

MA TESOL Program Assessment Report

Analysis of the writing quality of Teaching Philosophy Statements in the SF State MA TESOL Program

RATIONALE

Since 2003, the MA TESOL Program has required that all students produce a professional portfolio as part of their culminating experience requirements. Initially, students were required to include a Position Paper in their portfolio, an 8-10 page document that outlined the students' teaching beliefs and values, supported by references to the empirical/theoretical literature and classroom practice. Although the Position Paper provided students with a valuable opportunity to formally *articulate* their teaching beliefs, in recent years, many of our students and faculty expressed concerns that the Position Paper was (1) not engaging students in the integration of theory, methods, and tools learned in the TESOL curriculum; (2) did not give students the opportunity to generate a scholarly work (pedagogical or research-based) that showcased their individual passions and interests as teachers; and (3) was not an immediately practical document for job search purposes. Based on this feedback, the TESOL program opted in early 2011 to replace the Position Paper and instead require students to submit an original Capstone Project/Paper and a Teaching Philosophy Statement as part of their portfolio requirements.

For this program assessment, the TESOL faculty chose to examine the quality of the Teaching Philosophy Statements produced since this curricular change. The Teaching Philosophy Statement is a writing genre commonly assigned in teacher education programs and required in real-life professional settings (e.g., job applications, tenure dossiers). However, while there is an ample supply of how-to essays, advice columns, and sample statements on the Web, there is surprisingly little literature which systematically examines the genre features of Teaching Philosophy Statements, or the process of writing and assessing statements (see Alexander et al, 2012, for similar commentary).

Our program assessment aimed to answer this question: What is the quality -- in terms of content and writing features -- of our students' MA TESOL Teaching Philosophy Statements? With this assessment, we hoped to strengthen our understanding of the variation in quality of students' statements, and inform our own pedagogy for supporting students' TPS writing process.

WRITING SAMPLE

We based our analysis on a convenience sample of Teaching Philosophy Statements (TPS) written by students in their final semester in the MA TESOL program. We were able to collect 28 final TPS drafts from the Spring 2011 cohort, 15 from Fall 2011, and 12 from Spring 2012. Because the Spring 2012 cohort only included 12 graduates, we chose to randomly select 12 statements from the other two semesters, for a total sample of 36 statements. We removed names from all statements prior to scoring.

About our students

Our graduate program routinely attracts both international and residential students. Our writing sample reflects this demographic, with 25 US-born students who speak English as their first language (hereafter, L1) and 11 multilingual students (includes international F1 visa students) from Japan, Korea, China, Ukraine and Turkey who learned English as a second language (L2).

Table 1. Description of students in sample of Teaching Philosophy Statements (n=36)

Semester	# of L1 English speaking writers	# of multilingual writers
Spring 2011	7	5
Fall 2011	8	4
Spring 2012	10	2
Total	25	11

ANALYTIC APPROACH AND PROCEDURES

We first reviewed existing TPS evaluation rubrics gathered from Education and Applied Linguistics programs at several major universities (University of California-Santa Barbara, University of Pennsylvania, Penn State, University of Michigan, University of Iowa, University of New Mexico). We developed our own rubric, adapting one developed at Penn State University, to assess our sample of statements. As shown in our rubric (See Appendix 1), we chose to focus on 5 attributes of the Teaching Philosophy Statement: (1) logic/clarity of ideas; (2) specificity, referring to the use of evidence, details, and examples to support key ideas, (3) the examination of personal beliefs, (4) quality of reflection, and (5) length.

We rated the first 4 attributes on a Likert scale of 1 to 3, with 1 indicating weak performance, 2 satisfactory performance, and 3 exceptional performance. We chose to rate length of Statements based on an holistic judgement and on a polar-scale (yes/no). This decision was in response to faculty discussions of TPS length, which we found difficult, and even unnecessary, to disentangle from an evaluation of appropriate length from the writer's clarity of ideas, or the substantive content of those ideas. We have typically advised our students to keep their statements to 2 single-spaced pages, less than 1,000 words, although in practice, we have found that effective statements can be shorter or longer than 1,000 words. More discussion about the evaluation of length appears below.

