

Fall 2016

## **An Accurate Representation of What?: Student Perceptions of Standardized Writing Assessment**

Kristen Getchell

Lindsay Illich

Follow this and additional works at: <https://scholarworks.utrgv.edu/crosspol>



Part of the [Education Commons](#)

---

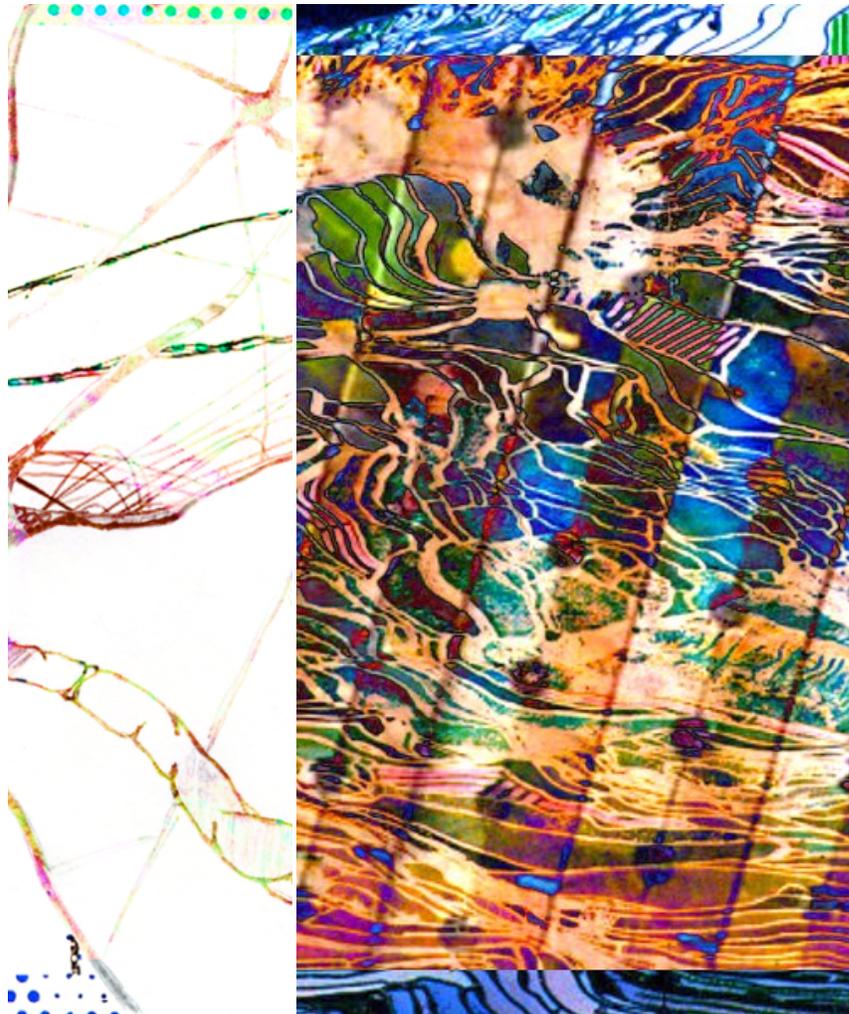
### **Recommended Citation**

Getchell, K. & Illich, L. (2016). An Accurate Representation of What?: Student Perceptions of Standardized Writing Assessment. *Crosspol: A Journal of Transitions for High School and College Writing Teachers*, 2(1), 38–49.

This Article is brought to you for free and open access by ScholarWorks @ UTRGV. It has been accepted for inclusion in Crosspol by an authorized administrator of ScholarWorks @ UTRGV. For more information, please contact [justin.white@utrgv.edu](mailto:justin.white@utrgv.edu), [william.flores01@utrgv.edu](mailto:william.flores01@utrgv.edu).

