# How Faculty Attitudes and Expectations toward Student Nationality Affect Writing Assessment

**Peggy Lindsey, University of Dayton, and Deborah Crusan, Wright State University**

**Abstract:** Earlier research on assessment suggests that even when Native English Speaker (NSE) and Non-Native English Speaker (NNES) writers make similar errors, faculty tend to assess the NNES writers more harshly. Studies indicate that evaluators may be particularly severe when grading NNES writers holistically. In an effort to provide more recent data on how faculty perceive student writers based on their nationalities, researchers at two medium-sized Midwestern universities surveyed and conducted interviews with faculty to determine if such discrepancies continue to exist between assessments of international and American writers, to identify what preconceptions faculty may have regarding international writers, and to explore how these notions may affect their assessment of such writers. Results indicate that while faculty continue to rate international writers lower when scoring analytically, they consistently evaluate those same writers higher when scoring holistically.

Crusan (2010) has argued that "All teachers, consciously or unconsciously, hold biases about virtually every aspect of the workings of their classroom" (p. 89). In this context, bias is not a pejorative term, but rather an instructor's "individual-agenda, or discipline-based preference" (p. 88). One instructor may favor group work while another prefers lecture; one discipline may tend to value individual creativity while another emphasizes collaboration. These biases extend to how different faculty members define good writing: a study by Crusan (2001), for example, revealed that medical faculty rank grammatical correctness as one of the most important features of good writing whereas conciseness and clarity were paramount for business faculty (p. 92). (For other studies of discipline-based bias, see Brown, 1991; Mendelsohn & Cumming, 1987; Roberts & Cimasko, 2008; Santos, 1988; Song & Caruso, 1996; Weigle et al., 2003.)

While all student writers may find adapting to the expectations of a particular discipline or instructor challenging, Non-Native English Speaker (NNES) writers, (that is, writers whose first language is not English), often struggle more than their native-speaker peers to understand and meet faculty expectations for good writing (Crusan, 2010, p. 91. See also Leki, 2006). NNES writers who struggle with grammatical correctness will not fare well if the faculty member tends toward what Crusan (2010) refers to as a left-brain assessor, one who "focus[es] more on mechanical aspects of testing and rating and emphasize[s] the logical continuity of thought and mechanical aspects of writing" (p. 91). NNES students with a good mastery of grammar may still struggle to meet teacher expectations

if their notion of how to organize an argument varies substantially from their instructor's. Connor (2002) notes that "the linear argument preferred by native English speakers may well represent what such speakers view as coherent, though speakers of other languages may disagree" (p. 497). A study by Kobayashi & Rennert (1996) evaluated 465 readers with different cultural backgrounds and concluded that "culturally influenced rhetorical patterns affected assessment of EFL [English as a Foreign Language] students" (p. 397). An NNES writer from a culture that values an indirect approach, with the point coming at the end of his paper, may find his work downgraded by a professor who expects the thesis to appear early in the writing.

Faculty may reduce the negative impact their biases may have on student grades and comprehension simply by making their preferences about what constitutes good writing explicit to students. Providing assessment criteria with an assignment, for example, offers a useful means for faculty to clarify any biases they have regarding mechanics and/or content (Crusan, 2010; Ferris & Hedgcock, 2005; Valdez-Pierce, 2000). Such a step benefits all students, not only NNES writers, by fronting what factors a faculty member will weight most heavily in assessment. Faculty can also take steps to consider the presence of cultural bias in their materials, that is, the extent to which a faculty member or a discipline privileges a particular way of learning or knowing as superior (Tyler, Stevens & Uqdah, 2009).

But what about any hidden or unconscious biases a faculty member may hold? How might these affect assessment? As Clark (2010) notes, "Teachers view themselves as 'good' and 'fair' people, and they are skeptical about being biased," but the millions of Implicit Association Tests, or IATs, administered since 2002 by psychologists for Project Implicit reveal that not only are implicit biases pervasive, but that people are unaware of them: "Ordinary people, including the researchers who direct this project, are found to harbor negative associations in relation to various social groups (i.e., implicit biases) even while honestly (the researchers believe) reporting that they regard themselves as lacking these biases" ("General Information," 2008). It stands to reason that at least some faculty might possess what we, for lack of a better term, refer to as ethnolinguistic bias: the tendency of faculty to evaluate students' ability with written language less or more favorably due to positive or negative bias triggered by markers, such as nationality.

## Does Ethnolinguistic Bias Exist?

Studies suggest that even when educators are cognizant of student diversity, their ability to translate that knowledge into new teaching and assessment practices has often moved slowly or not at all. Clair (1995) identified three studies that revealed that awareness does not necessarily change practice or attitude:

> Sleeter (1992) found that although many of the participating teachers [in a multicultural education program] perceived that they had learned much, there was little change in their attitudes and practice. Ahlquist (1992) noted that teacher attitudes and beliefs remained unchanged for the most part during a multicultural foundations course. . . . McDiarmid (1990) studied teachers' attitudes toward ESL students both before and after a 3-day workshop designed to influence these attitudes and found that the multicultural presentations had little influence on the teachers' beliefs about ESL students. (p. 193)

Such findings concur with a review of more recent studies by Tyler, Stevens & Uqdah (2009) that suggest evidence of cultural bias in teaching throughout the academy: "In addition to cultural bias found throughout public school curricula and standardized testing, cultural bias is believed to be

salient throughout the institutional practices promoted and executed by school teachers and administrators" (p. 293). Ndura's analysis of ESL instructional materials, for example, revealed that ESL textbooks "fail to reflect the growing diversity of students' life experiences and perspectives" (2004, p. 150). This finding parallels studies of textbooks used in Native English Speaker (NES) classrooms where "most contributions to academic subject matter . . . are made by members of the majority race or culture . . . and much of the text throughout this subject matter is used to reinforce the superiority of this group" (Tyler, Stevens, & Uqdah, 2009, p. 292). Studies by Boykin et al. (2006) and Tyler, Boykin, & Walton (2006) both concluded that many classroom teachers are biased in favor of students whose behavior reflects mainstream cultural values, such as competition and individualism, rather than those who demonstrate alternative ethnocultural values such as communalism. Teachers tended to see students who followed the mainstream values as more motivated and higher-achieving.

In terms of bias specifically in writing assessment—whether in first year composition or in courses across the curriculum—such perceptions can run deep despite efforts to train faculty across the curriculum to be more aware of how a student's linguistic and cultural background affects how they complete assignments. Scholarship in second language acquisition has repeatedly indicated that "second language acquisition is a slow and gradual process and that expecting ESL students' writing to be indistinguishable in terms of grammar from that of their NES [Native English Speaker] counterparts is naïve and unrealistic" (Silva, 1997, p. 362). Across the curriculum, many faculty may be unaware of this scholarship, or simply be uninterested because they don't perceive writing instruction as their responsibility (Salem & Jones, 2010). Rather than break down assignments into structured steps that guide students in developing the strategies required for success, many faculty across the curriculum continue to simply assign a paper and collect the final product with no discussion of how one moves from assignment to successful finished essay in a particular discipline. Even those explicitly trained to assess writing may continue to fall back on traditional methods that emphasize product above all else: "composition specialists have long suspected that many teachers, although they publicly eschew a focus on the final product of writing and celebrate process-centered writing instruction as excellent pedagogy, still practice current-traditional rhetoric in the classroom" (Crusan, 2010, p. 93).

Writing assessment studies since the 1980s have also repeatedly found that teachers modify their assessment strategies when grading NNES work. A 1982 review of 12 studies about Native English Speaker reactions to NNES writing concluded that, while faculty were generally able to successfully comprehend the message being conveyed, readers were significantly irritated by NNES errors. Vocabulary errors which occurred with grammatical mistakes were considered the most annoying (Ludwig, 1982). How such irritation affected final assessment, however, was variable, and seemed to be affected by multiple factors (Ludwig, 1982, p. 281). Studies since Ludwig have continued to suggest that educators tend to rate NNES writers differently from their NES peers (see Kobayashi, 1992; Lee, 2009; Kobayashi & Rennert, 2001; Milnes & Cheng, 2008). Huang's (2009) review of 20 major empirical studies on ESL writing assessment concluded that "[r]ater background, mother tongue, previous experience, amount of prior training, and types and difficulty of writing tasks have been found to affect the rating of written responses of ESL students" (p. 1). Huang cites an early study by McDaniel suggesting that "raters weight or emphasize the criteria differently while rating compositions written by ESL and NE [Native English] students" (2009, p. 6). Studies since McDaniel concur with this finding: Milnes & Cheng's 2008 comparison of evaluations of ESL vs. non-ESL students at a private Canadian high school revealed that "most participants modified their assessment strategies when marking the work of ESL students" (p. 49). While many of these studies offer suggestions for improvement in NNES writing assessment, research suggests that even those

most highly trained in working with NNES writers may still evaluate NNES writers differently: not only are raters likely to be influenced in varying degrees by previous assessment experiences and practices, "it may be relatively difficult for experienced raters to 'unlearn' or shift themselves away from particular criteria or procedures that they have become skilled at using" (Cumming, Kantor & Powers, 2002, p. 71).

