the security measures I mentioned above? Of course there is. Hiring more workers and buying new technology would cost the school and the community a lot of money that we may or may not be willing to spend. The question arises, what is more important to the good citizens of Floodrock, more money in our pockets or more security for our children?

The new security measures will benefit the school's security system tremendously. The new security plan doesn't even involve new security measures. Simply more of what is already in place.

In conclusion, I believe that we the citizens of Floodrock have an obligation to the students of the high school to provide them with a safe, educational environment. They shouldn't have to worry about getting beat up or having their valuables stolen. What we need to do is heighten the existing security measures already in place. It will cost the taxpayers some money, but it will be money well spent.

The quality of this paper is not unusually high for the posttest. Most students used some evidence to present their arguments. This writer has chosen some elements of the new security recommendations over others to support. That is clear in his emphasis on expanding the use of existent measures. He might have made that more clear and argued against the measures that he saw as undesirable. By and large, however, the paper is typical of those in the upper 20%.

Overall, the results were so positive that the argument unit and its use in other disciplines was mandated for use in the future. One key finding was that students wrote real arguments that included evidence, warrants, and counterarguments or rebuttals, not simply unsupported claims. This writing sample took longer to score than the fabled 1 minute that the commercial scoring services claim, but the score provided a

detailed account of how well students used evidence, how successful their use of warrants was, how frequently they used rebuttals, and so forth. Clearly, then it is not necessary to use writing assessments like those in Illinois and Texas, that work against the goal of learning how to think critically and argue persuasively.

### Speculations about Why Tests Are Counterproductive

**D**irectors of state writing assessments have told me that they cannot hold students responsible for content. That means that they cannot score a composition low for not presenting any data that serves as evidence in support of claims or the positions advocated. If that is true, they cannot hold students responsible for presenting evidence in support of their positions. To present a set of claims with no evidence is considerably easier than presenting real evidence in support of a claim. Perhaps that is why the Texas and Illinois scoring guides refer to unsupported claims as evidence. Perhaps state officials want to appear to be concerned with logical thinking and argument and, therefore pretend that student writing presents evidence even when it does not.

Any teacher who has asked students to write an essay that involves making a case knows how apparently reluctant students are to incorporate evidence. My students and I used the shopping mall prompt as a pretest with 10th graders in a Chicago public school last fall, and in many other fall workshops. Even though evidence is available, most students tend to avoid it. They believe that statements of their own feelings and opinions suffice as cogent arguments. And as the study of state writing assessment indicates, most have been taught that in Illinois schools. But on the posttest, most students have learned to use evidence. Thus, it cannot be simply that evidence is difficult to use that state tests tend to avoid it.

I believe that there is another factor operative here. Governors and legislators, not to mention presidents, are responsible for testing programs. If, like Bush, they make education a major plank in their platforms, they want evidence of educational progress, usually in the form of advancing test scores to show that their policies have indeed had a positive effect on the learning of students. However, they also know that people are naive and will assume that advancing test scores are actually indicative of learning. If scores do not advance, the politicians will have failed in their policies. If tests are demanding, scores will not go up so readily. Recently, when certain state exams in Illinois were made more complex, scores went down and now there is a new state superintendent. In short, it may behoove politicians and many others to keep tests easy in order to ensure that the majority of students pass and appear to be learning. Let us hope that is not really the case.

Although such weak tests of writing and thinking may suit the needs of politicians, they do not meet the needs of students and future citizens. In Nussbaum's words, argument, the kind of critical thinking involved in real persuasive writing, is "an essential tool of civic freedom" (p. 19). Tests like those in Illinois and Texas actively work against such thinking, encouraging teachers and students to believe that critical thinking and argument involve no more than asserting unsupported opinions, to believe that they have made a case when they have only blethered. I can think of no greater disservice to a democracy.



## REFERENCES

- Aristotle. (1932). *The rhetoric of Aristotle* (L. Cooper, Eds., L. Cooper, Trans.). New York: Appleton-Century-Crofts.
- Aristotle. (1964). *Prior and posterior analytics* (J. Warrington, Ed. and Trans.). London: Dent; New York: Dutton. Everyman's Library.
- Bereiter, C., & Scardamalia, M. (1982). From conversation to composition: The role of instruction in a developmental process. In R. Glaser (Ed.), *Advances in instructional psychology* (Vol. 2, pp. 1-64). Hillsdale, NJ: Erlbaum.
- Illinois State Board of Education. (1996). *Preliminary draft: Illinois academic standards for public review and comment, Vol. 1: State goals 1-10*. Springfield, IL: Author. Available online [www.isbe.net/ils/english/](http://www.isbe.net/ils/english/) with no change to material.
- New Jersey Department of Education. (2000). Grade eight proficiency assessment 2000 sample test form. In *2000 GEPA sample forms for language arts literacy and mathematics* [Online]. Available: [www.state.nj.us.njded](http://www.state.nj.us.njded).
- Nussbaum, M. (1997). *Cultivating humanity*. Cambridge, MA: Harvard University Press.
- Plato. (1999). *Euthyphro, Apology, Crito, and Phaedo* (H. N. Fowler, Trans.). Cambridge, MA; Harvard University Press
- Ramage, J. D., & Bean, J.C. (1992) *Writing arguments: A rhetoric with readings* (2nd ed.). New York: Macmillan.
- Texas Education Agency. (1993). *Exit level 1993-1994 writing collection scoring guide for persuasive writing*. Austin, TX: Author.
- Toulmin, S. (1958). *The uses of argument*. New York: Cambridge University Press.
- Toulmin, S. (2001). *Return to reason*. Cambridge, MA: Harvard University Press.