# an accurate representation of what?: student perceptions of standardized writing assessment

kristen getchell + lindsay illich

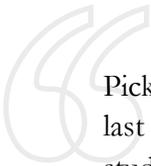


*Most of our students challenged the SAT structure and format and demonstrated a significant distrust of written assessment of writing, something they are still harboring as first-year college students.*

## **An Accurate Representation of What?:**

### **Student Perceptions of Standardized Writing Assessment**

Kristen Getchell + Lindsay Illitch



Pick at random any study of the testing of writing published in the last twenty years. What are the chances that it will investigate how students are affected by the testing? Test construction, yes. Test administration, yes. Test results, yes. Test impingement on student minds and lives, rarely (Haswell, WPA-L, March 24, 2014).

What do students talk about when they talk about testing? Since the debate over the SAT writing section has entered mainstream media, discussing testing has become a door for us (and we suspect, other writing teachers), a door opening up conversations about relevant and timely issues in writing studies with our students. We talk about what constitutes good writing, which leads to discussions of rhetorical situations. We talk about how tests like AP and SAT are scored and the problems with machine-scoring; we mention text-attributes. And what we found is that students have stories to tell about the testing culture they have experienced. At our institution, many our students self-identify as poor test-takers and feel that standardized tests such as the MCAS Massachusetts State tests or NY Regents exams and the SAT and ACT tests are poor representations of their ability to succeed in the classroom. Despite this, about 86% of our students take the SAT Writing exam, by far the most widely taken standardized test for our incoming students. Our incoming average GPA for enrolled students for Fall 2014 is a 2.8, with an average SAT verbal of 471 and SAT writing score of 470. Approximately 22% of our student body are enrolled in a fee-for-service program for language-based learning disabilities, and, of the remaining population, approximately 8% have registered for accommodations with our disability services office that may include extended time on tests and quizzes.

In our study, students wrote a diagnostic essay—a regular feature of the first-year writing course to spot-check placement—and a reflection on the diagnostic essay. The activities together serve as a sequence of activities aimed at encouraging reflection and critical evaluation of previous writing assessment. The assignments also helped us extend a line of inquiry related to student perceptions of assessment. We found that students have much to say about their own perceptions of the adequacy of the test to measure their writing ability—something they often refer to as

“fairness.” Further, they point to what they believe are deficiencies in the structure, format, and content of the testing experience as a test of writing. However, while there are certainly trends in their responses, students’ understanding of the construct of the test is varied, and their consideration of the influences, consequences, and/or decisions being made about the test are not explicit or sustained.

## Context

Haswell was right about the scantiness of literature attending to student perceptions of assessment. Kathleen Blake Yancey’s landmark 1999 “Looking Back As We Look Forward: Historicizing Writing Assessment” asked readers to consider the “role of the person/al” and to consider “what (else) might we learn from writing assessment,” inviting researchers to ask more questions along these lines (484-485). She understands these “humane and ethical” dimensions as the most significant in the history of assessment over the last fifty years, yet few studies have taken up Yancey’s charge to further study. Albertson and Martwitz analyzed testing artifacts, including prewriting, to document how students negotiated a timed proficiency exam required of all students before enrolling in upper level courses, extending arguments about timed writing tests as inadequate measures of students’ writing processes. However, Albertson’s and Martwitz’s study focused narrowly on pre-writing artifacts and the test to draw conclusions, without students’ accounts of their writing processes to answer questions about students’ relationship to writing assessment. Also, the study didn’t address the effects of the tests or how students experienced it. Petersen’s “‘This Test Makes No Freaking Sense’: Criticism, Confusion, and Frustration in Timed Writing” similarly analyzes the assessment artifacts of his institution’s timed writing portion of the junior portfolio. In this study, some students resisted the prompt, evidenced in the text of their responses to the prompt. While some resistance was covert, others overtly challenged the test’s assumptions (thus the titular quote from a student’s marginalia) and challenged the relevancy of the test.

To extend this line of inquiry, we used writing test assessment as the topic for the diagnostic reading, as well as added a reflection on the diagnostic essay. In the *diagnostic essay* (see teaching artifact 1), students were asked to read and respond to an article by Joanna Weiss in the *Boston Globe* titled “The Man Who Killed the SAT Essay,” which summarizes the research of Les Perelman at MIT on machine-grading (or robo-grading in Perelman’s words) that exposes critical inadequacies with the algorithms designed to “grade” the writing section of the SAT. Specifically, students were asked to read the article and first summarize the main points in 1-2 paragraphs, and respond to the ideas in the article by citing their own opinions or experiences. The assignment prompt explains in both the beginning and the end of the prompt that this will serve as a way for instructors to plan for the course.