While NNES students may not always be aware of the more unconscious forms of ethnolinguistic bias (any more than their instructor is) in the grading of their written work, these students report awareness that they are treated differently. Whether the difference translates into higher or lower grades, students express frustration at being underestimated: that their struggles with language lead to their work being discounted and to perceptions that their ability to handle the work is limited. As one student reported, "The academic skills of students who are not native speakers of English are not worse than academic skills of American students. In some areas it can be much better. Just because we have problems with language . . . that some professors hate . . . doesn't mean that we don't understand at all" (Zamel, 2004, p.9). Studies of basic writers attest to how a faculty member's belief that a particular type of student (such as a basic writer or an NNES writer) does or does not possess the ability to handle course material can lead to false judgments when applied to individual students. When that belief translates into faculty applying different criteria—either easier or harder—when grading, a student's ability to learn may be undermined. In some cases this takes the form of lowered expectations; in others it translates into faculty who simply take less time with students they perceive as less capable. (Zamel, 2004). In all cases, it suggests that faculty may demonstrate bias when evaluating students from outside the ethnolinguistic mainstream.

## The Role of Assessment Criteria in Creating Ethnolinguistic Bias

Indeed, various studies of NNES, L2 and ESL writers—those with ethnolinguistic backgrounds different from their American-born NES peers—suggest that faculty may assess these students differently. Huang (2009) notes that holistic scoring (that is, grading based on the overall effectiveness of a communication rather than on mechanical correctness) seems especially biased when used with NNES writers. Ironically, despite the fact that holistic scoring de-emphasizes mechanics, Russikoff suggested that raters tended to focus on language use as primary criteria for evaluation when rating holistically (Huang, 2009, p. 7). Since this is often a weak area for NNES writers, their scores suffered. As argued by Huang (2009), "When the same raters rated the same ESL compositions analytically [that is, by evaluating specific areas such as content and organization separately] they were surprised to see how strong the 'content' and 'organization' of these compositions were" (p. 7). Given that Ludwig (1982) discovered that evaluator irritation with errors in vocabulary and grammar were significant factors in NNES assessment, it seems logical that holistic graders could over-emphasize the effect of mechanical issues in their evaluations. Indeed, a study by Nairn (2003) revealed that non-ESL faculty treat NNES grammar errors more severely. Using Lane & Lange's (1999) classifications for categorizing "global ESL-type errors, local ESL-type errors, and other errors, which are also made by native speakers," Nairn discovered that non-ESL English department faculty tended to be least tolerant of ESL-types of errors: "Only 36.4% of errors that are also made by native speakers were marked as unacceptable in college-level writing, whereas 45.5% of the global ESL-type errors and 51.5% of the local ESL-type errors were marked" (Nairn, 2003). Faculty from the humanities and social sciences had a similar reaction: "They tended to overlook the errors that are made by both native speakers and ESL students more frequently than they overlooked the ESL-type errors" (Nairn, 2003). Although Nairn's study had a very limited scope, the findings parallel Ludwig's (1982) and Crusan's (2010) discoveries about rater irritation with NNES errors and would likely be duplicated in a larger study: it makes sense that readers are more likely to note

errors that they see less often or are unlikely to make themselves. Nairn argues that because faculty are less tolerant of ESL-type errors, they "perceive that the ESL students' writing contains a larger number of more serious sentence-level errors than they see in the writing of the native speaker students" (Nairn, 2003). Faculty already resentful about teaching writing in their courses may feel particularly bothered by such errors (Salem & Jones, 2010, p. 70-71). Even if such errors don't prevent an evaluator from comprehending the writer's message, the mistakes simply stand out more clearly, creating a larger impact in a holistic score.

It is important to note, however, that the impact is not always negative. Faculty demonstrate varying degrees of tolerance toward NNES errors. Some may find such errors irritating, but consciously choose not to be affected by them, and instead "bend over backwards and make extra allowances for NNS's composition difficulties" (Rubin & Williams-James, 1997, p. 141). Others will grade more harshly because they fear developing any kind of double standards that would allow an NNES writer with more errors to earn the same grade as a native speaker without such errors (Nairn, 2003).

Such findings suggest that simply using analytic rather than holistic scoring[1] would reduce ethnolinguistic bias by ensuring that faculty maintain a focus on writing content rather than language error. Huang's review and Nairn's study both suggest that language errors disproportionately affect holistic scores. Nairn (2003) noted that the perception of more sentence-level errors made faculty "more likely to develop a negative view of the ESL students' writing abilities." Yet holistic scoring methods offer the benefit of focusing on the overall quality of a communication, rather than simply scoring separate components. It is, arguably, a more authentic form of assessment in that it evaluates writing based on how well it communicates the author's ideas to the reader, not on how well it succeeds at individual components such as thesis statement, organization or logic. And given that several studies in Huang's review used analytic scoring and still discovered ethnolinguistic bias, a shift away from holistic grading appears unlikely to eliminate the problem. Having worked with faculty who spoke of feeling dread at the thought of reading a paper simply because they believed its author to be an ESL writer, not because of any firsthand knowledge of the student's writing, we wondered if, regardless of whether holistic or analytical assessment were used, faculty were at all predisposed to find more errors in NNES writing or to be less tolerant of errors of any type made by NNES writers.

## Faculty Presuppositions toward NNES Writers

To explore this idea, we held a faculty development workshop on working with ESL writers where we presented six faculty with brief writing samples by four students. We asked them to not only mark any errors they saw, but to identify which sample they believed was written by a native speaker of English. Errors were not intruded into the samples, which were of varying quality, but all four contained errors common to both ESL and native speakers as well as those likely to be made only by ESL writers. Not a single participant was able to correctly identify the NES writer. Perhaps most troubling was that while the participants unanimously identified the same sample as the strongest example of good writing, none of them anticipated that that sample was written by an NNES writer.

The extremely small scope of this workshop means that the results should not in any way be seen as conclusive, but it certainly raised questions for us about the possibility of faculty having an unconscious bias toward writers from a different ethnolinguistic background simply because they know them to be from a different ethnicity and assume they are NNES. Our findings supported other studies which indicated that "when raters do *not* know the language background of writers [...] similarly prepared NNS [non-native speakers] and native English-speaking (NES) writers are evaluated similarly by ESL and mainstream composition teachers alike" (Rubin & Williams-James,

1997, p. 141) and that contact with nonnative speakers of English may influence teachers' rating of their writing (Crusan, 2010; Roberts & Cimasko, 2008).

The most complete study we found on the subject (Rubin & Williams-James, 1997) was more than ten years old and surveyed 33 graduate teaching assistants from departments of English at four different universities. That study asked participants to evaluate six papers in which six types of errors were intruded. Each essay was assigned a nationality from the US, Denmark or Thailand, so that raters read two essays from each nationality. Readers were also given TOEFL scores for the Danish and Thai writers, creating a strong implication that these students were NNES. The nationalities were rotated "to avoid confounding essay text with nationality" (Rubin & Williams-James, 1997, p. 144). The study concluded that NES and NNES writers were rated differently: "ratings of NNS writing were best predicted by the number of surface errors they detected. Ratings of NES writing, in contrast were justified by marginal notations and comments" (Rubin & Williams-James, 1997, p. 139). Yet, interestingly, the study did not discover a bias *against* NNES writers: on the overall quality of the composition, "the Asian NNS writers were judged superior to the U.S." (Rubin & Williams-James, 1997, p. 148). Such a finding suggests that ethnolinguistic bias may not be based solely in NNES writing errors, but that a faculty member's ideas and attitudes about a particular culture or ethnicity may also play a role in assessment.

Because the study by Rubin and Williams-James was published 14 years ago, we wondered whether their results would be the same today. In particular, we were interested in how much a faculty member's judgments about a student's linguistic ability are colored by assumptions he or she makes based on the student's nationality, ethnicity, and culture. In particular, we wondered whether simply knowing a student's country of origin could make a difference in how a faculty member assesses a student.

## Methods

### Participants

In order to more thoroughly explore the current impact of nationality on rating students' writing and to discover if ethnolinguistic bias, either positive or negative, might still be a factor in writing assessment, we adapted the methods used in Rubin and Williams-James (1997) and surveyed faculty from across the curriculum at two and four-year institutions. Three versions of a survey including a scale for rating student essays were posted on Survey Monkey (http://www.surveymonkey.com/). Each version included the same demographic questionnaire (see Appendix A), essays (see Appendix B) and rating scale (see Appendix C). All participants were given the essays in the same order. The only variation in the surveys was the student profile assigned to the essays. Participants had the option to provide written comments after evaluating each essay and at the end of the survey. We used electronic mail to contact faculty from across the curriculum and ask them to complete the survey.

### Demographic Overview

In all, 87 participants finished the survey: 73 native speakers of English (NSE) along with nonnative speakers of five other languages (Bengali, German, Japanese, Romanian, and Spanish). Respondents also reported several other dialects of English including British and Canadian English and Ebonics. Participants reported experience studying 30 different languages with French and Spanish being the most common; 34 professed fluency in at least one language other than English. Again French and Spanish topped the list. Among the respondents on each survey, at least one person had traveled to each region of the world listed in the survey. Europe was the most popular destination with over

three-quarters having lived or traveled there. About one-quarter of all respondents reported travel to East Asia, Mexico, and Latin America, the regions of the world on which this study focuses. The fact that so many respondents have studied, visited, and/or lived in a foreign culture suggests that this group is open to diversity.

The majority of respondents were from the humanities (61%) and social sciences (18%), as shown in Table 1. Despite direct appeals, we were unable to generate any responses in Math or Nursing; however, we had some participation from Fine Arts (2%), Business (5%), Science (5%), Psychology (2%) and Engineering (1%). Five respondents (or 6 %) did not identify their primary teaching area.