Six MA TESOL faculty participated in the scoring of the statements, with each of the 36 statements scored twice by independent raters. We completed a norming session to train ourselves in use of the rubric; as a result of this norming session, we were able to identify 3 anchor statements, one for each rating level (1, 2, and 3). For data analysis

purposes, we used the average of the raw scores assigned by the 2 raters on the 4 attributes.

Also, as part of the evaluation process (see page 2 of the rubric in Appendix 1), raters selected up to 4 key content areas that were addressed in the statement. This survey of the content areas was not tied to evaluative judgments about what was desirable for a good TPS. Rather, we felt this survey would provide useful information about the ways our students frame themselves in relation to issues taught in our program. Raters selected from a list of 16 content areas commonly found in beliefs statements, based on our anecdotal impressions as well as our knowledge of major issues in the TESOL field (e.g., learner motivation, English teaching in a globalizing world); raters were also given the option to add a content area if a statement addressed a topic not included in the list of 16. This inventory of content areas aimed to provide us with a sense of the commonly occurring themes, as well as the breadth of themes covered, in our students' Teaching Philosophy Statements.

Rater agreement

Across the sample, on average, faculty ratings were in exact agreement for 60% of the 36 statements. As shown in Table 2, faculty ratings were in agreement for 81% of the statements with respect to Length. Scoring agreement on the other four attributes hovered around 50%, ranging from 53% to 57%.

Table 2. Summary of rating agreements

Attribute	# rating agreements (n=36)	% of total ratings
Logic and clarity	18	53%
Specificity	19	53%
Foundations in beliefs about learning and teaching	20	56%
Quality of reflection	17	57%
Length	29	81%
<i>Overall average</i>	<i>20.6</i>	<i>60%</i>

We resolved 5 cases where there was a discrepancy of 2 points in scores. Thus, in our analysis, “rater disagreement” refers to cases when two raters differed by 1-point in their scoring on the 3-point scale. With respect to the 4 attributes related to content and writing quality, there were 28 statements where there was agreement in scores on at least 3 or more attributes. However, there were also 8 statements where there was minimal to zero agreement in scores.