The *reflection on the diagnostic essay* (see teaching artifact 2) was developed in response to two concerns. First, the *Boston Globe* article was a new text for diagnostic

essay, and we were interested in getting student perceptions of the article and of the diagnostic assignment itself. Second, we were interested in providing a space for students to respond honestly to the diagnostic experience. We hoped that by asking students about their experience with the diagnostic, we were inviting them into a conversation about assessment and encouraging them to see writing assessment as a complicated act that should be open to their own critical perspectives. We planned for the reflective assignment to serve as a transition between the diagnostic essay and their first major essay (the literacy narrative), and we hoped it would allow students an opportunity to express their ideas, feelings, and frustrations with their experiences with writing evaluation. Our goal was that these writing episodes would all work in concert to create space for a conversation in the course about literacy development and writing assessment.

In the following section, we will provide examples of student responses to these two assignments to illustrate students' perceptions of standardized tests like the SAT as a measure of ability—writing or otherwise—and their views on the appropriateness of the instruments used in these types of tests.

### **A Comfort Zone**

One of the categories that emerged from student writing was repeated discussion of perceptions of tests and whether or not they adequately addressed what it was purporting to measure (in some educational testing circles referred to as “face validity”). There was a range of opinions about whether the test seemed to measure what it was supposed to assess. However, among our students, there was also some disagreement about what exactly the test intended to measure or was supposed to show about them as writers, students, or individuals.

We'll start with Robert, who found that both the SAT and later the diagnostic essay assignment were good representations of his writing. In response to The Boston Globe article, he characterizes what he views as Perelman's argument about the SAT as “absurd considering that the Essay is not meant to evaluate a student's life experiences or opinions” but instead “to gauge a student's ability to organize their thoughts and display the abilities they were meant to be learning in twelve years of schooling. The essay shows that a student can express their thoughts and develop a written piece while following a prompt.” Robert's confidence is welcome and something we see only occasionally in our first-year writers, but he is supporting an argument against one of Perelman's criticisms of the SAT, that the prompts do not encourage students to provide actual, real world knowledge. In his disagreement with Perelman, Robert is also indicating what he believes the SAT is supposed to measure—specifically, organization and expression of thoughts and a student's ability to follow a prompt. To Robert, those seem to be important elements of writing process; therefore, the SAT is an appropriate test.

In his reflection, Robert writes that the diagnostic, with “a specific prompt and a time limit,” put him in his “comfort zone for writing.” It is because of this

comfort with the testing format that he believes that the diagnostic was “an appropriate diagnostic tool for [his] abilities.” Robert continued his support of standardized testing, including our essay prompt in this category, because these types of test are “good indicators of a student’s abilities because they are so basic in their framework. Simple prompts and specific guidelines for the essays not only to be a test of writing, but at the same time, of organization, time management, and ability to follow instructions.” Again, we see Robert willing to defend the defining characteristic of a standardized test: the standardization. In his mind, clarity and simplicity are equal to fairness. Robert also admitted that it was under these circumstances that he feels most comfortable as a writer, using his own experience to support his claim.

While Robert does not quite make it to the point of explaining the decisions being made on behalf of the test score, he does make a clear connection between format of the test and the criteria that are being examined. The majority of the rest of the class, unfortunately, would disagree with Robert. Many of our students extracted and agreed with several of Perelman’s major issues with the SAT as outlined in the reading: prompts that support empty speech, lack of time for pre-writing and invention, and the context for scoring which paints a harried and underpaid scoring staff.

