

What does fairness mean when assessing multilingual writers?



Mya Poe (m.poe@neu.edu)
Northeastern University

MATSOL
Framingham, MA
May 9, 2014

What is writing assessment?

Writing assessment is the gathering of information about students' written literacy for a variety of formative and summative purposes but primarily to support effective teaching and learning.

Gallagher, "What Do WPAs Need to Know about Writing Assessment? An Immodest Proposal," 2009

CCCC Position Statement

A statement on an education issue approved by the CCCC Executive Committee

Writing Assessment: A Position Statement

1. Writing assessment is useful primarily as a means of improving teaching and learning. The primary purpose of any assessment should govern its design, its implementation, and the generation and dissemination of its results.

As a result...

A. Best assessment practice is informed by pedagogical and curricular goals, which are in turn formatively affected by the assessment. Teachers or administrators designing assessments should ground the assessment in the classroom, program or departmental context. The goals or outcomes assessed should lead to assessment data which is fed back to those involved with the regular activities assessed so that assessment results may be used to make changes in practice.

B. Best assessment practice is undertaken in response to local goals, not external pressures. Even when external forces require assessment, the local community must assert control of the assessment process, including selection of the assessment instrument and criteria.

C. Best assessment practice provides regular professional development opportunities. Colleges, universities, and secondary schools should make use of assessments as opportunities for professional development and for the exchange of information about student abilities and institutional expectations.

What is *fair* writing assessment?

Does fair assessment mean that everyone is measured the same?

YARDSTICK UNIVERSE



1. Does consistency yield fairness?

Not necessarily

2. Is standardization unfair?

Not necessarily

1. Does consistency yield fairness? Not necessarily.

Consistency ensures **RELIABILITY**.

But that's not the same thing as fairness.

2. Is standardization unfair? Not necessarily

Standardization allows comparison for basis of judgment.

Standardization becomes a problem when (1) it limits our ability to think beyond standardization (e.g., beyond the rubric) and (2) it leads to unexamined negative consequences.

3. What is unfair?

Assessment that:

- Trivializes the richness of writing while placing inordinate attention on scores
- Devalues linguistic and cultural identity
- Denies personal engagement with writing
- Leads to poor curricular innovation
- Perpetuates racism and xenophobia

What makes for fair writing assessment?

Lack of bias, equitable treatment in the testing process, equality in outcomes of testing, and opportunity to learn.

-Standards for Educational and Psychological Testing, 1999

“A concern for fairness should **guard against any disproportionate social effects** on any language minority group. Writing assessments that are used to make important decisions about individuals and the material and educational conditions that affect these individuals should provide an **equal opportunity** for students to understand the expectations, roles, and purposes of the assessment”

-National Council of Teachers of English

Lack of bias

Equitable treatment

Measurement has focused here

Equality in outcomes of testing

Opportunity to learn

**Guard against disproportionate
social effects**

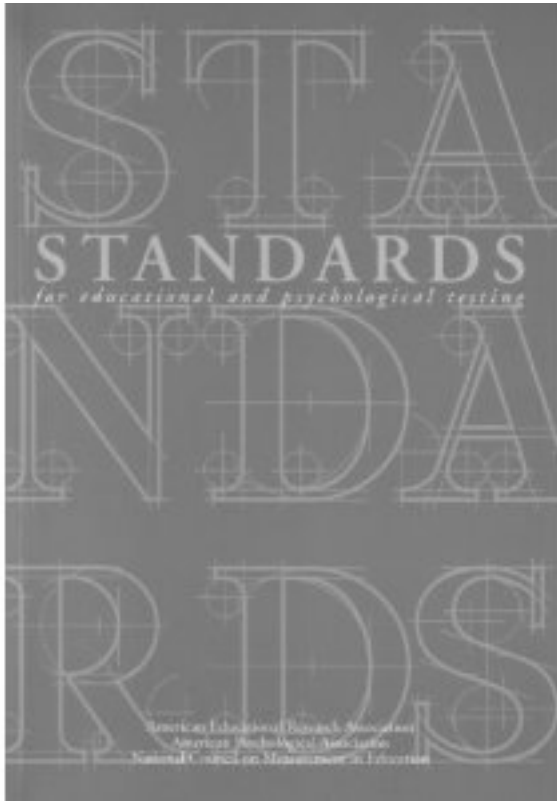
Equal opportunity

Here is where we should focus

What is the connection
between decisions &
consequences?

