

24

MASS-MARKET WRITING ASSESSMENTS AS BULLSHIT

Les Perelman

Massachusetts Institute of Technology

In Summer 2007, one of the almost 1,000 scorers of the English Language and Composition Advanced Placement, Edward M. White (2008), wrote on the airplane home a 5-paragraph essay that reflected and responded to the hundreds of 5-paragraph essays he read while participating in the grading of 280,000 tests written by high school students. The fourth and “clincher” paragraph of this essay reads:

The last reason to write this way is the most important. Once you have it down, you can use it for practically anything. Does God exist? Well you can say yes and give three reasons, or no and give three different reasons. It doesn't really matter. You're sure to get a good grade whatever you pick to put into the formula. And that's the real reason for education, to get those good grades without thinking too much and using up too much time. (p. 525)

Such an approach to truth goes back at least to the Sophists, and the attitude expressed by White's student's direct ancestor, Polus, who in Plato's *Gorgias* states that a rhetorician need have no special knowledge of a specific subject when persuading people about that subject (460b-460c) (Plato, 2004).

Socrates, of course, attacks this position, and much more recently, the contemporary philosopher Harry G. Frankfurt (2004) has attempted to develop a fairly rigorous philosophical definition of what these students are doing in his seminal monograph *On Bullshit* and his response (Frankfurt, 2002) to Cohen's (2002) essay "Deeper into Bullshit." As Frankfurt (2004) begins his monograph, "One of the most salient features of our culture is that there is so much bullshit." (p. 1). As Frankfurt argues, bullshit is actually more dangerous than outright lying because while the liar knows the truth, although he wants to lead his audience away from it, the bullshitter is unconcerned with the truth, "but that the motive guiding and controlling it is unconcerned with how the things of which he speaks truly are" (p. 55).

Bullshit is dangerous. Eubanks and Schaeffer (2008), however, claim in their essay that some forms of unprototypical academic bullshit may "be both unavoidable and beneficial" (p. 372) (but they never fully support that assertion in their essay nor do they refer to Cohen's extension of Frankfurt's definition [which bears some similarity to their own analysis] nor of Frankfurt's reply to Cohen). Similarly, Smagorinsky, Daigle, O'Donnell Allen, and Bynum (2010) argue based on the protocol analysis of one essay by one student that "bullshitting can serve as a key developmental tool in its promotion of exploratory thinking and speech through which learners may approach tasks at new levels of complexity" (p. 401). Both Tietge (2006) and Fredal (2011) present thoughtful analyses of the relationship of rhetoric to bullshit.

The examples of harmful bullshit, however, are everywhere. One only has to look at the historical examples in Tuchman's (1984) book, *The March of Folly*. A more contemporary example can be found in the rationales for the American Invasion of Iraq in 2002-03. Significantly for this essay, the George W. Bush Administration had exactly three supporting arguments to justify the Invasion of Iraq: (a) Saddam Hussein had Weapons of Mass Destruction; (b) Saddam Hussein supported Osama bin Laden and Al Qaida; and (c) Saddam Hussein brutally repressed his own people. (Someone in the White House must have done well on the SAT Writing Test.) The first assertion has been proven false by postwar inspections (S. Rep. No. 108-301, 2004). There is no real evidence to support the second assertion and some evidence to refute it (S. Rep. No. 109-331, 2006). The third assertion is clearly true, but it is equally true for many other countries that the United States did not invade. Thus we can classify two of the three reasons for the United States invading Iraq as bullshit. It is just not that these statements are false. Truth or falsity does not determine bullshit, but rather the bullshitter's intention to be unconcerned with truth or falsehood. There may have been some people in the White House who knew that Saddam Hussein did not possess nuclear arms capability, but for most of them, they just accepted the statement and did not want to hear any evidence to the contrary (Wilson, 2004).

Bullshit does not only harm public affairs; it is bad for business and manufacturing. Elsewhere, I have documented how the explosion of Space Shuttle *Challenger* was caused by the management of Morton-Thiokol ignoring the categories of truth or falsehood, opting instead to “put on their management hats,” which meant essentially that they were supposed to be unconcerned with the probability of the O-rings failing in the extremely cold weather and be concerned with the economic profits of the company (Perelman, 1994).