**Table 1: Distribution of Participants by Discipline**

| Discipline | Survey A | Survey B | Survey C | Total | % of Total |
|---|---|---|---|---|---|
| Fine Arts | 2 | 0 | 0 | 2 | 2 |
| Humanities | 10 | 23 | 20 | 53 | 61 |
| Social Sciences | 3 | 5 | 8 | 16 | 18 |
| Business | 3 | 0 | 1 | 4 | 5 |
| Science | 2 | 1 | 1 | 4 | 5 |
| Math | 0 | 0 | 0 | 0 | 0 |
| Psychology | 0 | 0 | 2 | 2 | 2 |
| Engineering/Computer Science | 1 | 0 | 0 | 1 | 1 |
| Nursing | 0 | 0 | 0 | 0 | 0 |
| No answer | 1 | 2 | 2 | 5 | 6 |
| Total | 22 | 31 | 34 | 87 | 100 |

The large percentage of humanities faculty in our sample pool could potentially skew our results in ways that a set of participants more evenly distributed across the curriculum would not. The preponderance of humanities and social science faculty also helps to explain why more than one-third of participants reported some background in TESOL or linguistics since these areas would encompass faculty who work with language as teachers of composition, foreign language, and linguistics.

Respondents reported substantial teaching experience, with the average participant having at least 12 years in the classroom. Nearly all of them include writing in their courses: only one participant reported not requiring any writing. The most frequent types of assignments required were essays of fewer than five pages (80%) and in-class writing (65%). 68% of the respondents reported having WAC training in the form of graduate course work or faculty development workshops. Fewer faculty

had training in teaching ESL students. Only 35% reported any kind of training working with ESL writers: 28% had graduate course work in TESOL and 9% had attended faculty development workshops in ESL. Most, however, had taught ESL students. Only 10% of respondents reported never teaching ESL students. Other respondents averaged approximately 12 students per year.

## Essays

After faculty completed demographic questions on the survey, they were then presented with six student writing samples consisting of the student's name and a short paragraph written by the student in response to the study prompt. Essays for this study were written in a required first year composition class at a medium-sized Midwestern university. Students were native speakers of English who wrote to the following prompt:

*Many parents have begun to worry that young adults today spend too much time using social networking technologies such as Facebook and texting at the expense of face-to-face interaction.*

*In response, the university has proposed that students participate in a voluntary social media boycott for one week.The feeling is that less time spent on technology will mean more time to become involved in other activities on campus. Write a letter to your peers in which you attempt to convince them to participate or not participate in the experiment.*

Students wrote responses as an in-class exercise, so they were typically shorter than outside assignments; most essays were between 125 and 175 words in length. Using the rubric and evaluation criteria (see Appendix C), the researchers chose six medium-quality essays to use in the study, three supporting the boycott and three against. While brief, we believe the texts offered enough information for a basic evaluation of a student's ability to communicate an idea in writing. First year composition teachers frequently use short writing exercises as a starting point for longer writing tasks (Roberts & Cimasko, 2008). The social networking topic for the short essay in this study might possibly serve as a very rough draft of a longer argument either for or against boycotting social networking or for discussing pros and cons of social networking in general.

## Error Intrusion

Vann, Meyer, and Lorenz (1984) categorized common ESL learner errors and perceived error gravity. Using their work and the error selection process of Rubin and Williams-James (1997), we intruded the following surface-level errors into the six native speaker essays:

- word order reversals
- wrong tense
- wrong preposition
- lack of pronoun/antecedent agreement
- minor spelling errors (mainly homonym confusions), and
- omitted articles.

Replicating the methods of Rubin and Williams-James (1997) and the work of Connors and Lunsford (1988), we intruded errors at the rate of five errors per 100 words. Each essay contained the same types of errors. As in the Rubin and Williams-James study, we left the first and last sentence of each essay error-free and roughly scattered the others throughout the remainder of the essays. A copy of the six papers with intruded errors can be found in Appendix B.

We intruded only surface-level errors into the writing samples because such errors in our estimation are typically the most visible distinction for many readers when identifying NNES writing and therefore can serve as a marker of ethnolinguistic difference. They are also the types of errors most easily agreed upon as correct or incorrect and therefore less prone to variations in faculty preferences regarding style and rhetoric. We recognize that other rhetorical differences often exist between NNES and NES writing such as ways of organizing information, understanding the role of the reading in creating meaning from the text, and approaching argument. While survey participants were asked to rate rhetorical features, in this study, all of these features were kept as the original NES author had intended. This allowed us to focus on how much common NNES surface-level errors may affect ethnolinguistic bias.

## Fabricated Writer Names

**Table 2: Survey Distribution for Survey A**

| Sample # | Position | Gender | Name | Nation |
|---|---|---|---|---|
| 1 | Con | Male | Jesus Gonzalez | Mexico |
| 2 | Con | Male | Brandon Douglass | USA |
| 3 | Con | Male | Zhou Ming | China |
| 4 | Pro | Female | Ana Martinez | Colombia |
| 5 | Pro | Female | Katie Brekinridge | USA |
| 6 | Pro | Female | Chie Miyagi | Japan |

**Table 3: Survey Distribution for Survey B**

| Sample # | Position | Gender | Name | Nation |
|---|---|---|---|---|
| 1 | Con | Female | Katie Breckinridge | USA |
| 2 | Con | Female | Chie Miyagi | Japan |
| 3 | Con | Female | Ana Martinez | Colombia |
| 4 | Pro | Male | Brandon Douglass | USA |
| 5 | Pro | Male | Zhou Ming | China |
| 6 | Pro | Male | Jesus Gonzalez | Mexico |

**Table 4: Survey Distribution for Survey C**

| Sample # | Position | Gender | Name | Nation |
|---|---|---|---|---|
| 1 | Con | Male | Zhou Ming | China |
| 2 | Con | Female | Ana Martinez | Colombia |
| 3 | Con | Female | Katie Breckinridge | USA |
| 4 | Pro | Female | Chie Miyagi | Japan |
| 5 | Pro | Male | Jesus Gonzalez | Mexico |

| 6 | Pro | Male | Brandon Douglass | USA |
|---|-----|------|------------------|-----|

Although all of the essays were written by native speakers of English, in order to investigate bias, we attributed different nationalities (through assignment of names associated with Mexican, Columbian, Chinese, Japanese, and American students) to the essays. We opted to focus on Asian and Latin American nationalities because a majority of international students entering U.S. universities today come from these regions ("International Student," 2010). In choosing our names for the Asian and Latin American students, we did an internet search to find the most common names from the selected countries. To create the names for American students, we simply chose two first names common among today's undergraduates and two Anglo-American last names. The name of the writer was repeated many times during essay rating and participants were asked to identify the student by name when writing any optional comments. By using names common to each of the chosen countries and repeating the names frequently, we hoped, like Rubin and William-James (1997), to "strengthen the perceptual salience of these constructed ethnolinguistic identities" (p. 144).

Before reading each sample, participants were informed of the name, age, gender, and country of origin of the writer. For example, a participant might see "This sample was written by Zhou Ming, an 18-year-old male student from China." Rubin and William-James (1997) also provided TOEFL scores for the non-American students, creating a perception that these were NNES writers. In the informed consent portion of our survey, we notified participants that our purpose was to examine faculty attitudes about writing by students who are nonnative speakers of English. We did not identify the writers as specifically international students, residents, immigrants, or generation 1.5. Because the study by Rubin and William-James (1997) provided raters with both nationality and evidence of the student's ESL proficiency, it was impossible to distinguish whether biases among raters were due to expectations based on a student's TOEFL scores or attitudes about the student's nationality. In designing our study, our purpose was primarily to investigate how much knowing only a student's nationality influences assessment. By not explicitly identifying students from Asia and Latin America as L2 international visa students or immigrants, we left it to the faculty to draw their own conclusions, thus making it easier to detect how much nationality, rather than language proficiency scores, colors assessment. Using the same six essays, we distributed names in the ways shown in Tables 2-4.

## Focus Group

In addition to the online survey, we also invited six survey takers to participate in a focus group to further discuss preconceptions that faculty have regarding second language writers. The group consisted of faculty from humanities, social science, business and education. Half the group was American-born, native speakers of English; half was foreign-born, but all had spoken English since childhood.

## Results

Research has repeatedly shown that discrepancies continue to exist between assessments of NES and NNES writers. Our study focuses on determining whether identifying a writer as an international student contributes to such discrepancies. Our survey results suggest they do, and that such discrepancies may be due solely to the perceived ethnolinguistic identity of the writer rather than because of any measurable differences in the writing itself.

The first five items in the rating scale asked respondents to rate the sample analytically on a scale from "Much better than most first year students" to "Below average for a first year student." We converted the four choices to a 4.0 scale where "Much better" = 4 and "Below average" = 1. The next three items asked respondents to evaluate holistically by simply assigning a letter grade. We also converted these responses to a 4.0 scale with A=4 and F=0. The question asking for a subjective evaluation of the student as a writer was converted similarly to the analytic questions with "Very good writer" = 4 and "bad writer" = 1. The question about intelligence had only three choices. "Very intelligent" = 4 and "Not very intelligent" = 2. This effectively converted the data to the scale used at most universities for calculating grade point average.