Validity

The Standards for Educational and Psychological Testing (1999)



Part 1: Test Construction, Evaluation, Documentation

Validity, “degree to which evidence and theory support the interpretation of test scores entailed by proposed uses of the test.” (*Standards*, 1999, p. 9)



Part 2: Fairness in Testing

Fairness in testing and Test Use

The Rights and Responsibilities of Test Takers

Testing Individuals of Diverse Linguistic Backgrounds

Testing Individuals with Disabilities

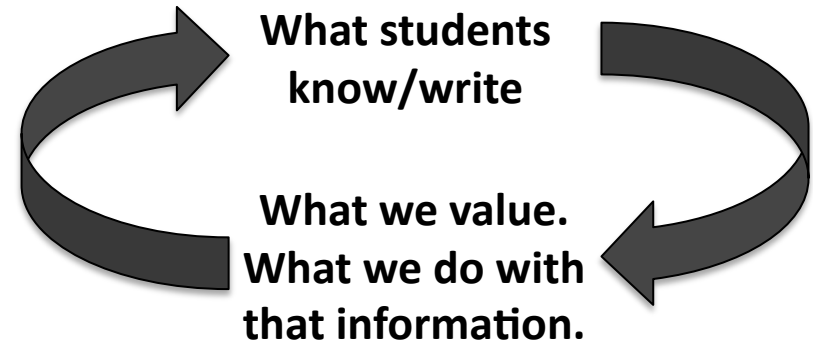
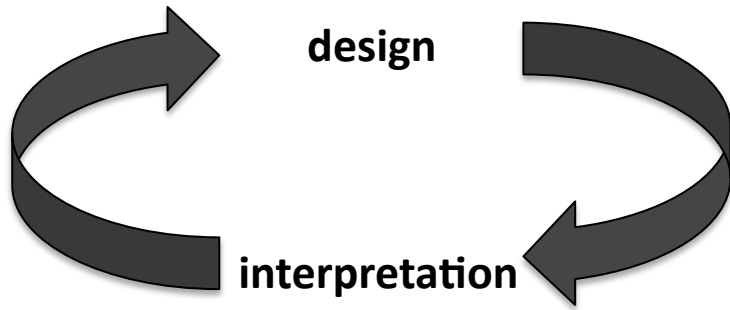
“Concerns about fairness arise from the intended and unintended consequences of testing. Fairness is thus not a property of a test per se, and for this reason, investigations of fairness are framed by test use.” (Camilli, Ed. M., p. 251)

How do we make good decisions?

How do we make good decisions?

- Guillermo Solano-Flores: cultural validity
- Liz Hamp-Lyons: portfolio assessment
- Deborah Crusan: ethnolinguistic bias
- Sara Cushing Weigle: computerized scoring
- Dana Ferris: error analysis
- Suresh Canagarajah: classroom assessment
- David Slomp: consequential validity

- *Journal of Second Language Writing, Assessing Writing, Journal of EAP, Modern Language Journal.*
 - Peer feedback
 - Response to test prompts, test-taking strategies
 - Reviewer feedback
 - Score comparisons
 - Stereotypes and affective dimensions
- **Center for Culturally Responsive Assessment Evaluation**



What consequences do we want?

Fairness is about making good **decisions**.

Decisions that are connected to **consequences**.

Consequences that yield better **outcomes** for students and society.