Education should be the enemy of bullshit. Although there may, in some disciplines, be multiple truths, the intention of educated discourse, be it for academic or nonacademic purposes, is to represent at least one truth. Yet mass-market testing practices and organizations, especially when assessing writing, have whole-heartedly embraced bullshit.¹ In the remainder of this chapter, I demonstrate how testing organizations within the area of writing assessment both encourage and practice bullshitting in three distinct ways. First, the timed impromptu essay not only invites students to bullshit, it encourages the practice. Second, the distortion of holistic scoring practiced by most mass-market organizations, with its reliance on conformity and reliability at the cost of ignoring intellectual content is, itself, a form of bullshit. Finally, the research conducted by these testing organizations often produces reports containing substantial amounts of bullshit.

In 2005, I discovered that several of the top sample student essays in the booklet handed out by the College Board at the 2005 meeting of the Conference on College Composition and Communication, *ScoreWrite™: A Guide to Preparing for the New SAT® Essay* (The College Board, 2005), contained facts that were wildly untrue or completely irrelevant. I contacted the College Board and discovered in a phone conversation with Wayne Camara, the Vice President for Research and Development, that the official scoring guide for readers explicitly instructed readers not to penalize students for presenting incorrect information. A few weeks later, Michael Winerip, an education reporter for the *New York Times*, shared with me the “Official Guide for Scorers of the SAT Writing Essay,” which explains: “Writers may make errors in facts or information that do not affect the quality of their essays. For example, a writer may state ‘The American Revolution began in 1842’ or ‘Anna Karenina,’ a play by the French author Joseph Conrad, was a very upbeat literary work. ... You are scoring the writing, and not the correctness of facts” (Winerip, 2005). Clearly, individuals who knew the truth about the historical event or the novel did not intentionally lie. Simply, there was no reason to do so. In most cases, except in the bizarre universe of mass-market testing, giving correct information is always safer than presenting information that may be easily identified as false. We may assume then that the SAT Writing Test, like the test taken by White’s 5-paragraph essay student, rewards putative facts, regardless if they are true or false. In terms of this discussion, they reward bullshit.

Test preparation companies and authors of test preparation guides also advise students to make up information. Indeed, one subheading in the chapter on the SAT Writing Section in a popular SAT prep book is “Making Stuff Up” (Berger, Colton, Mistry, & Rossi, 2008). The section explicitly advises students not to lie. It instructs them to state as supporting facts information that could possibly be true; that is bullshit.

There are other strong indications that the SAT, and most probably similar tests, reward bullshit. At the end of the College Board booklet, there is an advertisement for “The Official SAT Online Course,” which in addition to its other features “offers auto essay scoring for practice SAT essays including the *ScoreWrite*TM essay.” A new version of the *ScoreWrite*TM Manual exists and is available online (The College Board, 2010).

“Auto essay scoring” means machine scoring, and although I have learned from my colleagues at MIT that machine scoring can do very well in scoring short written responses in very circumscribed knowledge domains, it is extremely invalid in scoring arguments or longer technical responses. What machine scoring can do is what is implicitly or explicitly asked of human graders in mass-market assessments to ensure interrater reliability: count. Scorers look for support and development in such features as length, the number of proper nouns, and the number of direct quotations. They look for infrequently used words such as *plethora*, *myriad*, and *egregious*. Moreover, they are explicitly instructed to not notice errors of fact. In sum, they are taught to grade like machines. The result is that the testing companies can show a close correlation between these graders and Automated Essay scoring (AES). What machines cannot do, especially in analyzing essays that are not constrained by a very narrow content domain, and human graders are forbidden to do while grading the SAT, is to differentiate false statements from true ones. That is, both humans and machines happily accept bullshit.