The ratings for Essay #1 do not demonstrate considerable variability in terms of ratings for each item, but reveal significant patterns regarding ethnolinguistic bias (Table 5). For example, the ratings for Survey A, B, C for the student's ability to convey the overall point clearly were 2.50, 2.63, and 2.47 (or roughly C+/B-). In Survey A, the essay writer was identified as Jesus Gonzalez from Mexico; in Survey B, the writer of the identical essay was identified as Katie Breckinridge, an American; in Survey C, the writer of the identical essay was identified as Zhou Ming, from China. In evaluating the samples analytically—that is, by judging elements such as clarity, organization, evidence and MUGS (mechanics, usage, grammar and style) separately—Survey B respondents consistently ranked the writer identified as American higher than respondents to Surveys A and C ranked the same writing sample. The sample by the Asian male was consistently rated lowest. Yet when rating the sample holistically (Table 6), Survey B respondents rated the U.S. female lower than other survey takers rated the students from other countries for the same sample. The Asian student, identified as a Chinese male for respondents to Survey C, was rated highest in 3 out of 5 holistic rankings.

**Table 5: Essay #1 Analytic Score Ratings**

| Essay #1: Analytic scores | Survey A: Jesus Gonzalez | Survey B: Katie Breckinridge | Survey C: Zhou Ming | Highest ranking | Lowest ranking | Variation between high & low |
|---|---|---|---|---|---|---|
| Ability to convey overall point clearly | 2.50 | 2.63 | 2.47 | U.S. female | Asian male | .16 |
| Ability to present ideas in a clearly organized manner | 2.60 | 2.74 | 2.43 | U.S. female | Asian male | .31 |
| Ability to develop major points with appropriate, convincing supporting evidence | 2.74 | 2.69 | 2.57 | Latin American male | Asian male | .17 |
| Mastery of the conventions of written English (mechanics, usage, grammar, style) | 3.00 | 3.15 | 2.97 | U.S. female | Asian male | .18 |

| | | | | | | |
|---|---|---|---|---|---|---|
| Overall quality of the writing | 2.85 | 3.15 | 2.73 | U.S. female | Asian male | .42 |
| **Average score** | **2.74** | **2.87** | **2.63** | **U.S. female** | **Asian male** | **.25** |

**Table 6: Essay #1 Holistic Score Ratings**

| Essay #1: Holistic scores | Jesus Gonzalez | Katie Breckinridge | Zhou Ming | Highest ranking | Lowest ranking | Variation between high & low scores |
|---|---|---|---|---|---|---|
| Overall grade for content | 2.55 | 2.39 | 2.65 | Asian male | U.S. female | .26 |
| Overall grade for MUGS | 2.30 | 2.07 | 2.23 | Latin American male | U.S. female | .23 |
| Overall grade for entire sample | 2.60 | 2.32 | 2.48 | Latin American male | U.S. female | .28 |
| Assessment of writer | 2.4-0 | 2.25 | 2.70 | Asian male | U.S. female | .45 |
| Assessment of intelligence | 3.16 | 2.93 | 3.37 | Asian Male | U.S. female | .44 |
| **Average overall grade scores** | **2.48** | **2.26** | **2.45** | **Latin American male** | **U.S. female** | **.26** |
| **Average all holistic scores** | **2.60** | **2.39** | **2.69** | **Asian male** | **U.S. female** | **.33** |

The possibility that these results are merely due to the respondents to Survey C being more severe analytic graders and more generous holistic graders is negated by the fact that these same patterns emerge in the scores for Essays #2-6. In Essay #2 (Table 7), respondents to Survey A had the American student identified as the author and rated him higher when evaluating analytically than respondents to Surveys B and C rated the Asian and Latin American students. Yet, once again, the American student suffered when rated holistically (Table 8), receiving the lowest ranking on all five assessments. Once again, the Asian student, this time identified as a Japanese female for takers of Survey B, received the highest ranking three out of five times.

**Table 7: Essay #2 Analytic Score Rankings**

| Essay #2: Analytic scores | Survey A: Brandon Douglass | Survey B: Chie Miyagi | Survey C: Ana Martinez | Highest ranking | Lowest ranking | Variation between high & low scores |
|---|---|---|---|---|---|---|
| Ability to convey overall point clearly | 2.90 | 2.78 | 2.93 | Latin American female | Asian femal | .15 |
| Ability to present ideas in a clearly organized manner | 3.15 | 2.81 | 2.86 | U.S. male | Asian female | .29 |
| Ability to develop major points with appropriate, convincing supporting evidence | 3.25 | 3.07 | 3.18 | U.S. male | Asian female | .18 |
| Mastery of the conventions of written English (mechanics, usage, grammar, style) | 3.35 | 3.33 | 3.07 | U.S. male | Latin American female | .28 |
| Overall quality of the writing | 3.20 | 3.07 | 3.04 | U.S. male | Latin American female | .16 |
| **Average score** | **3.17** | **3.01** | **3.02** | **U.S. male** | **Latin American female** | **.16** |

**Table 8: Essay #2 Holistic Score Rankings**

| Essay #2: Holistic score | Survey A: Brandon Douglass | Survey B: Chie Miyagi | Survey C: Ana Martinez | Highest ranking | Lowest ranking | Variation between high & low scores |
|---|---|---|---|---|---|---|
| Overall grade for content | 1.80 | 2.21 | 2.12 | Asian female | U.S. male | .41 |
| Overall grade for MUGS | 1.80 | 1.96 | 1.86 | Asian female | U.S. male | .16 |

| Overall grade for entire sample | 1.75 | 2.19 | 2.03 | Asian female | U.S. male | .44 |
|---|---|---|---|---|---|---|
| Assessment of writer | 1.85 | 2.13 | 2.14 | Latin American female | U.S. male | .29 |
| Assessment of intelligence | 2.70 | 3.03 | 3.09 | Latin American female | U.S. male | .39 |
| **Average overall grade scores** | **1.78** | **2.12** | **2.00** | **Asian female** | **U.S. male** | **.34** |
| **Average all holistic scores** | **1.98** | **2.03** | **2.25** | **Asian female** | **U.S. ale** | **.32** |

This pattern continues in Essay #3 where, when evaluators score analytically (Table 9), once again the sample identified as being written by an American is consistently ranked higher than the sample identified as written by an Asian or Latin American. When asked to rate the samples more holistically (Table 10), the American student receives the lowest rankings. This is significant because by the time all survey takers had read Essay #3, they had all had the opportunity to read a sample identified as being written from each one of the world regions in the study—Asia, Latin America and the United States—with the U.S. sample ranking highest each time. Thus, the results are not due to one particular group of survey takers favoring or not favoring a particular group. When given the opportunity, all three groups favored the U.S. writer when rating analytically and favored international writers when rating holistically.

## Table 9: Essay #3 Analytic Score Rankings

| Essay #3: Analytic scores | Zhou Ming | Ana Martinez | Katie Breckinridge | Highest ranking | Lowest ranking | Variation between high & low scores |
|---|---|---|---|---|---|---|
| Ability to convey overall point clearly | 2.70 | 2.59 | 2.96 | U.S. female | Latin American male | .37 |
| Ability to present ideas in a clearly organized manner | 2.70 | 2.67 | 3.11 | U.S. female | Latin American female | .44 |
| Ability to develop major points with appropriate, | 2.74 | 2.56 | 3.11 | U.S. female | Latin American female | .55 |

| | | | | | | |
|---|---|---|---|---|---|---|
| convincing supporting evidence | | | | | | |
| Mastery of the conventions of written English (mechanics, usage, grammar, style) | 3.55 | 3.26 | 3.37 | Asian male | Latin American female | .29 |
| Overall quality of the writing | 3.10 | 2.70 | 3.15 | U.S. female | Latin American female | .45 |
| **Average score** | **2.96** | **2.76** | **3.14** | **U.S. female** | **Latin American female** | **.42** |

**Table 10: Essay #3 Holistic Score Rankings**

| Essay #3: Holistic scores | Zhou Ming | Ana Martinez | Katie Breckinridge | Highest ranking | Lowest ranking | Variation between high & low scores |
|---|---|---|---|---|---|---|
| Overall grade for content | 2.75 | 2.43 | 1.89 | Asian male | U.S. female | .86 |
| Overall grade for MUGS | 1.50 | 2.07 | 1.64 | Latin American female | Asian male | .57 |
| Overall grade for entire sample | 2.20 | 2.36 | 1.89 | Latin American female | U.S. female | .47 |
| Assessment of writer | 2.15 | 2.29 | 2.03 | Latin American female | U.S. female | .26 |
| Assessment of intelligence | 3.25 | 3.12 | 3.05 | Asian male | U.S. female | .20 |
| **Average overall grade scores** | **2.15** | **2.29** | **1.24** | **Latin American female** | **U.S. female** | **.63** |

| Average all holistic scores | 2.37 | 2.45 | 2.10 | Latin American female | U.S. female | .47 |
|---|---|---|---|---|---|---|

The evaluations for Essay #4 continue this pattern for analytic scoring (Table 11). The U.S. male, Brandon Douglass, earns higher scores; the Asian female earns the lowest scores. When evaluated holistically (Table 12), the pattern continues for the highest rankings with the Asian female scoring the highest ratings for all but one question. For the first time, however, an international student –the Latin American female—earned the majority of the lowest scores when rated holistically. Yet, while raters assign the lowest overall grade to Ana Martinez of Colombia, the U.S. male was still scored lower than his international counterparts when evaluated on his overall ability as a writer and his intelligence. This continues the pattern of raters perceiving higher overall ability and intelligence in students from other countries and rating them more highly when grading holistically even as they grade their written work lower than their U.S. counterparts when scoring analytically. Indeed, while the Asian female earned the highest holistic scores in most areas, the Latin American female—who earned the lowest overall grade scores—earned the highest score for perceived intelligence. While it is beyond the scope of this study to explore gender bias in detail, this result suggests that such bias may also be a factor.