Educational opportunity

Awesome writing

Novel assessment

Improvement in teaching

Social justice

Disparate Impact Analysis and Writing Assessment: The Case of Brick City

Mya Poe

Northeastern University

John Aloysius Cogan Jr

UCONN School of Law

Norbert Elliot

New Jersey Institute of Technology

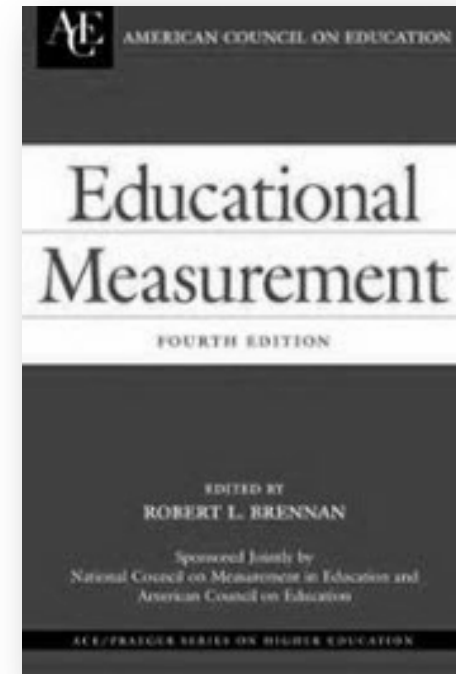
Tito G. A. Nurudeen Jr

Rutgers School of Law-Newark

Poe, M., Elliot, N., Cogan, J.A., and Nurudeen, T. (June 2014). "The Legal and the Local: Using Disparate Impact Analysis to Understand the Consequences of Writing Assessment," *College Composition and Communication*.

Disparate impact analysis in Educational Measurement is not helpful.

1. focuses on federal law and does not take state law into consideration (Davis; Camilli; Pollock; Verdun; Ryan; Kidder and Rosner)
2. contains incorrect statements of the law (Sireci and Parker; Phillips and Camara)
3. fails to consider important legal developments, such as relevant Supreme Court decisions (Sireci and Parker)
4. mentions Title VI or disparate impact with little or no discussion of the applicable legal standards (Elliot et al.; Verdun)



Can disparate impact analysis be used for the purposes of writing program self-assessment?

**The question is not if students need additional support.
The question is if our assessment practices are leading to unintended discrimination.**

Disparate impact discrimination is the unintended racial differences in outcomes resulting from facially neutral policies or practices.

Disparate impact is applicable to ESL and international writers.

Case Study of Brick City, a 4-year, public college

Students Admitted

Group	Number and Percent
African American	108 (13%)
Native American	9 (1%)
Asian	191 (23%)
Hispanic	200 (24%)
White	337 (39%)

Admitted white students (39%) do not outnumber other groups (61%). African American, Asian, and Hispanic students now form the majority of admitted students.

The Brick City writing placement exam remediates many African American and Hispanic students.

Those students graduate at much lower rates than other students. Is remediation creating a barrier to graduation?

Group	Mean Writing Placement Test Scores	Placement	
		Basic Writing Number and Percent	First Year Writing Number and Percent
African American (<i>n</i> =107)	6.86 <i>SD</i> = 1.48	50 (47%)	57 (53%)
Native American (<i>n</i> =9)	7.33 <i>SD</i> = 1.5	2 (22%)	7 (78%)
Asian (<i>n</i> =191)	7.87 <i>SD</i> = 1.41	29 (15%)	162 (85%)
Hispanic (<i>n</i> =200)	7.42 <i>SD</i> = 1.35	57 (28%)	143 (72%)
White (<i>n</i> =337)	7.99 <i>SD</i> = 1.22	35 (10%)	302 (90%)

Brick City College used the “burden shifting” approach to conduct a self-study:

3 steps, according to the Dept. of Education Office for Civil Rights:

- 1. Do the assessment policies or practices result in an adverse impact on students of a particular race as compared with students of other races?**
2. Are the assessment policies or practices necessary to meet an important educational goal?
3. Even in situations where a college or university can demonstrate that the assessment policies or practices are necessary to meet an important educational goal, are there comparably effective alternative policies or practices available that would meet the school’s stated educational goal with less of a burden or adverse impact on the disproportionately affected racial group?
(Department of Education)

Disparate Impact Test #1: The Four-Fifths Rule

= if the outcome of a test on a particular race is less than 80%, or 4/5th, of the rate of effects of test on other races, then there is disparate impact.

	Total Students	White Students	Asian Students	Hispanic Students	Native American Students	African American Students
Total Population	879	337	192	192	192	192
Number of Students Tracked to First Year Writing	700	302	162	162	162	162
Percent of Students Tracked to First Year Writing	80%	90%	85%	72%	78%	53%
Four-Fifths Threshold (.8 x Percentage of White Students Tracked to First Year Writing)	72% (.8 * .9 = .72)					
Four-Fifths Rule Violated?	--	N/A	No	No	No	Yes

Asian students adverse impact ratio: 0.85. Hence, the 4/5 rule is not violated.

Hispanic students adverse impact ratio: 0.72. Hence, the 4/5 rule is violated.