### **Good Writing Takes Time**

As we reviewed our student responses, we found that the most frequent complaint about the SAT was with the amount of time students were given to complete the test. When students address time constraints, they did so in interesting ways. Two students discussed time constraints extensively in both the diagnostic essay and the diagnostic reflection. When they did so, time became a lens through which ideas became things to be found, and further, that information from outside sources were also figured as things, so when one student wrote about needing more time, she meant not only time to access ideas in her mind but also time to access information from outside sources. Emma wrote about time constraints inhibiting her ability to “gather” thoughts to be displayed: “with only twenty-five minutes to gather your thoughts and put them down on paper in an essay format, it is difficult to put your best work on the table.” She goes on to describe her frustration: “I have felt pressured and rushed during the essay portion. When writing an essay, I need time to think my thoughts through and come up with good ideas to write about. I feel that writing an essay in only twenty five minutes makes me rush my thoughts and that it is not my best work.” In an interesting reversal of the Ginsberg phrase “first thought, best thought,” Emma here is a pearl diver, going down into the depths of her brain to “come up” with good ideas, an accomplishment that takes time. The ideas aren’t readily or easily available to her. Like a set of lost keys, her ideas need to be found.

In her reflection, she picks up the “gathering” metaphor again, but this time pushes it further to include outside sources as well as her own ideas: “I like having

the time to gather my thoughts and look up information to support my answer.” Here, the time constraint is conflated with another constraint of the timed-writing environment: no outside sources can be consulted. If writing is figured as a Burkean parlor where writers join in the conversation, what happens when you’re talking to yourself in a mirror? Though this isn’t an issue related to time constraints, this student sees a relationship between the two, especially given the cultural metaphor of ideas as things (Lakoff and Johnson 10). If it takes time to locate my own (even in my head), then it stands to reason that it also takes time to find others’ ideas (in the form of sources). This idea is further articulated later in the reflection: “I would have had more time to look for quotes and find out more information about the topic. When writing papers I like to have enough information to thoroughly support my answer and have enough evidence to prove a point.” Here, again, “information” comes from inside and outside the writer.

Emma’s diagnostic also reflected somewhat mechanistic ideas about process. Specifically, she wrote that not having prior knowledge of the topic made getting started difficult. Here, invention is figured as an engine that needs to warm before taking off. Writing “cold,” without “knowing what you’re writing about before the test,” like sight reading music for musicians, represents the most challenging kind of writing: “You have no knowledge of the topic beforehand and it may be challenging to think of good points to back up your reasoning.” In this case, Emma is identifying the issue of empty responses as an issue of time constraint, not necessarily a deficiency with the prompt itself. She takes specific issue with Perelman’s point that SAT essays seem to be scored largely based on length: “Grading someone based on how long their essay is, is not accurate and it is unfair to other students. One could ramble on about the topic making it sound like they know what they are talking about, when in fact they do not.” For Emma, longer essays are not better essays.

For Kelsey, the better you are at writing, the faster you can do it: “I have never been a good at writing, let alone a good, fast writer.” In this figuration, the speed at which a task can be executed is related to how well you can execute a task. In other words, there’s not-so-good, slow writers, good, average speed writers, and good, fast writers. She goes on to suggest that time necessary can be segmented into smaller tasks: “In order for me to write an essay, I need time to think and plan ahead, like all of my teachers have taught me to do. Now, this testing is taking away these tools I have been taught by not allowing me time to do so.” That is, writing as mechanized process that takes time at each station in the factory line to assemble the end product. Kelsey’s quote reflects a quasi-process pedagogy, one that itself has become at the same time Fordian and rickety (even reified). In her reflection, she suggests that time constraints mean that she has to take shortcuts, skipping necessary steps in the manufacturing process: “I was rushed and forced to write quickly without being able to plan my essay out in advanced [sic]. I was unable to make a structured thought with details and examples because I did not have time for it.” How could she do well when necessary steps couldn’t be executed? Her assertions

about the test, however, were supported by her score, which she describes as “decent” and “average,” but not a good representation of her ability.