**Table 11: Essay #4 Analytic Score Rankings**

| Essay #4: Analytic scores | Ana Martinez | Brandon Douglass | Chie Miyagi | Highest ranking | Lowest ranking | Variation between high & low scores |
|---|---|---|---|---|---|---|
| Ability to convey overall point clearly | 2.35 | 2.44 | 2.33 | U.S. male | Asian female | .11 |
| Ability to present ideas in a clearly organized manner | 2.35 | 2.48 | 2.19 | U.S. male | Asian female | .29 |
| Ability to develop major points with appropriate, convincing supporting evidence | 2.40 | 2.44 | 2.37 | U.S. male | Asian female | .07 |
| Mastery of the conventions of written English (mechanics, usage, grammar, style) | 2.95 | 2.96 | 2.78 | U.S. male | Asian female | .18 |

| Overall quality of the writing | 2.56 | 2.67 | 2.35 | U.S. male | Asian female | .32 |
|---|---|---|---|---|---|---|
| **Average score** | **2.52** | **2.60** | **2.40** | **U.S. male** | **Asian female** | **.14** |

**Table 12: Essay #4 Holistic Score Rankings**

| Essay #4: Holistic scores | Ana Martinez | Brandon Douglass | Chie Miyagi | Highest ranking | Lowest ranking | Variation between high & low scores |
|---|---|---|---|---|---|---|
| Overall grade for content | 2.60 | 2.61 | 2.71 | Asian female | Latin American female | .11 |
| Overall grade for MUGS | 2.20 | 2.29 | 2.36 | Asian female | Latin American female | .16 |
| Overall grade for entire sample | 2.40 | 2.50 | 2.50 | Asian female/U.S. male (tie) | Latin American female | .10 |
| Assessment of writer | 2.55 | 2.46 | 2.80 | Asian female | U.S. male | .34 |
| Assessment of intelligence | 3.35 | 3.26 | 3.33 | Latin American female | U.S. male | .09 |
| **Average overall grade scores** | **2.40** | **2.47** | **2.52** | **Asian female** | **Latin American female** | **.12** |
| **Average all holistic scores** | **2.61** | **2.62** | **2.74** | **Asian female** | **Latin American female** | **.16** |

The results for Essays #5 and #6 reinforce earlier results: The U.S. students consistently earn the highest ratings when scored analytically (Table 13 and Table 15) and the lowest ratings when scored holistically (Table 14 and Table 16); the students from other countries consistently earn the lowest ratings when scored analytically and the highest ratings when scored holistically.

**Table 13: Essay #5 Analytic Score Rankings**

| Essay #5: Analytic scores | Katie Breckinridge | Zhou Ming | Jesus Gonzalez | Highest ranking | Lowest ranking | Variation between high & low scores |
|---|---|---|---|---|---|---|
| Ability to convey overall point clearly | 2.40 | 2.08 | 2.33 | U.S. female | Asian male | .32 |
| Ability to present ideas in a clearly organized manner | 2.45 | 2.15 | 2.41 | U.S. female | Asian male | .30 |
| Ability to develop major points with appropriate, convincing supporting evidence | 2.58 | 2.12 | 2.48 | U.S. female | Asian male | .46 |
| Mastery of the conventions of written English (mechanics, usage, grammar, style) | 2.65 | 2.62 | 2.81 | Latin American male | Asian male | .19 |
| Overall quality of the writing | 2.53 | 2.19 | 2.48 | U.S. female | Asian male | .34 |
| **Average score** | **2.52** | **2.23** | **2.50** | **U.S. female** | **Asian male** | **.32** |

**Table 14: Essay #5 Holistic Score Rankings**

| Essay #5: Holistic scores | Katie Breckinridge | Zhou Ming | Jesus Gonzalez | Highest ranking | Lowest ranking | Variation between high & low scores |
|---|---|---|---|---|---|---|
| Overall grade for content | 2.45 | 2.93 | 2.74 | Asian male | U.S. female | .48 |
| Overall grade for MUGS | 2.35 | 2.51 | 2.30 | Asian male | Latin American male | .21 |

| | | | | Latin American male | U.S. female | .08 |
|---|---|---|---|---|---|---|
| Overall grade for entire sample | 2.60 | 2.63 | 2.68 | Latin American male | U.S. female | .08 |
| Assessment of writer | 2.50 | 2.78 | 2.56 | Asian male | U.S. female | .28 |
| Assessment of intelligence | 3.16 | 3.34 | 3.19 | Asian male | U.S. female | .18 |
| **Average overall grade scores** | **2.47** | **2.70** | **2.57** | **Asian male** | **U.S. female** | **.26** |
| **Average all holistic scores** | **2.61** | **2.84** | **2.69** | **Asian male** | **U.S. female** | **.25** |

**Table 15: Essay #6 Analytic Score Rankings**

| Essay #6: Analytic scores | Chie Miyagi | Jesus Gonzalez | Brandon Douglass | Highest ranking | Lowest ranking | Variation between high & low scores |
|---|---|---|---|---|---|---|
| Ability to convey overall point clearly | 2.25 | 2.12 | 2.70 | U.S. male | Latin American male | .58 |
| Ability to present ideas in a clearly organized manner | 2.25 | 2.23 | 2.59 | U.S. male | Asian female | .36 |
| Ability to develop major points with appropriate, convincing supporting evidence | 2.25 | 2.15 | 2.78 | U.S. male | Latin American male | .53 |
| Mastery of the conventions of written English (mechanics, usage, grammar, style) | 2.75 | 2.72 | 3.00 | U.S. male | Latin American male | .28 |
| Overall quality of the writing | 2.33 | 2.23 | 2.70 | U.S. male | Latin American male | .49 |

| Average score | 2.37 | 2.29 | 2.75 | U.S. male | Latin American male | .45 |

**Table 16: Essay #6 Holistic Score Rankings**

| Essay #6: Holistic scores | Chie Miyagi | Jesus Gonzalez | Brandon Douglass | Highest ranking | Lowest ranking | Variation between high & low scores |
|---|---|---|---|---|---|---|
| Overall grade for content | 2.75 | 2.89 | 2.54 | Latin American male | U.S. male | .35 |
| Overall grade for MUGS | 2.35 | 2.48 | 2.29 | Latin American male | U.S. male | .19 |
| Overall grade for entire sample | 2.75 | 2.78 | 2.40 | Latin American male | U.S. male | .38 |
| Assessment of writer | 2.80 | 2.74 | 2.40 | Asian female | U.S. male | .40 |
| Assessment of intelligence | 3.40 | 3.26 | 3.14 | Asian female | U.S. male | .26 |
| **Average overall grade scores** | **2.62** | **2.72** | **2.41** | **Latin American male** | **U.S. male** | **.31** |
| **Average all holistic scores** | **2.81** | **2.83** | **2.55** | **Latin American male** | **U.S. male** | **.32** |

## Discussion

The variation among the scores assigned, regardless of nationality, may seem small, but could potentially be quite significant when looked at in terms of a student's grade. When scoring analytically, the average variation between the highest and lowest ratings—that is, the scores given to the American students compared to their peers from other countries—was 0.29. If one thinks of our 4-point rating scale in terms of a student's GPA, that amount roughly constitutes the difference between a 3.0 or B average and a 3.3 or B+. As a grade on a single paper in a single course, the consequences of the bias are probably non-existent. But our results suggest that students from other countries subjected to such a discrepancy throughout a course will probably earn a lower grade if the professor consistently employs analytic scoring methods simply because the student's nationality has marked him or her as ethnolinguistically different. The reverse is true when students are rated holistically. Here the average variation for overall grade scores is 0.32 in favor of students from other countries.

Perhaps most significant is the marked variation between analytic and holistic scoring. While students from other countries consistently earned the lowest rankings when scored analytically (Table 17), they just as consistently earned the highest ranking when scored holistically (Table 18).

**Table 17: Analytic Score Rankings by Gender and Country of Origin**

| # OF ANALYTIC SCORES EARNED | Male: Highest ranking | Male: Lowest ranking | Female: Highest ranking | Female: Lowest ranking |
|---|---|---|---|---|
| U.S | 14 | 0 | 11 | 0 |
| Asia | 1 | 10 | 0 | 9 |
| Latin America | 2 | 4 | 1 | 7 |

**Table 18: Holistic Score Rankings by Gender and Country of Origin**

| # OF HOLISTIC SCORES EARNED | Male: Highest ranking | Male: Lowest ranking | Female: Highest ranking | Female: Lowest ranking |
|---|---|---|---|---|
| U.S | 1 | 12 | 0 | 13 |
| Asia | 13 | 1 | 5 | 0 |
| Latin America | 6 | 1 | 5 | 3 |

This finding contradicts studies from the 1980s in which evaluators asked to rate ESL writers holistically focused on language use and were unable to recognize strengths in content and organization until they were asked to rate the writing analytically (Huang, 2009, p. 7). In our study, the opposite occurred: analytic ratings for writers from other countries were lower compared to their American peers; holistic ratings were higher. When scored holistically, the U.S. students' scores suffer, indicating a bias in favor of non-native students. Possibly raters may expect Americans to outperform international student writers and, consequently, when asked to make general judgments (as opposed to evaluation of particular components in an essay) score the U.S. students lower when they do not meet rater expectations. Faculty may also be responding more positively to institutional rhetoric about globalization than they might have 20 or 30 years ago when the number of students from other countries was much lower and their benefit to the university much less widely understood.