I have been testing this hypothesis for several years by coaching now a total of 15 students retaking the SAT and 11 students taking the GRE, the GMAT (both graded by a machine and human grader), and the MCAT. [Full disclosure: I do not ask for any compensation for my coaching but ask the students (all of them are over 18) to send me their score along with a copy of the essay if the particular test makes it available to them. If they feel that my coaching has monetary value to them, I ask them to make a contribution of any amount to FairTest.org.] I coach them to include lots of detail and proper nouns and quotations regardless of whether they make sense. “Don’t worry about the truth,” I tell them. In sum, I tell them to bullshit and, frighteningly, it works very well. In all 15 cases in which I coached students retaking the SAT, I have raised their essay scores in the writing section, even though all of their original scores were already above the mean. The following is a transcript of part of an essay written by one of my early participants:

A major reason why cooperation is a preference to competition is because competition induces civil struggle at a time of crisis while cooperation reduces tension. In the 1930's, American businesses were locked in a fierce economic competition with Russian merchants for fear that their communist philosophies would dominate American markets. As a result, American competition drove the country into an economic depression and the only way to pull them out of it was through civil cooperation. American president Franklin Delenor [sic] Roosevelt advocated for civil unity despite the communist threat of success by quoting "the only thing we need to fear is itself," which desdained [sic] competition as an alternative to cooperation for success. In the end, the American economy pulled out of the depression and succeeded communism.

Because of the spirit of unity it induces, cooperation is the key to success. People unified work as a larger and stronger than those separated by competition, allowing utmost success to transpire.

The student who wrote this essay received two independent scores of "5" (the 92nd percentile) compared to the two "4s" the student received the first time he/she took the SAT. I interviewed the student after the return of the score and paper online, and the student responded that he/she was unsure of the details about the Great Depression so some hastily and partially thought out details from the Cold War were added. The student was amazed my strategy worked so well.

When I recounted this incident to my students at MIT, they were not nearly so amazed. Their almost unanimous response was, "Didn't you know that you always make up supporting evidence when taking standardized writing tests?" "Of course you make up personal stories," one of my students told me, "it takes less time than trying to remember one and you can always make it fit your thesis." My students at MIT, of course, are the students who excelled on the SAT, making me realize the paradox that students need to disregard the truth to get into an institution like MIT that venerates it. Mass-market testing makes them into bullshitters in order to get into a place that has little tolerance for bullshit.

THE FORMAT OF THE TIMED-IMPROMPTU AND BULLSHIT

The timed impromptu is often justified as being very similar to the essay tests that students will take as part of many college courses. But such a comparison is inherently flawed. The prompts for college essay tests are based on having students display and use knowledge, modes of analysis, or both particular to the field of study as the content and engine driving the writing of the essay. The prompts used in mass-market writing assessments are not based on

the extensive readings, lectures, and class discussions that provide the content for the college essay test. The mass-market writing prompt is designed to be accessible to anyone taking the test without any additional input such as readings longer than a paragraph. In essence, the topics invite students to write about subjects for which they have little if any real information and about which they may never have given much consideration. As Frankfurt (2004) notes in his essay, “Bullshit is unavoidable whenever circumstances require someone to talk without knowing what he is talking about.” (p. 63). It was precisely this extreme construct underrepresentation that led me and others to develop iMOAT (see Peckham, Chapter 9).

White (1995) places the timed impromptu between two other assessment instruments, indirect multiple-choice tests and portfolio assessments. He then argues that in certain venues, such as freshman placement tests, the timed impromptu is an appropriate assessment tool because for many schools, “Most of us would be happy if our entering freshmen really knew how to predicate and read” (p. 33). White’s analysis, however, is illuminating but incomplete. He essentially positions writing assessment upon two axes. The first is indirect (e.g., multiple choice) versus direct (i.e., writing). The second axis differentiates first-draft writing (the timed impromptu) from collections of revised writing (the writing portfolio). There is a third and very important axis, writing with knowledge of the topic and writing without knowledge of the topic, that is, writing bullshit. In real life, there are many instances in which an individual has to write on demand. In most of the cases, such as a student taking an exam in a specific subject, it is assumed that the writer has some familiarity with the content. I may get an email from a Dean, for example, asking for a quick response on strategies for reducing plagiarism or on the possible use of a specific writing test in the admission process. Because these are issues I have knowledge of, I can send a rapid reply that is based on my knowledge of the subject. I am not writing bullshit. No Dean, Provost, or President has, however, emailed me asking “Is failure necessary for success? Please reply in 25 minutes.” If I received such a request, my reply would certainly be bullshit.