Statements with 0 disagreements = 3
 Statements with 1-2 disagreements = 25

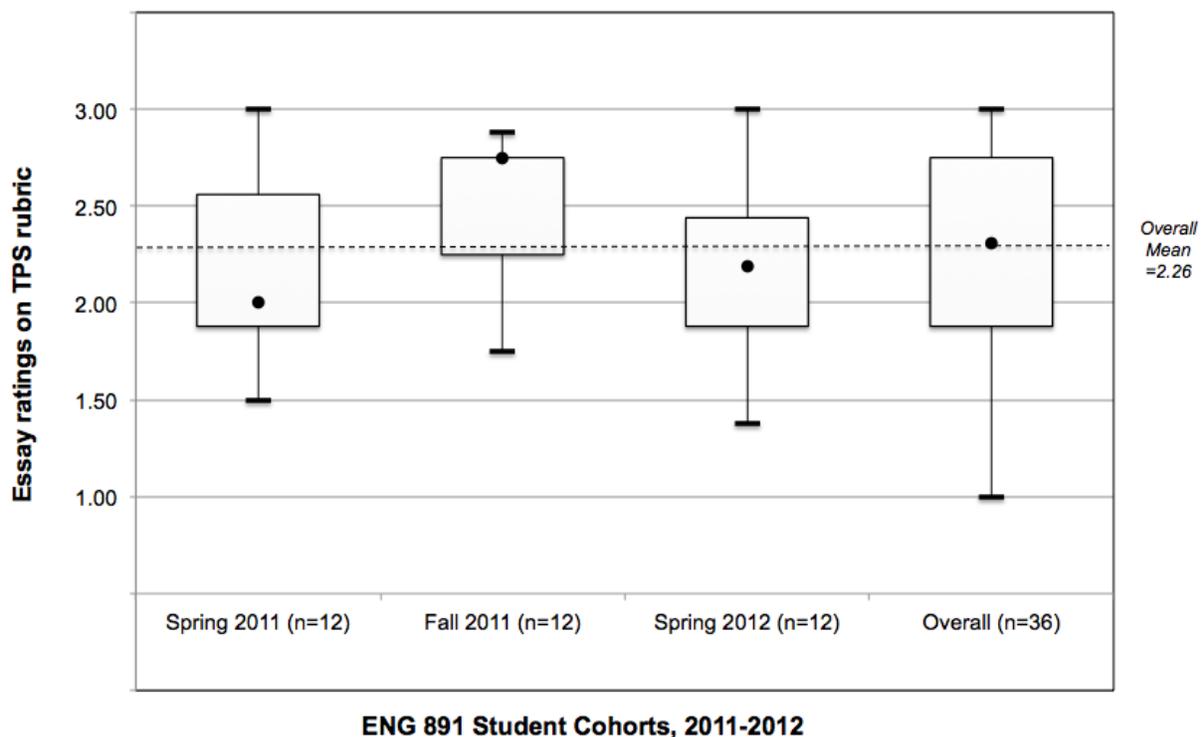
Statements with 3 disagreements = 4
Statements with 4 disagreements = 4

It's important to note that the 8 statements with low agreement were scored low (scores of 1) on at least one of the content/writing attributes, which suggests that these higher rates of disagreement occurred with relatively weaker statements. We will return to this issue in the Implications section.

RESULTS

As shown in Figure 1, the average score in the overall sample of Teaching Philosophy Statements was 2.26 (sd=.51). The average scores for the Spring 2011 cohort (mean=2.20, sd=.51) and the Spring 2012 cohort (mean=2.11, sd=.59) were relatively lower. The average score for the Fall 2011 cohort (mean=2.48, sd=.39) was higher than the overall average, with relatively less variation around the mean compared to other cohorts.

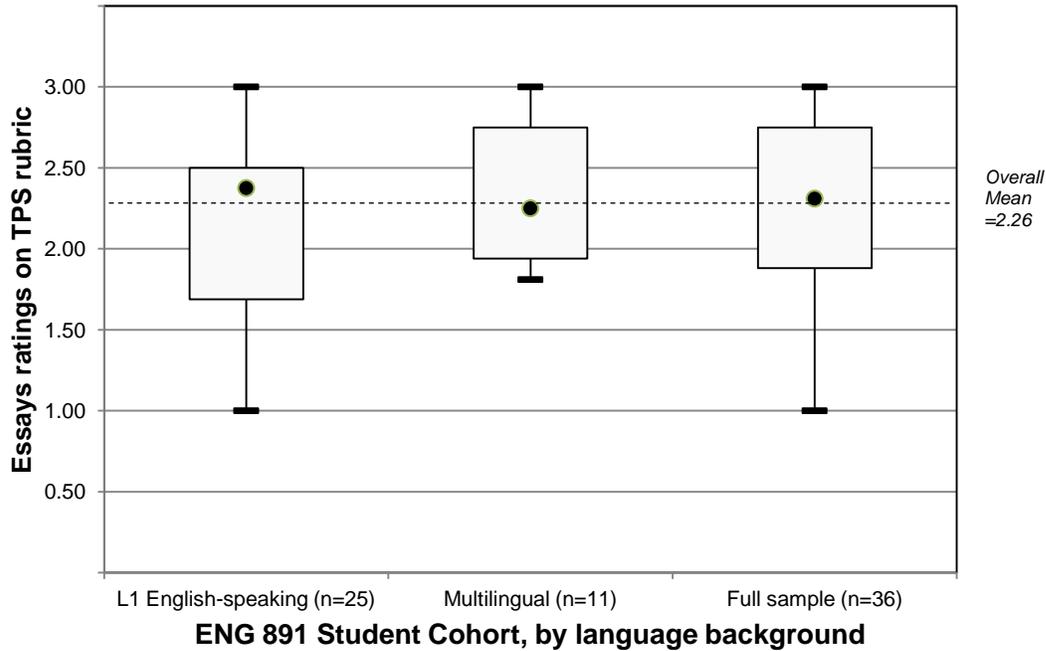
Figure 1. Writing quality of Teaching Philosophy Statements in the MA TESOL Program (n=36). *Dots indicate median scores.*