This does not address the issue of why the students with Anglo-American names consistently scored better when rated analytically. We considered the possibility that variations in the amount of training in teaching writing to NNES students could be a factor in explaining these results. Despite our attempts to randomly solicit responses, it is clear that there is a difference in the amount of ESL training in our groups of survey takers. 60% of Survey C respondents reported having some kind of ESL training while only 19% responding to Survey A reported ESL training, and only 28% of those responding to Survey B. This variation can be addressed somewhat by looking at the results of each survey instead of comparing results across surveys. In Survey A (Table 19), for example, respondents

rated sample 2 (identified as written by the American male, Brandon) highest followed by the Chinese male. Overall, the American female, Katie, rated more highly than her female international counterparts. In this case, the raters were reading different essays, so arguably the different ratings could be based in the actual content of the text, not in any ethnolinguistic bias. Nevertheless, the results suggest a bias toward American writers. Since this group had the least amount of training in ESL assessment, if the surveys with more highly trained evaluators did not show this bias, we could conclude that the result is related to ESL training.

**Table 19: Ratings for Survey A**

| Survey A | Sample 1: Jesus Gonzalez | Sample 2: Brandon Douglass | Sample 3: Zhou Ming | Sample 4: Ana Martinez | Sample 5: Katie Breckinridge | Sample 6: Chie Miyagi |
|---|---|---|---|---|---|---|
| Ability to convey overall point clearly | 2.50 | 2.90 | 2.70 | 2.35 | 2.40 | 2.25 |
| Ability to present ideas in a clearly organzed manner | 2.60 | 3.15 | 2.70 | 2.35 | 2.45 | 2.25 |
| Ability to develop major points with appropriate, convincing supporting evidence | 2.74 | 3.25 | 2.74 | 2.40 | 2.58 | 2.25 |
| Mastery of the conventions of written English (mechanics, usage, grammar, style) | 3.00 | 3.35 | 3.55 | 2.95 | 2.65 | 2.75 |
| Overall quality of the writing | 2.85 | 3.20 | 3.10 | 2.56 | 2.53 | 2.33 |

In fact, this was not the case. While the raters in Survey C (Table 20)—the group with the highest amount of training in ESL—also rated Sample 2 fairly high (suggesting that it is indeed one of the better writing samples), their highest ratings went to Katie, the American, and they paralleled the patterns found in Survey A in that the American female outranked her international counterparts, and Brandon Douglass, the American male, outranked his counterparts.

**Table 20: Ratings for Survey C**

| Survey C | Sample 1: Zhou Ming | Sample 2: Ana Martinez | Sample 3: Katie Breckinridge | Sample 4: Chie Miyagi | Sample 5: Jesus Gonzalez | Sample 6: Brandon Douglass |
|---|---|---|---|---|---|---|
| Ability to convey overall point clearly | 2.47 | 2.93 | 2.96 | 2.33 | 2.33 | 2.70 |
| Ability to present ideas in a clearly organzed manner | 2.43 | 2.86 | 3.11 | 2.19 | 2.41 | 2.59 |
| Ability to develop major points with appropriate, convincing supporting evidence | 2.57 | 3.18 | 3.11 | 2.37 | 2.48 | 2.78 |
| Mastery of the conventions of written English (mechanics, usage, grammar, style) | 2.97 | 3.07 | 3.37 | 2.78 | 2.81 | 3.00 |
| Overall quality of the writing | 2.73 | 3.04 | 3.15 | 2.35 | 2.48 | 2.70 |

The respondents to Survey B (Table 21)—28% of whom had training in ESL—continued this trend:

**Table 21: Ratings for Survey B**

| Survey B | Sample 1: Katie Breckinridge | Sample 2: Chie Miyagi | Sample 3: Ana Martinez | Sample 4: Brandon Douglass | Sample 5: Zhou Ming | Sample : Jesus Gonzalez |
|---|---|---|---|---|---|---|
| Ability to convey overall point clearly | 2.63 | 2.78 | 2.59 | 2.44 | 2.08 | 2.12 |
| Ability to present ideas in a clearly organzed manner | 2.74 | 2.81 | 2.67 | 2.48 | 2.15 | 2.23 |
| Ability to develop major points with appropriate, | 2.89 | 3.07 | 2.56 | 2.44 | 2.12 | 2.15 |

| | | | | | | |
|---|---|---|---|---|---|---|
| convincing supporting evidence | | | | | | |
| Mastery of the conventions of written English (mechanics, usage, grammar, style) | 3.15 | 3.33 | 3.26 | 2.96 | 2.62 | 2.72 |
| Overall quality of the writing | 3.15 | 3.07 | 2.70 | 2.67 | 2.19 | 2.23 |

The bias toward writers with Anglo-American names does not appear to depend on the amount of training a responder has in ESL assessment.

We then considered whether the preference for the Asian and Latin American writers in holistic writings may be due to the nature of our survey pool, the majority of whom had experience with foreign languages and/or travel. Faculty who have struggled to write in another language or live abroad may feel particularly sympathetic to the challenges international writers face and make an effort to "counter stereotypes by softening or reducing criticism" (Roberts & Cimasko, 2008, p. 137). This interpretation of sympathetic bias for international writers was supported by the comments of our focus group where participants expressed primarily positive attitudes about ESL students of any type (international, refugee, immigrant, etc.) and an inclination to tolerate errors they would not accept from NES writers. Even though the faculty saw similarities between NNES writers and underprepared NES students in that both groups struggle with similar grammatical issues, they admitted that they may be more lenient when grading ESL writers: "As a teacher, I think I'm probably more generous in terms of some of the writing," noted one participant. The group believed ESL students to be intelligent and brave. One participant noted that, "It boggles my mind that they would try to do a degree in a second language when it was difficult enough in your first. I had a positive attitude toward them in terms of them being courageous people." This perception appears to lead faculty to slightly inflate the grades of NNES writers simply because they are NNES writers. Such inflations may be easier to make when grading holistically.

If our focus group can be taken to be at all indicative of the overall whole, our raters may be likely to show bias in favor of international writers when grading holistically because they perceive these students as particularly intelligent and hard-working under difficult conditions. We also asked raters to assess how intelligent they perceived the writers to be as a means of detecting hidden biases about the intelligence of particular ethnic groups. Zamel (2004) notes that some faculty who work with NNES writers may confound language use with intellectual ability: "'bad language' and 'insufficient cognitive development' were being conflated" (p. 5). Our findings (Table 22), however, supported the perceptions of the focus group that NNES writers are generally highly intelligent. Students identified as American, rather than the non-native writers, were generally perceived as less intelligent than their international peers. In Essay #4, the Latin American female, despite receiving the lowest holistic rankings overall, earned the highest score for assessment of intelligence. On Essay #3, the Asian male received the highest scores for intelligence despite a low ranking for mechanics and grammar. Significantly, this sympathetic bias does not extend to American students in holistic grading. An American student earned the lowest rankings on intelligence on all six essays.

**Table 22: Assessment of Intelligence**

| ASSESSMENT OF INTELLIGENCE | Sample 1: Jesus Gonzalez | Sample 2: Brandon Douglass | Sample 3: Zhou Ming | Sample 4: Ana Martinez | Sample 5: Katie Breckinridge | Sample 6: Chie Miyagi |
|---|---|---|---|---|---|---|
| Survey A | 3.10 | 2.70 | 3.25 | 2.35 | 3.16 | 3.40 |
| Survey B | 2.93 | 3.03 | 3.12 | 3.26 | 3.34 | 3.26 |
| Survey C | 3.37 | 3.09 | 3.05 | 3.33 | 3.19 | 3.14 |

This suggests that some ethnolinguistic bias may be due to faculty consciously wanting to acknowledge and accommodate the challenges faced by international writers. Unfortunately, this recognition appears to come at the expense of the American student who may be facing equal challenges with his or her writing.

The low numbers of faculty participants from disciplines outside the humanities made it impossible for us to draw any conclusions about bias within specific disciplines. Indeed our survey was designed to identify bias across the curriculum rather than within particular academic areas. However, the high concentrations of humanities faculty led us to consider whether Survey B, where 80% of respondents were from the humanities, and Survey C, where 63% of respondents were from the humanities, would reveal any significant differences in evaluations when compared to Survey A, where 47% of respondents were from the humanities. Because the number of respondents varied for each survey, we converted the total number of responses to each question to a percentage of the total responses to that question within each survey. We then compared the percentages for each response from Surveys A, B, and C to establish whether any patterns emerged. We did not find any significant patterns that would justify concluding that the humanities-dominant Survey B respondents were statistically more or less likely to demonstrate bias than the respondents to Survey A (where fewer than half of respondents were from the humanities).

For example, Table 23 (below) details the responses participants gave when asked to rate an essay's overall grade for mechanics, usage, grammar and spelling. As the table reveals, in many instances, the percentage of respondents choosing a particular response was identical or nearly so. When larger discrepancies (which we defined as discrepancies greater than 10% between one survey group and the group with next highest percentage) do occur, they are not consistently caused by a particular group of survey takers. At times, the respondents are nearly equal in their ratings; in other instances, any one of the survey samples might be higher or lower. For example, on Sample 6, the same percentage of participants from surveys A and B chose an "A" grade as their response. In that same sample, the same percentage of participants from surveys B and C chose "B" as their answer. In Sample 3, Survey B takers chose "C" as their response more than 30% more often than respondents to Surveys A and C. In Sample 4, however, Survey C takers were more than 30% more likely than respondents in Surveys A and B to choose "C."