Native American students adverse impact ratio: 0.78. Hence, the 4/5 rule is violated.

African American students adverse impact ratio: 0.53. Hence, the 4/5 rule is violated.

Brick City College used the “burden shifting” approach to conduct a self-study:

3 steps, according to the Dept. of Ed. Office for Civil Rights:

1. Do the assessment policies or practices result in an adverse impact on students of a particular race as compared with students of other races?
- 2. Are the assessment policies or practices necessary to meet an important educational goal?**
3. Even in situations where a college or university can demonstrate that the assessment policies or practices are necessary to meet an important educational goal, are there comparably effective alternative policies or practices available that would meet the school’s stated educational goal with less of a burden or adverse impact on the disproportionately affected racial group? (Department of Education)

How can Brick City articulate the educational goal of the placement exam?

- determining whether the elements of writing that the test measures are important components of student writing ability
- ensuring that the test in fact captures those components
- documenting that the basic writing course remediates the identified writing problems and addresses those problems
- demonstrating that the test is predictive of or significantly correlated with students' performance in college writing.
How do Basic Writing students do in First Year Writing?

Brick City College uses the “burden shifting” approach for self-study:

3 steps, according to the Dept. of Ed. Office for Civil Rights:

1. Do the assessment policies or practices result in an adverse impact on students of a particular race as compared with students of other races?
2. Are the assessment policies or practices necessary to meet an important educational goal?
3. Even in situations where a college or university can demonstrate that the assessment policies or practices are necessary to meet an important educational goal, **are there comparably effective alternative policies or practices available that would meet the school’s stated educational goal with less of a burden or adverse impact on the disproportionately affected racial group?** (Department of Education)

Brick City looks for a method of assessment that results in equally good outcomes for all its students but without the racialized distribution.

- Change the placement exam?
- Keep basic writing but change curriculum?
- Mainstream all students into first year writing?
- Mainstream and add tutors (“move in” approach)?
- Some combination of these?

**Altered the placement exam to align w/curriculum.
Mainstreamed all students and offered additional support through a tutorial approach.**

Civil Rights Laws as Self-Study Heuristic

Legal frameworks are good models for self-study. Question the apparatuses by which students are sorted.

National conversation about justice as fairness. It's not a zero-sum game.

Local conversation about the connection between assessment and curriculum. How can we think boldly about writing assessment?

The Consequences of Writing Assessment: Using Qualitative Research to Understand Fairness in Context

Mya Poe

Northeastern University

Charlie Lesh

Northeastern University

John Walsh

Northeastern University

Kathleen Baldwin

University of Massachusetts Amherst

Linh Dich

Miami University

In my study, I want answers to 4 questions:

- 1. Writing histories:** What are students' histories with writing assessment?
- 2. Writing in college:** Where and what writing assessment do students encounter in college?
- 3. Consequences:** How do assessment practices effect writing development and learning?
- 4. Implications** for teaching and assessment?

Interviewed

Students in the Study
1. Ben
2. Lan
3. Alex
4. Xuyen
5. Menghan
6. Scott
7. Sapna
8. Catherine
9. Miranda
10. Leydrichah
11. Lauren
12. Jacob
13. Cara
14. Sean
15. Vasin

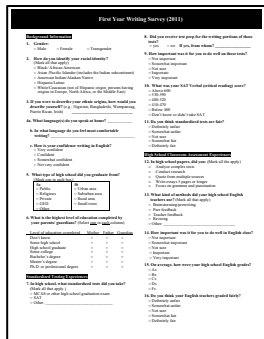
2 international Chinese students

1 int.-born student/immigrant parents (late arriving)
- 1 Haitian

6 U.S.-born child/immigrant parents (early-arriving)
- 1 Spanish lost (Cuban)
- 2 oral only (Thai; Farsi)
- 2 oral/writing (Vietnamese)
- 1 oral/writing mix (Gujarati/English)

6 U.S.-born child/U.S. parents
- 1 Polish lost w/in 1 generation
- 1 Yiddish speaker
- 1 Hebrew speaker/writer
- 3 no immediate multilingual memory (Irish; French-Canadian/Scottish; American)

- Institutional documents
- External data from ETS
- Archival documents
- Interviewed FY teachers
- Surveyed approx. 1,200 students in FYC



Portraits

Sapna / Lan

Milton, Massachusetts

English speaker

“American” (Polish)

Born and raised in Massachusetts, U.S. (Indian)

First language is Gujarati/English mix

SAT (Above 600 Critical Reading score)

Has studied Spanish for 4 years

“I’m not particularly good at English.”