DISTORTED ESSAY SCORING AS BULLSHIT

How did the essay on “Franklin Delenor Roosevelt” receive two scores of 5? According to the current rubric on the College Board web site:

An essay in this category demonstrates reasonably consistent mastery, although it has occasional errors or lapses in quality. A typical essay:

- Effectively develops a point of view on the issue and demonstrates strong critical thinking, generally using appropriate examples, reasons and other evidence to support its position
- Is well organized and focused, demonstrating coherence and progression of ideas
- Exhibits facility in the use of language, using appropriate vocabulary. (The College Board, 2010)

The essay in question clearly has more than “occasional errors or lapses in quality.” Indeed, it would be difficult to argue that this essay’s claim that the Great Depression was caused by competition between American business and Russian merchants with communist philosophies “demonstrates strong critical thinking, generally using appropriate examples ... to support its position.”

In both his landmark essay “Holisticism (1984) and the subsequent two editions of *Teaching and Assessing Writing* (1994), which incorporate much of the original essay, White warns of many of the pitfalls that can occur during a holistic scoring session, especially the problems produced when readers “feel intimidated or coerced by insensitive leaders, or harassed by an uncomfortable or autocratic working environment.” (p. 408). The grading environments for most mass-market writing assessments are what Robert Schaeffer of Fair Test calls “cyber sweatshops” (Owen, 2006, p. 26). Graders are not part of an interpretative community as envisioned by White, but moonlighting teachers and graduate students. Often they are not even grading together, but are grading alone on their home computer in minimum segments of 4 hours and are expected to grade 20 essays each hour, or one essay every 3 minutes. The all too human reactions to these more than uncomfortable work environments is chronicled by Farley (2009) in his memoir of 15 years in the standardized testing industry. Describing his first experience in a holistic reading for National Computer Systems (now part of Pearson Education) he recounts how he and other readers started just skimming the first paragraph of essays and then scored mainly on length and spelling. Readers and table leaders stopped caring about insightful rank ordering; they just wanted to get the job done the easiest way possible. Thus, their scores did not reflect a thoughtful evaluation of student writing, but successful attempts to quickly give an essay a score that would not appear to be deviant. In short, the scoring itself was bullshit.

Not all holistic scoring sessions are bullshit. I have been running sessions for 30 years, and agree with White (1996) that when done well, with sensitivity and respect for the readers and a primary concern about the quality of the reading rather than its economic cost, holistic scoring is a reliable method for assessing writing. As White (1996) notes, however, when testing firms are in control they desire procedures that will produce scores quickly and cheaply.

RESEARCH ON WRITING ASSESSMENT AS BULLSHIT

In March 2006, newspapers began reporting that a mechanical error had caused the misscoring of thousands of SAT tests from October 2005. Within a 2-week period, the College Board announced three times that their earlier assessments of the scale of the problem were low, and, in the end reported that approximately 4,400 students received official scores lower than what they actually earned, and 613 students received scores that overreported their performance. Some critics, such as one Dean of Admissions, implied that the College Board staff was lying (Arenson, 2006). I believe a better explanation was that they were bullshitting—they had no real estimation of the extent of the problem and kept trying to minimize it because they wanted it to be so, not that they knew it was untrue.

This same mindset infects the institutional research of mass-market testing organizations. Elliot (2005) chronicles many excellent researchers in the field of writing research. Some of the greatest, including Paul Diederich, Fred Godshalk, and Hunter Breeland worked for the Educational Testing Service. Even before ETS was created, however, giants in the field such as Carl Brigham (1937) had warned of the tendency in a large new testing organization to replace solid research with propaganda. Propaganda is not necessarily bullshit, but bullshit often serves as Propaganda.

Confirming some of Brigham's fears, few recent College Board Research, ETS Research, or Pearson Educational Research Reports discuss failings or problems in any of the standardized tests owned, developed, or administered by these companies. Like Lake Woebegone, these are communities where all the children are above average. Moreover, it has become commonplace for the abstracts of many of these research reports to include assertions that are not supported by data in the actual report. Because of space limitations, I focus on only one report, which is representative of the bullshit contained in many of them.