**Table 23: Respondents by Percentage Evaluating Mechanics, Usage, Grammar and Spelling**

| Overall grade for MUGS | % of respondents who chose A | % of respondents who chose B | % of respondents who chose C | % of respondents who chose D | % of respondents who chose F | Nationality assigned to sample |
|---|---|---|---|---|---|---|
| **Sample 1** | | | | | | |
| Survey A (47% humanities) | 0 | 35 | 60 | 5 | 0 | Latin American |
| Survey B (80% humanities) | 4 | 21 | 54 | 21 | 0 | American |
| Survey C (63% humanities) | 4 | 30 | 53 | 13 | 0 | Asian |
| **Sample 2** | | | | | | |
| Survey A (47% humanities) | 0 | 21 | 47 | 32 | 5 | American |
| Survey B (80% humanities) | 4 | 18 | 50 | 28 | 0 | Asian |
| Survey C (63% humanities) | 4 | 18 | 39 | 39 | 0 | Latin American |
| **Sample 3** | | | | | | |
| Survey A (47% humanities) | 0 | 5 | 40 | 55 | 0 | Asian |
| Survey B (80% humanities) | 0 | 15 | 78 | 7 | 4 | Latin American |
| Survey C (63% humanities) | 0 | 11 | 43 | 46 | 0 | American |
| **Sample 4** | | | | | | |
| Survey A (47% humanities) | 0 | 30 | 60 | 10 | 0 | Latin American |

| | | | | | |
|---|---|---|---|---|---|
| Survey B (80% humanities) | 4 | 28 | 61 | 7 | 0 | American |
| Survey C (63% humanities) | 4 | 46 | 32 | 18 | 0 | Asian |
| **Sample 5** | | | | | | |
| Survey A (47% humanities) | 5 | 45 | 30 | 20 | 0 | American |
| Survey B (80% humanities) | 8 | 44 | 41 | 7 | 0 | Asian |
| Survey C (63% humanities) | 8 | 31 | 46 | 15 | 4 | Latin American |
| **Sample 6** | | | | | | |
| Survey A (47% humanities) | 15 | 20 | 50 | 15 | 0 | Asian |
| Survey B (80% humanities) | 15 | 33 | 37 | 15 | 0 | Latin American |
| Survey C (63% humanities) | 4 | 33 | 52 | 11 | 0 | American |

A survey designed to more finely distinguish among disciplines with more respondents from across the curriculum may contradict our findings, but based on our limited sample, there does not seem to be a substantial difference in bias between humanities and non-humanities faculty.

## Conclusion

This study was undertaken purely as a preliminary investigation to discover what ways (if any) bias related to nationality continues to be a factor in writing assessment across the curriculum. While our survey pool was larger and represented a broader range of experiences than the study by Rubin and Williams-James (1997) that provided the basis of our survey, 87 is still a small number that serves more to highlight areas for further research than offer hard and fast conclusions. Moreover, the numbers of participants outside humanities was not large enough to offer any convincing patterns of how faculty in specific disciplines responded. It was the intention of the authors to include only descriptive and qualitative data analysis in this paper in order to *begin* to give a general picture of the state of ethnolinguistic bias in WAC writing assessment.

While sweeping generalizations are difficult to make from any study, the findings nevertheless suggest that there is a great deal of variation in scores when students are evaluated analytically (with raters providing separate scores for elements such as clarity and organization) as opposed to

holistically (with raters offering an overall grade). The descriptive data from this project reinforces earlier findings that ethnolinguistic bias exists, but indicates that, at least when faculty grade holistically, the bias is now in favor of rather than against international student writers. While raters seem to have avoided conflating language and intelligence when rating international student writers, U.S. students are perceived as less intelligent when their writing has the same amount of errors as an international writer. The data provided by our small survey pool suggests that these biases occur throughout the academy and are not discipline-based.

These findings, though preliminary, offer substantial material for WAC/WID faculty and administrators to consider when working with students from other countries. The fact that our findings found that the type of assessment (analytic vs. holistic) affected the level of negative bias toward a particular type of student (native vs. non-native) raises the question of whether different assessment tools should be employed for students from other countries. Silva (1993) has argued that NES and NNES writing "are different in numerous and important ways. This difference needs to be acknowledged and addressed by those who deal with NNES writers if these writers are to be treated fairly, taught effectively, and thus, given an equal chance to succeed in their writing-related personal and academic endeavors" (p. 671). While we concur with Silva's findings, our study raises the question of whether the recognition that international student writing is different from writing by American students has not created a climate in which international student writing is automatically seen as more or less proficient simply because it is identified as international. Our study suggests that may be the case since students were ranked differently based on the ethnicity of their names and assigned nationalities. We recommend that in courses where American and international student writers are mixed, as in WID courses, that faculty consider not having students identify themselves by name on their papers. This would help faculty avoid ethnolinguistic bias (and possibly other implicit biases such as gender) by eliminating the factor that created the discrepancy in evaluations in our study: the student's identity. While the differences between NES and NNES writing noted by Silva may in many cases still mark a student's effort as NNES, anonymous submission would make it more likely that such a conclusion is based on the writing itself and not on any preconceived notions the reader may have about the student's ability based on nationality. In many courses, faculty strive to create a relationship with students and discuss or comment on their writing at all stages of the writing process. When pre-writing exercises and multiple drafts are submitted, as in some writing-intensive courses, it may be more difficult for faculty to not recognize an author's identity simply because they remember discussing the topic with a student. Nevertheless, anonymous submissions could be effective in reducing ethnolinguistic bias in many courses where students submit less evidence of their writing process and on essay exams where only a single draft is submitted. Even using anonymous submission on a few assignments could help faculty evaluate whether or not they demonstrate ethnolinguistic bias. If anonymous submissions are not a viable option, faculty might consider varying the types of assessment they use. By using both analytic and holistic assessment in a course, faculty can reduce the likelihood that a student is subjected to a single form of ethnolinguistic bias (harsher analytic evaluations or kinder holistic ones) throughout the term.

For administrators, educating faculty about writing assessment is a vital key in reducing the variation among evaluators; however, research in reducing cultural bias suggests that this alone is not enough. When teachers understand the workings of assessment, the complications of writing in a second language, and the development of assessment skills, they can practice assessment which is more fair, equitable, reliable, and valid. But until faculty also recognize the possibility that their assessment may be affected by implicit biases, even the most tolerant and fair-minded may unconsciously rate international writers higher or lower simply because they are international writers. WAC/WID administrators should provide faculty development opportunities in the assessment of writing, the

creation of rubrics, and criteria generation in addition to training in second language acquisition theory and practice. Many programs already do so. But we suggest that concerted efforts to address the implicit biases at work for and against international writers must also be made. Since earlier studies suggest that diversity workshops do little to actually change practice and attitudes, eliminating ethnolinguistic bias is not something that is likely to be vanquished with a simple workshop. Nevertheless, offering faculty a variety of self-reflective and interactive practices can help teachers respond more fairly to international students (Tyler, Stevens & Uqdah, 2009). An essential first step is for faculty and administrators to examine their own biases, thus workshops involving activities that give faculty the opportunity to evaluate international and American student writing and compare differences in evaluations may be useful. Sharing the findings in this article could help some teachers consider how the type of assessment they use may be influenced by ethnolinguistic bias. We particularly encourage WAC programs to work with ESL specialists on their campuses to do a study like ours and share the results with their faculty of how such bias, if any, occurs on their campus. Creating a space—be it discussion groups, workshops, new faculty training workshops, or other forum—that give faculty the opportunity to be more aware of their own attitudes toward specific cultures without feeling judged for those biases is an important element: "For many, reducing cultural bias in teaching requires teachers to become more aware of themselves as cultural beings" (American, 2003).

Increasing faculty awareness that such biases exist and offering them tools for avoiding such bias is essential. The number of faculty who rarely or never see international students in their classes is certain to decrease with each passing year. Despite dips in international student enrollments following 9/11 and the economic downturn in 2008, overall, the first decade of the 21st century has continued a long-running trend of steadily increasing international enrollments. During the 2009-2010 academic year, international students comprised 3.5 percent of U.S. university enrollments. Nearly 275,000 of those students enrolled as undergraduates ("Open Doors 2010," 2010). Such statistics merely hint at the numbers of students enrolled in U.S. universities who could fall under the heading of international learners. A growing number of NNES students enrolled in U.S. universities are not international students who specifically come to the United States as temporary residents on student visas. An increasing portion of the population identified as having NNES or English as a Second Language (ESL) issues in their writing come from immigrant, refugee, and Generation 1.5 populations, the term used to describe "those who immigrate as young children and have life experiences that span two or more countries, cultures, and languages" (Roberge, 2009, p.4). Harriet Allison notes that "[i]n 1998, more than 90% of the ESL students at the southeastern U.S. community college where [she] taught had completed high school outside the U.S.; within 3 years, the college's ESL enrollment had grown from 16 to 160 students, almost 90% of whom were U.S.-educated English learners" (Allison, 2009, p. 75). Failing to educate faculty to evaluate the NNES writer without bias is not something higher education in the United States can afford to do.

## Appendix A - Demographic Survey

1. Are you a native speaker of American English? (A native speaker is someone who became fluent in the language in childhood. Typically one's native language is the first language learned from birth.)
2. If American English is not your native language, what is your native language?
3. What languages other than American English have you studied?