Resource Economics and Marketing

Sapna

Honors student

Lan

Beijing, China

Mandarin speaker

“Chinese” (Han)

Basic Writing student

Born and raised in Beijing, China (Han)

First language is Cantonese

Studied 1 year at U.S. high school (Arizona)

Took SAT in China (410-470 Critical Reading score)

Has studied English for 10 years

“I’m really good at English”

Psychology vs. Business

Mode/Reasoning
Genre
Process

Mode/Reasoning: Chinese
Grammar: English

Writing in High school

Sapna:

My freshman year, my English [honors] class was very **discussion-based**, and there wasn't a whole lot of writing. It was more responses. ... For the most part, it was very open. **Personal responses** mostly, and just discussing. That really helped me, because it allowed me to think and talk, and read things, then talk about them.

My sophomore year... That was the year we read a lot of books. We wrote a ten page **research paper**, and I didn't learn anything that year.

Junior year is the year I learned how to actually write [in] an interdisciplinary class. There was no textbook or anything like that. You would do just readings, come in the next day, talk about the readings, and then ... write an essay on the book.... **I sat down with [my teacher] multiple times, revised my papers, rewrote them**, and I learned how to write from her. I learned how to make a **thesis statement, support it**, and that was the year that I learned how to write.

Lan:

What Chinese teachers valued writing in *Chinese*:

The teacher gives us the question and there's no right or wrong; we need to state our point of view, we need to state our **opinions**. We have to use **examples** from history or from politics to support our point of views.

What Chinese teachers valued writing in *English*

In China, actually, we just normally pay attention to **grammar**. ...we didn't learn all of the context to learn words or phrases and grammar. The grammar is really the concentration in China. But when I come here, I found that grammar is a structure, not really practical [for college reading and writing].

What her U.S. high school teacher valued writing in

***English*:** I think **grammar** was a really important part of it. She just corrected me a lot on informal usage, grammar errors. That's called errors, right? Most of my problems are more about the grammar and informal usage.

Sapna

High School	Admissions	Fall 2011	Spring 2012	Fall 2012	Spring 2013
AP English	SAT	SOM Transitions	Leadership Seminar	Intro. to Business Systems	Inter. Micro Theory
	SAT II	Intro to Phil.	College Writing	Accounting I	Accounting II
	MCAS	Ideas that Changed the World	Social Problems	Inter. Macro Theory	Fundamentals of Marketing
		Calc I for Soc Sci	Intro. to Microeconomics	Intro. to Bus. Law	Consumer in Society
		Intro. Macroeconomics	Math 128 Calc II	Intro to. Stats	Intermediate Stats
			Africa since 1500		

Lan

High School	Admissions	Fall 2011	Spring 2012	Fall 2012	Spring 2013
Sr. Writing	SAT	Basic Writing	College Writing	Intro. Accounting	Statistics
English Lit.	TOEFL	Psych. 101	Nutrition	Intro. Physics	Finance
		Micro Econ.	Intro. to Statistics	Intro. to Law	Int. Econ.
		Art Appreciation	World History	Internet	Env. Econ.
			Mgt. Leadership	French 101	Dev. Psych.
					French

Sapna

**Sapna and Lan are writing
in 50% of classes in their
first 2 years of college**

High S					
AP Eng					
High S					
Sr. Wri					
English					
		Micro Econ.	Intro. to Statistics	Intro. to Law	Int. Econ.
		Art Appreciation	World History	Internet	Env. Econ.
			Mgt. Leadership	French 101	Dev. Psych.
					French

Sapna's writing

"Easy A's" on writing

College Writing

The assignment was to write about yourself, and analyze yourself. It was just like I have no idea, like, how to do this. You have to, like, define yourself, and then write that in an essay form. I was like I don't know how to do this. ... It is different so it's weird sometimes, because, like, I can use "I" in this essay? That's weird... I really liked this first assignment after I figured out how to do it.