In 2005, the College Board modified the SAT by adding what had been "The SAT II Writing Test" as "the Writing Section" of the SAT itself. The first College Board research report that year was *A Survey to Evaluate the Alignment of the New SAT® Writing and Critical Reading Sections to Curricula and Instructional Practices* (Milewski, Johnson, Glazer, & Kubota, 2005). Although the report begins by stating that "researchers are encouraged to freely express their professional judgment," several of the key conclusions do not reflect the common standards of judgment in social science based on the data presented. Instead, the report asserts a strong alignment between the new SAT Writing and Critical Reading sections, high school and college curricula, and instruction because the authors wish it so. These assertions are, of course, bullshit.

The report also contains some outright factual errors, but, because the most glaring of these errors in the Abstract is contradicted on the same page, I attribute it more to a lack of concern for accuracy than to an active intention to lie. The abstract reports that a total of 2,351 high school students and college teachers were surveyed about student reading and writing skills. Yet on the same page, it is later stated that the survey was sent out to approximately 38,000 teachers. (At the end of the report, the exact number, 38,848, is given.) We also learn in the introduction that not all of the 2,351 respondents came from the original survey, which was conducted on the web. Because the return rate was low, over a quarter of these respondents were readers for the Advanced Placement (AP) English Examinations, who used a paper form rather than the web-based instrument. Of course, these readers, who are being trained and paid by the College Board may not be a representative sample of the over 37,000 surveyed teachers who chose not to respond to the survey.

Although the descriptive information on all the participants is given on page four of the report, the statistics do not segregate the AP readers from the ordinary responders. Moreover, although this section reports that the high school teachers had more education than national samples and that the college professors overrepresented public institutions, it asserts, “Despite the over and underrepresentation, the current sample can still be considered representative of high school and college faculty” (p. 4).

No mention of the low response rate or the possible bias in the addition of the AP readers is made until the end of the report in a section entitled “Limitations and Next Steps.” The actual response rate to the online survey was 4.45%, well below the median return rate of 30% for online surveys and less than half of the minimum acceptable return rate of 10% even for extremely large surveys like this one (Armstrong & Overton, 1977; Punch, 2003; Sheehan, 2001). The authors of the report acknowledge at least part of the problem at the end by stating, “The low response rate limits the generalizability of the survey results and suggests that volunteer bias might be at play” (p. 23). Moreover, it is only in this section, at the end of the report, that the authors give the exact number of respondents who were AP graders, 622 or 26.5% of the total sample. Although they admit that these teachers might not be representative of the general population of teachers, the authors neglect to mention that the table training they received as graders that same week might increase their bias.

There is a question, however, whether we can consider these kinds of deceptions as bullshit. The authors do admit that their study has limitations; they just do not mention that these limitations make the data highly limited if not worthless, and they hide the extent of the problem until the end of the report. But the primary assertion of the study, that the reading and writing abilities the SAT claims to measure align with high school and college cur-

ricula, is, in the terminology of Scottish Law, Not Proven. The assertion could be true, or it could be false. The methodology of the study, however, is largely unconcerned with that central issue, leading to the conclusion that as a whole, the report falls under Frankfurt's definition of bullshit.

This conclusion is reinforced by other assertions in the report. Of the college instructors participating in the study, 814, who taught writing, language arts, or English, were classified as *English professors*. Another 230 college faculty, who taught history, political science, psychology, or biology were classified as *Humanities professors*. A graph in the report exhibits that slightly more than 20% of English professors employed multiple-choice tests, whereas slightly fewer than 60% of the Humanities professors gave multiple-choice exams. Then, even though the English professors constituted 78% of the college professors, the report states that both high school teachers and college professors give multiple-choice tests. The report also asserts that the test emphasizes Writing Across the Curriculum because prompts such as "Are people motivated to achieve by personal satisfaction rather than by money or fame?" are "relevant to a wide range of fields and interests rather than narrowly related to specific topics or disciplines" (p. 23). In addition, the report asserts that essay prompts, which contain passages of four or five sentences totaling approximately 75-80 words for an essay to be written in 25 minutes measures students' abilities to use "writing and reading as tools for critical thinking." Such claims are clearly bullshit. It is not that the authors know that the claims are false; they, like the high achieving students writing the SAT essays, just do not really care about the truth.