As shown in Figure 2, on average, the multilingual writers (n=11, mean=2.37, sd=.46) scored higher than the L1-English speaking writers (n=25, mean=2.11, sd=.62). With the caveat that these findings are based on a relatively small sample size, it is notable

that the multilingual writers generated statements that were deemed on par or better than the statements of their L1-English speaking peers.

Figure 2. Writing quality of Teaching Philosophy Statements in the MA TESOL Program, by student language background (n=36). *Dots indicate median scores.*



As shown in Table 3, the Teaching Philosophy Statements in our sample achieved, on average, satisfactory (scores of 2) to exemplary (score of 3) on the 4 Likert-scaled attributes.

Table 3. Writing quality of Teaching Philosophy Statements with respect to 5 attributes (n=36)

Attribute	Spring 2011	Fall 2011	Spring 2012	Overall
Logic and clarity	2.17 (.54)	2.46 (.45)	2.15 (.84)	2.26 (.54)
Specificity	2.08 (.76)	2.50 (.60)	2.06 (.69)	2.22 (.70)
Foundations in beliefs about learning and teaching	2.38 (.43)	2.54 (.50)	2.23 (.67)	2.38 (.54)
Quality of reflection	2.17 (.49)	2.42 (.42)	2.00 (.56)	2.19 (.51)
Length (# of statements that were deemed of appropriate length)	11 out of 12	10 out of 12	9 out of 12	29 out of 12

While these data indicate that our students, overall, generate quality Teaching Philosophy Statements, with respect to the rubric components, the variation in the data also suggests that students may need more explicit guidance and support with various aspects of this genre.

For example, we found relatively more variation in scores associated with specificity in writing, referring to the student's incorporation of specific language and details (examples, evidence, details) to support their ideas about L2 teaching and learning. Although it was beyond the scope of this report to do a close content analysis of the sample, we found that several of the statements that scored a 3 on the specificity attribute began with a brief but detailed description of an actual classroom context, as illustrated by this essay from the Fall 2011 cohort:

Sample Excerpt Teaching Philosophy Statement, Fall 2011 Cohort

There are 19 women perched on children's chairs, crowded around three different tables, and loosely grouped according to their English level. Their children go to elementary school at this site and their babies are being watched in an adjacent classroom. The students are working on writing stories about their experiences on their first day in this country. They will write about the food, the weather, the journey, and how they and their families felt on that day. I will compile all their stories and give each student a book at the end of the term. Adeiba is from Yemen. She comes to class every day-- always early-- and sits right in front of the old rickety chalk board that I wheel in before each class from the kindergarten room next door. The final version of Adeiba's First Day story will read, "*The first day I came to this country I felt ashamed because I was wearing a veil. And when people talked to me I didn't understand. It was the first time I saw women wearing shorts. I felt strange because I don't want my husband to look at other women.*"

I spent 4 years teaching with the San Francisco Community-Based English Tutoring (CBET) Program. The experiences I had with CBET students shaped me both as a person and as a teacher. I learned about parents and immigrants and about the struggles they face as they do everything they can to do well in this country and provide for their children. I learned that the materials and activities need to meet their real needs, and not their needs as perceived or imposed by people whom they have never met.