### **Structural Flaws: Empty Prompts Beget Empty Prose**

Despite the fact that time was the most frequent culprit cited as students explained the perceived inappropriateness of the test as a measure of their writing ability, other students presented complicated responses to other aspects of the test. Sara, for example, expressed concerns in her diagnostic similar to many students who found that time was an issue. She also identified problems with accuracy of scoring when she writes about a section of the article that explains how many SAT scorers often feel rushed when evaluating an essay: “It is not fair to the grader to have to rush their readings because they have so many more ahead of them to analyze. That just causes stress and will affect the grading of the upcoming essay.” While she identified problems with the “fairness” of the scoring, she does so in a way that questions whether the process is fair to the reader/scorer, not the student test takers themselves.

Sara did point to what she believed to be a strength of the SAT and other standardized tests: the perceived emphasis on the five-paragraph essay to evaluate student writing. She writes, “One small point that I disagree with Perelman about is getting rid of the five paragraph essay. I personally believe that it is an effective way to get young people to write. It organizes their thoughts and teaches them what components they need to write a basic paper.” While Sara considered this a “small point,” it is actually very important that students consider the format of their response to be one that is a good representation of their writing. She does concede later on, however, that “as [students] get older, there definitely should be more to their writing than just five basic paragraphs.” Sara, while she has respect for the classic organization of the five-paragraph essay and clearly sees a place for it, wants us to know that she doesn’t see this as the only form. Instead, the five-paragraph essay has provided scaffolding for young writers—perhaps this is what Robert refers to when he refers to the SAT structure as his “comfort zone.” Further, she feels that as students mature, they should grow out of this “basic” format to more complex structures.

In her reflection, Sara explains that she felt that the diagnostic represented her writing well because she was asked to express her own ideas and she was able to relate to the essay prompt with her own experiences. She contrasts her experience with the diagnostic to a test that she did not perceive to be a good representation of her writing, The Massachusetts Comprehensive Assessment System (MCAS) test. While she felt that she had enough time to complete the essays for the MCAS, she found them “tedious and redundant.” As she explains, “There would be prompts about the most random topics with only a limited amount of pages to write on. I feel like these types of questions were okay to answer for earlier years, but as we got older it would be nice if there was one relatable topic to write about for the whole

day.” Here Sara’s sentiment echoes Perelman’s argument: she is asking for writing prompts that are relatable and writing prompts that matter. In her estimation, it is this type of prompt that is important in eliciting good writing from students who are transitioning to college.

Allison also identified issues with the writing prompts and was less forgiving of the SAT and other standardized writing tests. She expressed her doubts at the adequacy of the SAT by characterizing the SAT as part of the larger category of standardized testing. She writes in her diagnostic: “Standardized testing in all forms, the SAT, the ACT, is not the best way to judge a student’s ability or intelligence. I feel these are simply testing the student on what the creator of the test thinks is important.” Allison feels that this test is not one that is only meant to measure writing, but, broadly, ability and intelligence. Questioning the value of the test creator, she is speaking to the perceived arbitrary nature of the test and goes on to illustrate this point further a few paragraphs later: “Sure we all learned about white dwarfs and red stars, and the composition of various rocks, and various other pieces of knowledge, but these tests are designed to trick up, to be deceiving.” Allison views the SAT as not just failing to do a good job of testing writing skill, but something more nefarious, a test that is meant to trick students. Allison’s willingness to up the ante, so to speak, when it comes to her perceptions of the SAT demonstrate her willingness to allow her own experiences to support an argument not made by Perelman or the author Weiss, in the article for the diagnostic. She deftly uses Perelman’s argument related to subject matter to support her own understanding of what the test makers are doing when creating the SAT essay test.

Both Allison and Sara challenge the prompt as a way to demonstrate that the type of writing produced by the SAT test—often decontextualized, empty prose—is a result of the test itself. In a sense, they are arguing that the College Board is getting exactly the type of writing they are asking for with their test.

### **A Broken System with Real World Consequences**

While Allison and others managed to classify the SAT as part of a larger category of standardized tests, Lauren’s response to the diagnostic essay and the reflection demonstrated an understanding of how these tests had higher stakes involved than just a demonstration of writing ability. Her response was unique in its awareness of the position of the SAT within a complex system of education, one where decisions are being made that come to bear on a student’s future.