History: Africa Since 1500

Identifications and an essay on the book. ... My professor is ..really fair. So as long as you show really good understanding of the book and use examples to support your points. They're not looking for five to 10 paragraphs, perfectly organized, mostly content, which makes sense. ... they're not looking for what all writing test my look like, organization, and stuff like that. As long as it's clear and make sense. It is organized to a point. It's mostly content.

Business Law

Not getting feedback

It's just like doing like legal research... eight page research paper on the case I'm doing. It is a research paper. It'll be the case that I choose. ... 10 page research paper due for my final

Introduction to Statistics

short answers, so like problems that you have to solve. So that on you get feedback, like the TA will always write things.

Intermediate Macro Theory

Reflections ...we're creating a database for the Amherst Survival Center ...so you just reflect and write on [articles about poverty].

Sapna: Process vs. lack of support

Outline: I'll write down a bunch of bullet points and try to make an outline, and then from that I'll write. The way we learned with my junior year [high school] teacher is you write your thesis first, and then based on that, you make your outline and everything, you make sure everything is supported by that. I've always done the whole brainstorming topics, sentences, and trying to organize it. That's my generative writing.



Time management: If it's like a long paper, I'll probably do it in two or three days. ...I end up just probably making an outline in my head writing down thesis, "OK, this is what I'm going to write about." Then, I'll go back to it the next day, refresh what I looked at, start writing it, probably get really distracted over a couple of hours. ... Keep writing it again, then I'll probably finish it on that 2nd day because it's either due the next day or I'll finish it on the 3rd day.



Handing over review: I did [my cover letter] from Googling sample cover letters. I did that, I wrote one, then I went down to Chase Career Center, and I met with one of the people. I was like, "I need help. Read my cover letter. Change it." She changed it. We went back and forth like, "What if you said things like this, instead of this?" Then we drafted a final thing, and that's what I sent out last year to get my internship. This year... I just tweaked it. I never really had anyone look at this one.

Didn't buy books 2 semesters



Lack of feedback on writing

Lack of peer support (Intro to Marketing peer team didn't respond to email or fb requests)

There's a lot of Accounting resources and tutoring help and things like that available. **I never make use of them just because I wouldn't know what questions to ask.**

Lan's writing

Basic Writing*

Reasoning

We need to write an essay about **our own experience related to [the class] articles**. . . .I don't think [the teacher] pays attention to grammar. I think our writings reflect our ability to think critically. . . Grammar is one part, but not the most important part. (12/1/11)

Physics

Lab report: It's not a research paper; it's, like, ... **your process in the experiment and what you found, what you discovered**, like a report, so, it's kind of, yeah, I mean a report out of knowledge, physical knowledge about [the lab project].

Psychology

Reflection papers: We need to **write our thinking**, or reflections towards the papers. It's like our own stories related to this topic. I think psychology is really practical and related to people, so that's really interesting.... I could talk about my opinion; there's no right or wrong.

Intermediate economics and environmental economics

Short answer questions: We don't have any papers, but the style of the exams. We need to state the theory and we need to state the logic, we need the **reasoning** there. A lot of reasoning, a really good way to practice reasoning to, practice analytical skill.

French: [My French teacher] said you have some grammar problem, and I think it's the biggest problem in my French study. And that is I have a kind of poor vocabulary. I need to just enhance my vocabulary...it's really important, like hard in any language writing.

Grammar

Lan: Negotiating grammar vs. critical thinking

What teachers valued in writing in China: critical thinking; it's like the SAT the SAT writing test. The teacher gives us the question and there's no right or wrong; we need to state our point of view, we need to state our opinions. We have to use examples from history or from politics to support our point of views.



History: She did not read the book on grammar. She never did correct the grammar. She found the point, the profile she needed to answer the question. It's like when you provide the right history facts or illustrate your point of view or critical thinking. Critical thinking about reasoning, the point of view of.

What her U.S. high school teacher valued: I think grammar was a really important part of it. Probably, I'm not really English as a first language student. She just corrected me a lot on informal usage, grammar errors. That's called errors, right? Most of my problems are more about the grammar and informal usage.



The value of basic writing for writing in other courses? Not really because [basic writing is] not concentrating on the grammar or word usages. It's just concentrating on the content we need to write about and some sort of examples or some kind of logic things. So it's not really helpful.