In comparison, a few statements that scored only a 1 on the specificity attribute did not include vignette-type material; another statement that scored 1 on this attribute began with a metaphor (an introductory paragraph on how I learned to play the cello) but also did not include a vignette about a specific L2 learning context. Given that we did not cross-reference these TPS scores with learner background information, we do not know for certain whether experienced teachers utilized vignette-type material more frequently in their TPS compared to less experienced teachers. At the same time, these results provide useful evidence for teachers – regardless of years of experience – to raise their awareness of the rhetorical value of narrative material and the ways they can use them as an organizing element in their TPS writing.

With respect to the content of Teaching Philosophy Statements, we found a wide range of topics addressed in the sample. Below are the top 10 most frequently occurring content areas:

Table 4. Range of TESOL topics addressed in the sample of Teaching Philosophy Statements (n=36)

Topic	# of statements that addressed this topic
a. Learner-centered teaching	26
b. Creation of a safe learning community; interpersonal relationships	24
c. Learner motivation	19
d. Own experiences as an L2 learner	18
e. Learner autonomy	16
f. Communicative language teaching approaches	13
g. Learning styles and strategies	12
h. Authenticity in L2 teaching	10
i. Learner identity; L2 learning as identity transformation	8
j. Assessment approaches (incl. corrective feedback, testing, evaluation)	7

These themes reflect emphases in our MA TESOL coursework, so their prevalence in the sample is not surprising.

A relatively small number of statements explicitly addressed various pedagogical skill areas (reading/writing, listening/speaking, grammar, vocabulary, pragmatics). Reading/writing pedagogy was discussed in 6 statements, while pedagogy in the other 4 skill areas was discussed in a total of 6 statements. No statement in the sample addressed an increasingly pressing topic the TESOL field, “English teaching in a globalizing world”.

The average number of topics found in individual statements was 5, with some statements incorporating only 2 topics, while other statements made use of up to 8 topics. We did not find that high scoring statements were associated with a higher number of themes, or that low scoring statements relied on a low number of themes. In

other words, the quality of the statements did not appear to be dependent on the number of topics the writer chose to address.

IMPLICATIONS

Our assessment work provides an empirical basis for future discussions among faculty and graduate students about the generic conventions that are associated with effective Teaching Philosophy Statements. Here is a list of possible future research and pedagogical implications:

1. **Rubric development.** Our rubric helps to establish a set of salient criteria for evaluating TPS statements that faculty and students can learn to apply consistently. Given the variation in scores on some attributes (e.g., specificity in content), it will be important to examine our rubric components and the scoring before we make any definitive claims about the variation in quality in TPS essays.

For future assessments, it will be useful to apply more sophisticated tools of inter-rater reliability. While the agreement percentages suggest that we achieved a decent level of consistent, we are still left with several methodological questions:

- Do the limitations in agreement (e.g., the number of disagreements on some areas of the rubric, e.g., specificity) indicate we need to ultimately reach a higher level of consistency across raters? To what degree does this reflect the need for changes in the rubric versus more norming practice?
- Does the fact that the lack of agreement was largely with low-scoring (poorly rated) TPSs mean that we have more work to do to define that end of the scale?
- Do we consider being within one point across raters constitutes agreement between raters? Do we need more agreement to justify claims about our findings?

At the same time, the faculty felt that the current rubric would be a useful tool to share with MA TESOL students as a way to jumpstart discussions about content and formatting expectations for the TPS. We now plan to devote class time, in several courses, to explicit discussions of distinctions across weak, satisfactory, and exemplary performance, so that students will be able to use the rubric to assess their own TPS strengths and weaknesses and make improvements. Particularly for novice teachers, the rubric provides a framework for meeting the expectations of the TESOL discourse community where they hope to launch a career.