Lauren first worked to establish her ethos on the topic: “I have extensive experiences with writing evaluations throughout my academic career. I have previously taken the New York State Regents exams which have included document based questions (DBQs), other standardized tests, the writing assessment, and etc.” From this experience, Lauren identified issues with the SAT characterizing the test as “an outdated, overly standardized, overemphasized, overly competitive, and unfair measure in determining college readiness in students for the twenty-first century.”

For Lauren, the SAT is less about specific writing abilities and more about overall academic preparedness for college. Her perceptions of the test do not rest merely on the unfairness of the test to represent her “readiness,” but on its perceived lack of adaptation to the changing educational landscape: “It is incompatible and irrelevant for the twenty-first century learning experience and environment.” Lauren even goes further to say that the lack of alignment with the SAT and the learning objectives of the twenty-first century are a “a missed opportunity to take into account a comprehensive and complete scope of students’ individual (academic) experiences and potential to learn.” Much like Sara, Lauren identifies a lack of assessment of individual learning experiences, ideas, and opinions in favor of standardized responses that can often be augmented by tutoring. Lauren’s perceptions of the SAT as “outdated” and “missing an opportunity” speak to the idea of the SAT as a whole as a broader systemic problem where the test, if administered well, could actually be a learning opportunity for students.

In her reflection, she explains that she doesn’t believe that the majority of tests she has taken “represented [her] writing ability genuinely or accurately,” and of our sixteen student writers, Lauren was one of two students who connected the SAT essay as an assessment of writing ability to the decisions being made based on the test. She finds this connection when discussing the inequalities of the testing system: “The best potential student is not always the one with the highest number. People who can afford a quality tutor have an unequal advantage to be accepted to a top college over those who can’t.” Here Lauren articulates the types of decisions made from the SAT score; one of these real world results of the test being college admissions. She concludes by making a case for the detrimental effects of these tests: “The expectations and pressure to do well are stressful and enormous on our youth.” In her mind, these looming decisions and the associated “overwhelming stress load” have the potential to “lead to detrimental consequences to their health.”

Another student to make the connection between the test and these decisions was Mia, who also identifies the stress that comes with such high stakes tests: “The fact that colleges look so highly at SAT scores freaks out high school students, and as it comes down to the time for applying to college and students aren’t happy with their scores it makes the process that much more stressful.” Mia also explains to us that there are significant, personal consequences of this testing experience other than its measure of writing ability.

## **Conclusion**

Unsurprisingly, our students demonstrate a complicated and personal relationship to the writing assessment. What we see from this preliminary data is that, for many of our students, written tests like the SAT are decontextualized and seemingly arbitrary. While they can construct arguments for why the SAT is or is not a good representation for their own writing, students were less able to discuss how

the testing experience influenced their writing and how the decisions being made on behalf of the test were appropriate or not.

The variety of words our students used to describe what the SAT test is purported to measure—ability, intelligence, readiness, skill, knowledge, potential, success—show that they are not entirely clear about the assumptions being made on behalf of the test. And how could they be? When and how is the SAT discussed in any meaningful way for high school students? Instead, they are left to rely on their own perceptions of the test, its similarity or dissimilarity to other tests they've taken, and perhaps a good sprinkling of misleading cultural folklore to round out their understanding of the test. As a result, student views of the test tend to run along two lines: some focus on elements of the test itself and some focus on the broad concepts that they believe the test is measuring.

Only two of our students explicitly discussed the decisions being made based on the test, or in the case of the SAT, the stakes for college admissions and placement decisions. To most of our students, writing assessments exist in a bubble where the score and its implication were rarely discussed, but Lauren and Mia identified that the test is part of a larger process of college admissions, a sorting process to Lauren. While neither addressed the appropriateness of the SAT in contributing to these decisions, they did identify a very troubling effect of testing: overwhelming stress on the student test taker.

Despite not clearly explaining how these tests changed or influenced their writing and writing process, we do see that the SAT essay has an impact on how students view writing assessment. None of our students pointed to ways that they read or write differently based on the test. While Sara identified that there was value in the five-paragraph format, she did not seem to consider that her writing had been changed by the five-paragraph essay test training she had encountered. Most of our students instead challenged the SAT structure and format and demonstrated a significant distrust of written assessment of writing, something they are still harboring as first-year college students. This distrust has potential to shape the way they view writing activities in all of their classes.