Carl C. Brigham, the originator of the SAT and a eugenicist who renounced eugenics and the entire concept of aptitude testing when confronted with convincing and contradictory data, saw the dangers of mass market testing over 70 years ago. Discussing the possible creation of ETS, he wrote (1937):

It is easy for a powerful organization to set up false ideals. The new organization must be so contrived that it will always remain the servant of education and never become its master. It should inquire into the nature of values but it should not determine those values.

At the present time there are men of learning who sense these values intuitively yet are unable to put them on canvas with pigments which stand reproduction. Testing situations, when properly formulated, and with responses fully analyzed, constitute the most searching system of lenses yet contrived for photographing the canvas and making it generally available to mankind. The artist must work with the technician to get the results for which he is striving.

To-day most testers are content to buy a cheap, ready-made camera with a poor lens, fixed focus and no range finder, and to do a develop

while-you-wait business with stragglers on the sidewalk in front of the studio. Furthermore, they claim that the present camera can reproduce every detail of the completed masterpiece within—although they have merely peeked in the door and have never seen the canvas.

The new organization must plan to discard most of its present apparatus within five to ten years and approach its larger problem in a spirit of humility. The ideals of research are lofty, but its spirit is meek. An organization with major purpose research—and not propaganda—will be of great service to education. Education can not to-day afford to let the sidewalk vendor dictate its objectives, but it can properly help him adapt his apparatus to its own purposes. (p. 758)

Brigham's metaphor of the sidewalk photographer who claims that his little box camera can capture every detail of a masterpiece he has never seen is a clear description of a master bullshiter (although Brigham would have never used that term). I prefer a more contemporary metaphor. Mass market testing organizations want mass sales, giving their customers not nutritious food that will make them wise and healthy, but quickly produced products that may look attractive but actually hurt the mind and body rather than improve it. They do not foster education; they promote McLearning, high in sugar, salt, and fat, with low, if any, real nutritive value. But like the hamburger chain, the aim of mass-market testers is not education; it is an obese bottom line on the balance sheet.

ENDNOTE

1. Two books that discuss the inherent lack of concern with truth among large-scale testing companies are Farley (2009) and Lemann (1999).

REFERENCES

- Arenson, K. W. (2006, March 23). SAT problems even larger than reported. *The New York Times*, p. 1.
- Armstrong, J. S., & Overton, T. S. (1977). Estimating nonresponse bias in mail surveys. *Journal of Marketing Research*, 14(3), 396-402.
- Berger, L., Colton, M., Mistry, M., & Rossi, P. (2008). *Up your score: The underground guide to the SAT 2009-2010 edition*. New York, NY: Workman Publishing Company.