2. **Impact on TPS advising.** Although all essays in our sample were final-drafts (and thus presumably were approved by portfolio advisors), there was still notable variation in quality (as measured by our rubric). As a result of this program assessment process, the faculty felt that the rubric and future iterations will help to ensure we are providing consistent feedback and advising to students on their TPS essays.

3. **Anchor Essays.** This assessment process has helped us to identify a sub-set of anchor Teaching Philosophy Statements. After securing permission from the student writers, we will be able to share high-scoring statements with future cohorts so that our students are able to see models of quality statements in the TESOL field.

4. **Future assessment directions.** To further inform the writing pedagogy around Teaching Philosophy Statements in our program, we have identified several topics for future study:

- a. Applying a corpus approach, we could analyze the linguistic/rhetorical moves in these statements (e.g., the moves a writer uses to convey his/her teaching style, or to convey expertise). It would also be useful to explore how high-scoring versus low-scoring statements differ in their use of such moves.
- b. Through diary studies or qualitative interviews, we could explore the factors that influence the development of students' TPS (e.g., what teaching experiences, course readings, theories, etc) over the course of their time in our program.
- c. We would like to better understand how this TPS assessment links to a broader assessment of how our program cultivates students' academic/professional writing skills.

Source

Alexander, P (2012). Teaching with Technology: Remediating the Teaching Philosophy Statement. *Computers and Composition*, 29(1), 23-38.

APPENDIX 1. Teaching Philosophy Evaluation Rubric¹

EVALUATION CRITERIA

	1	2	3
1. Logic and clarity	Paragraphs are not developed, or they lack clear transitions. Overall, the statement lacks focus.	Paragraphs are generally well-developed, but there may be inconsistencies in terms of logic, clarity, paragraph development, flow, or focus.	Paragraphs are exceptionally well-developed with good transitions between them. Overall, the statement is clearly focused and flows logically.
2. Specificity	Ideas are presented in broad, general language and may rely on minimal anecdotal evidence.	Ideas are sufficiently supported by specific language and evidence (details, examples), with some areas under-developed.	Ideas are exceptionally well-supported by specific language and evidence (details, research findings, classroom examples).
3. Foundation in beliefs about teaching and learning	Statement reflects almost no examination of personal beliefs about the complexities of teaching and learning.	Statement reflects examination of personal beliefs about <i>either</i> teaching or learning with some higher-level connections.	Statement clearly reflects examination of personal beliefs about the complexities of teaching and learning.
4. Reflection	Statement expresses little original thought or connection to professional development.	Statement adequately demonstrates original thought and connection to professional development, but a few areas may merit further reflection.	Statement contains much original thought demonstrating exceptionally strong connection to professional development and meaningful reflection.

5. **Length: Is the length of the TPS appropriate and effective given its purpose and content?** YES NO

¹ Chas Brua Schreyer, Institute for Teaching Excellence, Penn State University, www.schreyerinstitute.psu.edu, Sept 2011

Content

Check up to 4 content areas addressed in the philosophy statement:

- 1. Own experiences as an L2 learner
- 2. Learning styles and strategies
- 3. Learner motivation
- 4. Learner autonomy
- 5. Learner identity; L2 learning as identity transformation
- 6. Learner-centered teaching
- 7. Communicative language teaching approaches
- 8. Skill development (*indicate emphasis*)
 - ___ Reading/Writing
 - ___ Speaking/Listening/Pronunciation
 - ___ Grammar
 - ___ Pragmatics
 - ___ Vocabulary
- 9. Authenticity in L2 teaching
- 10. Assessment approaches (incl. corrective feedback, testing, evaluation)
- 11. Creation of a safe learning community; interpersonal relationships
- 12. Creation of a linguistically rich learning environment (attention to input, interaction, output)
- 13. English teaching in a globalizing world
- 14. English teaching and technology
- 15. L2 teaching as social action; social inequities in L2 teaching
- 16. Teacher identity; L2 teaching as identity transformation
- 17. OTHER: _____