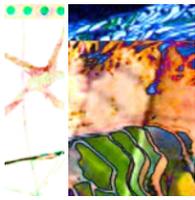
So, how do we respond to this type of feedback about standardized testing? We would argue that this type of feedback from students makes the work of first-year composition courses more complicated, but also essential for deconstructing some of the feelings of distrust and the misconceptions about writing assessment that students develop from their history with standardized testing. Engaging students in these types of conversations through writing assignments and classroom discussions that ask students to provide their opinions about testing is a way to begin to provide students with a voice when it comes to writing assessment. Even students like Robert, who reported feeling at ease in the testing environment, need to be able to reflect on their experience with different tests of writing to break down the test into its elements and identify their own strengths and weaknesses as a writer. Furthermore, a continued focus on providing students with assignments that clearly

demonstrate our values will help support students as they move away from the standardized models. Clearly articulated prompts with criteria for grading, thoughtful and considerate feedback that demonstrates a close reading of their work, a negotiated space for process in response to a variety of rhetorical situations (beyond the five-paragraph essay): these best practices that those in the field have been advocating for years need to be examined for their role in rebuilding our students trust and strengthening their understanding of what it means to write confidently and write well.



## References

- Albertson, Kathy, and Mary Marwitz. "The Silent Scream: Students Negotiating Time Writing Assessments." *TETYC* 29.2 (2001): 144-53. Print.
- Haswell, Richard. "debate on writing assessment" WPA-L. Web. March 24, 2014. 1 Nov. 2014.
- Lakoff, George, and Mark Johnson. *Metaphors We Live By*. Chicago, IL: University of Chicago Press, 2003. Print.
- Peterson, Jerry. "'The Test Makes No Freaking Sense': Criticism, Confusion, and Frustration in Timed Writing." *Assessing Writing* 14 (2009): 178-93. *Ebsco*. Web. 1 Nov. 2014.
- Yancey, Kathleen Blake. "Looking Back as We Look Forward: Historicizing Writing Assessment." *College Composition and Communication* 50.3 (1999): 483-503. *JSTOR*. Web. 7 Dec. 2011.



## **An Accurate Representation of What?: Student Perceptions of Standardized Writing Assessment Teaching Artifacts**

*Teaching Artifacts*

Kristen Getchell + Lindsay Illitch

### **1. Diagnostic Essay Assignment**

Welcome to ENG 1280: Writing Workshop I, the first course in a two-sequence required writing program to introduce you to the conventions of academic writing, practice your writing skills, and learn new strategies to improve your writing. The first step on our journey together is to get a snapshot of your current writing abilities. Think of it as your “before” picture. At the end of the course, you will write a similar essay, a snapshot of your writing at the end of the course--your “after” picture.

For this assignment, read the article by Joanna Weiss from *The Boston Globe*, “The man who killed the SAT,” an article that discusses the research focusing on the SAT Writing test: <http://www.bostonglobe.com/opinion/2014/03/13/the-man-who-killed-sat-essay/L9v3dbPXewKq8oAvOUqONM/story.html>. Next, write a summary of the article in one or two paragraphs. Then, respond to the author’s claims in the article by citing personal opinion, experience, or other evidence. Your response should be brief (no more than two or three paragraphs). Remember, a response can be a critique of one or more of the writer’s claims, elaboration of one or more specific points the writer makes, or a discussion of the way or ways in which the writer makes and argues his or her point.

As you should with any piece of writing, take time to proofread your essay carefully. Since I will use it to guide my planning for the class and to personalize instruction to meet your learning needs, it is important that the diagnostic essay represents your best efforts.

### **2. Reflection on Diagnostic Essay**

*Short reflective writing assignment (250 word minimum, typed)*

Was the diagnostic essay a good representation of your writing? Why or why not? What other kinds of writing tests have you taken? Do you feel these tests were a good representation of your writing ability?