- Brigham, C. C. (1937, December). The place of research in a testing organization. *School and Society*, 756-759.
- Cohen, G. A. (2002). Deeper into bullshit. In S. Buss & L. Overton (Eds.), *Contours of agency: Themes from the philosophy of Harry Frankfurt* (pp. 321-336). Cambridge, MA: MIT Press.
- The College Board. (2010). *SAT essay scoring—How essays are scored*. Web.
- The College Board. (2005). *ScoreWrite: A guide for preparing for the new SAT essay*. New York, NY: The College Board.
- Elliot, N. (2005). *On a scale: A social history of writing assessment in America*. New York, NY: Peter Lang.
- Eubanks, P., & Schaeffer, J. D. (2008). A kind word for bullshit: The problem of academic writing. *College Composition and Communication*, 59(3), 372-88.
- Farley, T. (2009). *Making the grade: My misadventures in the standardized testing industry*. Sausalito, CA: PoliPoint Press.
- Frankfurt, H. G. (2002). Response to G.A. Cohen. In S. Buss & L. Overton (Eds.), *Contours of agency: Themes from the philosophy of Harry Frankfurt* (pp. 337-339). Cambridge, MA: MIT Press.
- Frankfurt, H. G. (2004). *On bullshit*. Princeton, NJ and Oxford, UK: Princeton University Press.
- Fredal J. (2011). Rhetoric and bullshit. *College English*, 73(3), 243-259.
- Lemann, N. (1999). *The big test: The secret history of the American meritocracy*. New York, NY: Farrar, Straus, and Giroux.
- Milewski, G. B., Johnson, D., Glazer, N., & Kubota, M. (2005). *A survey to evaluate the alignment of the New SAT Writing and Critical Reading sections to curricula and instructional practices* (College Board Research Report 2005-1). New York, NY: College Board. Web.
- Owen, D. (2006, April 3). The S.A.T.'s watchdog. *The New Yorker*, 26.
- Perelman, L. (1994). The rhetoric of a major malfunction: The institutional and rhetorical dimensions of the explosion of the space shuttle Challenger. In C. Sides (Ed.), *Technical communication frontiers: Essays in theory* (pp. 43-59). St. Paul, MN: Association of Teachers of Technical Writing.
- Plato. (2004). *Gorgias* (C. Emlyn-Jones, Ed.) (W. Hamilton & C. Emlyn-Jones, Trans.). London, UK and New York, NY: Penguin.
- Punch, K. F. (2003). *Survey research: The basics*. London, UK: Sage.
- S. Rep. No. 108-301. (2004).
- S. Rep. No. 109-331. (2006).
- Sheehan, K. (2001, January). E-mail survey response rates: A review. *Journal of Computer Mediated Communication*, 6(2). Web.
- Smargorinsky, P., Daigle, E. A., O'Donnell-Allen, C., & Bynum, S. (2010). Bullshit in academic writing: A protocol analysis of a senior's process of interpreting *Much Ado about Nothing*. *Research in the Teaching of English*, 44(4), 368-405.
- Tietge, D. J. (2006). Rhetoric is not bullshit. In G. L. Hardcastle & G. A. Reisch (Eds.), *Bullshit and philosophy*. Chicago, IL and LaSalle, IL: Open Court.
- Tuchman, B. (1984). *The march of folly: From Troy to Vietnam*. New York, NY: Knopf.
- White, E. M. (1984). Holisticism. *College Composition and Communication*, 35(4), 400-409.

- White, E. M. (1994). *Teaching and assessing writing* (2nd ed.). San Francisco, CA: Jossey-Bass.
- White, E. M. (1995). An apology for the timed impromptu essay test. *College Composition and Communication*, 46(1), 30-45.
- White, E. M. (1996). Power and agenda setting in writing assessment. In E. M. White, W. Lutz, & S. Kamusikiri (Eds.), *Assessment of writing: Politics, policies, practices* (pp. 9-24). New York, NY: Modern Language Association.
- White, E. M. (2008). My five-paragraph-theme. *College Composition and Communication*, 59(3), 524-524.
- Wilson, J. (2004). *The politics of truth: A diplomat's memoir: Inside the lies that led to war and betrayed my wife's CIA identity*. New York, NY: Carroll & Graf Publishers.
- Winerip, M. (2005, May 4). SAT essay test rewards length and ignores errors of fact. *The New York Times*. Web.

Assessment as Learning. Assessment of Learning. Slide Number 20. To come to a common understanding of the meaning of Assessment for, as, and of Learning. To gather reliable and valid evidence based on informed professional judgment. Take Time to Think. Mass-market Writing Assessments as Bullshit more. by Les Perelman. Research Interests: Writing Assessment. Download (.pdf). Save to Library. Symposium: Assessment. more. It includes: (1) "Closed Systems and Standardized Writing Tests"; (Chris M. Anson); (2) "Information Illiteracy and Mass Market Writing more. This article presents four symposium papers on assessment. It includes: (1) "Closed Systems and Standardized Writing Tests"; (Chris M. Anson); (2) "Information Illiteracy and Mass Market Writing Assessments"; (Les Perelman); (3) "Genre, Testing, and the Constructed Realities Publisher: eric.ed.gov. Request PDF on ResearchGate | On Jan 1, 2012, L. Perelman and others published Mass-market writing assessments as bullshit. Assessment of educational outcomes through purchased tests is commonplace in the evaluation of individual student ability and of educational programs. Focusing on the assessment of writing performance in a longitudinal study of first-time, full-time students (n = 598), this research describes the design, use, and assessment of an open-source scoring platform. Augmenting usability testing, the research design relies on a framework of inter-reader agreement, inter-reader reliability, and coefficients of determination.