French: [My French teacher] said you have some grammar problem, and I think it's the biggest problem in my French study. And that is I have a kind of poor vocabulary. I need to just enhance my vocabulary...it's really important, like hard in any language writing.

Consequences

Sapna

Reflecting on herself as a writer: Because whenever I get a grade in anything, like on a paper or something like that, and it doesn't have a lot of comments, I get really frustrated. Why are there no comments?

Reflecting on herself as a learner: It was just, I guess to do better, and I guess keep up with better study habits, and things like that. If I have something due, do it ahead of time. Don't get as distracted while you're doing your work and things like that. I'm trying to hone in on better study skills, this semester.

Reflecting on external pressures: [My parents] aren't very happy, but they can't really do anything about it besides say, "Oh, we're not happy. You have to do better. You have to make sure you're doing this, this and this." I just have to sit there, and listen and go, "OK."

I am trying to do better, because I want to do better for myself. I, also, need to maintain at least a 3.4 GPA in order to keep scholarships. If I lose those scholarships, then I'm very much screwed.

Lan

Reflecting on herself as a writer: I think reading is interesting. Through reading I can learn more about how to express my ideas. Any writing of any language is from reading, so through the process, through reading, I know that I will improve so the process is really worth it.

I will do a lot of writing like practicing, writing this summer. I really need to do that. I actually have a plan on that. Actually I searched online and I tried to find some proper writing seminars, but there's few.

Reflecting on herself as a speaker: Do you know **TOEFL**, the test? ... It's an exam to test foreign language speakers, how they're level is. It's really fair test, a really, truly reflect your ability. I took this two years ago, and I took this now. I got a higher score.. much better than before.... I just want to see [if] my elocution really get better, did it get better, my English....

Reflecting on grammar: I could argue now that grammar is really important. To formalize our writing, grammar is important, but I think style is no need to be normalized, you know? I think style could be free.

Qualitative research (teacher research) can provide important evidence about more fairly assessing multilingual writers:

- how students grow as writers
- how writing assessment shapes learning
- how we can make assessment more meaningful to students

And how we can use that information to make better decisions about the kinds of educational opportunities that we make available in our programs and classrooms.



Fair Writing Assessment: A Hope for the Future

Fairness as a shared language to talk about assessment

- Critique: Critical analysis of high stakes testing needs to be more nuanced
- Teacher research: How writing assessment works in your classroom and affects your students
- Practice: Bold new ways to assess writing

To make better decisions with consequences that students deserve.

- While Title VI prohibits discrimination on the basis of national origin, language and national origin are not interchangeable (*Mumid v. Abraham Lincoln High School*; *Hannoon v. Fawn Engineering Corp.*; *Soberal–Perez v. Heckler*). A policy that treats students with limited English proficiency differently than other students in the district does not facially discriminate based on national origin. However, in certain circumstances, the failure by a university that receives of federal funding to ensure that persons with limited English proficiency can effectively participate in or benefit from the university’s programs and activities may violate Title VI’s prohibition against national origin discrimination (United States, “Guidance to Federal Financial Assistance”; *Lau v. Nichols*). Title VI also covers to international students attending a university receiving federal funds (*Scarlett v. School of Ozarks, Inc.*).
- - deleted excerpt from “The Legal and the Local: Using Disparate Impact Analysis to Understand the Consequences of Writing Assessment,” forthcoming in *College Composition and Communication*, June 2014. Mya Poe, Norbert Elliot, John Aloysius Cogan Jr., Tito Nurudeen

Disparate Impact Test #2: Chi Squares

The chi-square value and the resulting p-value explain the likelihood that the disparities between placement rates based on race are due to chance.

Group	Total	FY	BW	%FY	%BW	Expected FY	Expected BW	χ^2	p-value
White	337	302	35	90%	10%	267.99	69.01	2.63	p > 0.10
Asian	191	162	29	85%	15%	151.89	39.11		
White	337	302	35	90%	10%	267.99	69.01	29.01	p < 0.001 X
Hispanic	200	143	57	72%	28%	159.04	40.96		
White	337	302	35	90%	10%	267.99	69.01	1.29	p > 0.10
Native American	9	7	2	78%	22%	7.16	1.84		
White	337	302	35	90%	10%	267.99	69.01	69.03	p < 0.001 X
African American	107	57	50	53%	47%	85.09	21.91		