4. Please list any languages other than American English in which you are fluent. (Fluency here means you are capable of understanding and being understood in the language easily and accurately.)
5. Have you visited or lived in any of the following regions? Choose all that you have visited or lived in.
   o Europe
   o East Asia
   o Southeast Asia
   o The Middle
   o EastAfrica
   o Africa
   o Oceania (Australia/New Zealand/Polynesia/Micronesia)
   o The Caribbean
   o Latin America/Mexico
6. What is your primary teaching area?
   o Fine Arts
   o Humanities
   o Social Sciences
   o Business
   o Science
   o Math
   o Psychology
   o Engineering/Computer Science
   o Nursing
7. Have you had any training in teaching writing across the curriculum such as graduate course work, faculty development workshops, or continuing education credits?
8. Please describe any training you have had in teaching writing across the curriculum such as graduate course work, faculty development workshops, or continuing education credits.
9. Have you had any training in teaching English as a Second Language such as graduate course work, faculty development workshops, or continuing education credits?
10. Please describe any training you have had in teaching English as a Second Language such as graduate course work, faculty development workshops, continuing education credits.
11. Approximately how many non-native speakers of English do you typically teach in a year?
12. Approximately how many years have you been teaching?
13. What types of writing do you require, if any, in your courses?
    o I do not require writing in my courses
    o Short essays (5 pages or fewer)
    o Long essays (more than 5 pages)
    o In-class writing
    o Essay exams
    o Journals, blogs or other informal at-home assignments
    o Lab reports

# Appendix B - Sample Essays

**Essay 1**

Although many adults claim Facebook and text messaging run our teenage lives, I strongly feel that taking these social media outlets away for a week would not be helpful in anyway. Adults are simply stuck in their generation and find it hard to adapt to this technological shift. In contrast to many beliefs, Facebook and text massaging actually helps people become more social and involved. By boycotting these media outlets, talking and hanging out with friends would be a lot more difficult. In regards to college students, phones and computers serve for a connection to old friends and family members. I would strongly advised my peers not to participate in the one-week boycott. The boycott would only distance teens from their peers by making them an outcast unreachable. This is our generation, our means of communication, and if adults don't like it, they don't have to be involved.

**Essay 2**

Dear peers,

Have you heard that the university wants you to stop using social media? It is an outrage that the university will push us to participate in a voluntary social media boycott for an entire week. The failure to check you're email for seven day straight could potentially result in various academic consequences. It will ultimately lead to procrastination and disorganization. Avoiding social media will also put a damper in your relationships. You could potentially miss the biggest social event of the year! Avoiding any type of social contact will also add much stresses to your parents' lives. Consider the medical risks including bloody high pressure. Your organization skills, social life, and your parents' serenity depend on your choice to engage in social media. Make the right choice.

Sincerely,
XXXXX

**Essay 3**

Dear students,

Listen up! We have a problem very serious that seems to be getting worse. I have been hearing more and more frequent that we are losing our ability to interact because of all of technology that has come to rule our lives. This is not true! I have plenty of friends. I know this because Facebook keeps track of them all! To ask us to give up those friends for a week is a pointless exercise. Social media is not just about touching bases with friends. We also use it to find out what's going on socially and in our classes. Some professors even have facebook pages and blogs for class. How would we do the work in those classes. Maybe some people are online too much, but to ask all of us to give up social media is ridiculous. I encourage you to despise the boycott and keep using your technology. Stand united!

**Essay 4**

Dear fellow students,

As we arrive in the age of Facebook, Twitter, and all around social networking, we sometimes need little reminder to get out and actually be social! Just take a step back and look at what all your university has to offer. Recreational sport, clubs, and huge amounts of activities away you. It's time to put down that laptop, put down that iPhone, and get outside with some friends.

For those very reasons I urge you to participate in our technology boycott here at the University. Understandably students, like yourself, may go through technology withdrawl. It's going to be alright.

Who knows, you might even really like what you discover from meeting more people on face to face interactions. Surely this is an experience we can all grow from so don't miss out!

Sincerely,
XXXXX

**Essay 5**

Dear Peers,

I highly encourage all of you to join in on this expirement. Although social medias are a great way to communicate with each other. I believe it would be much better to communicate with each other in person. The time spent in our rooms or looking at our phones takes away from spending time actually interacting with others. I feel that if we take away these distractions, then we will be much more likely to go out and get involved. Think of all the extra free time we will have if we couldn't text, tweet or Facebook. We would have much more time to be with others, and more time to join other activities. It doesn't really matter how many friends you have on Facebook; what matters is how many friends you have in real life. So I encourage you all to stop wasting your time on the internet or on your phones and go out and get involved in your community.

**Essay 6**

Dear peers,

Have you heard? Soon our campus will be taking part in social networking experiment. In this experiment, all involved will cut their ties with their social network and participate in a week long boycott. I urge all you to take part in this. There is a growing trend in our society to rely on facebook, texting and twitters to both communicate and build relationship. This is beginning to take away for the importance of face to face interaction. By taking part in the experimental boycott, we can raise awareness to both ourselves, and our peers the importance of being involved outside of our social networking walls. There is a world to be lived out there. There are people to see, meet, and talk to. All of this can be seen through boycott. So please consider joining our cause.

## Appendix C - Rating Scale

1.   Please rate the quality of [Student name]'s writing sample.

| 2. | Much better than most first year college writers | Above average for a first year college writer | Average for a first year college writer | Below average for a first year college writer |
|---|---|---|---|---|
| Ability to convey overall point clearly | 4 | 3 | 2 | 1 |
| Ability to present ideas in a clearly organized manner | 4 | 3 | 2 | 1 |
| Ability to develop major points with appropriate and convincing evidence | 4 | 3 | 2 | 1 |

| Mastery of conventions of written English (mechanics, grammar, style and usage) | 4 | 3 | 2 | 1 |
| --- | --- | --- | --- | --- |
| Overall quality fo the writing | 4 | 3 | 2 | 1 |

3. The overall grade for the CONTENT of [Student name]'s writing sample should be:
   o A=Superior effort for a first year college writer
   o B=Above average effort for a first year college writer
   o C=Average effort for a first year college writer
   o D=Below average effort for a first year college writer
   o F=No effort demonstrated/inferior effort for a first year college writer

4. The overall grade for the MECHANICS, USAGE, STYLE, and GRAMMAR of [Student name]'s writing sample should be:
   o A=Superior effort for a first year college writer
   o B=Above average effort for a first year college writer
   o C=Average effort for a first year college writer
   o D=Below average effort for a first year college writer
   o F=No effort demonstrated/inferior effort for a first year college writer

5. The overall grade for [Student name]'s writing sample should be:
   o A=Superior effort for a first year college writer
   o B=Above average effort for a first year college writer
   o C=Average effort for a first year college writer
   o D=Below average effort for a first year college writer
   o F=No effort demonstrated/inferior effort for a first year college writer

6. Based on this sample, [Student name] seems to be
   o A very good writer
   o A good writer
   o An okay writer
   o A bad writer

7. Based on this sample, Jesus Rodriquez appears to be
   o Very intelligent
   o Of average intelligence
   o Not very intelligent

8. Do you have any additional comments about this writing sample? If so, please provide them below. Please identify the sample writer (Student name) by name in your response so we can more easily connect your comments to the specific student. Thank you.

# References

Allison, Harriett. (2009). High school academic literacy instruction and the transition to college writing. In Mark. Roberge, Meryl Siegal, & Linda Harklau (Eds.),*Generation 1.5 in college composition: Teaching academic writing to U.S.-educated learners of ESL.* (pp. 75-90). New York, NY: Routledge.

American Psychological Association. (2003). Guidelines on multicultural education, training, research, practice and organizational change for psychologists. *American Psychologist*, *58*(5), 377-401.

Boykin, A.Wade, Tyler, Kenneth M., & Miller, Oronde A. (2005). In search of cultural themes and their expressions in the dynamics of classroom life. *Urban Education*, *40*(5), 521-549.

Boykin, A.Wade, Tyler, Kenneth.M., Watkins-Lewis, Karen M., & Kizzie, Karmen. (2006). Culture in the sanctioned classroom practices of elementary school teachers serving low-income African-American students. *Journal of Education of Students Placed At-Risk*, *11*(2), 161-173.

Brown, James Dean. (1991). Do English and ESL faculties rate writing samples differently? *TESOL Quarterly, 25*(4), 587-603.

Clair, Nancy. (1995). Mainstream classroom teachers and ESL students. *TESOL Quarterly*, *29*(1), 189-196.

Clark, Pat. (2010, Spring). I don't think I'm biased. *Teaching Tolerance*, *37*. Retrieved from http://www.tolerance.org/magazine/number-37-spring-2010

Connor, Ulla. (2002). New directions in contrastive rhetoric. *TESOL Quarterly* 36 (4), 493-510.

Connors, Robert J., & Lunsford, Andrea A. (1988). Frequency of formal errors in current college writing, or Ma and Pa Kettle do research. *College Composition and Communication, 39*(4), 395-409.

Crusan, Deborah. (2001, March). *Conflicting communities: ESL writers and disharmonious definitions of academic discourse.* Paper presented at the 52nd Annual Convention—College Conference on Composition and Communication, Denver CO.

Crusan, Deborah. (2010). *Assessment in the second language writing classroom*. Ann Arbor: University of Michigan Press.

Cumming, Alister, Kantor, Robert, & Powers, Donald E. (2002). Decision making while rating ESL/EFL writing tasks: A descriptive framework. *The Modern Language Journal*, *86*, 67-96.

The economic benefits of international education to the United States for the 2009-2010 academic year: A statistical analysis. (2010). *NAFSA Association of International Educators.* Retrieved from http://www.nafsa.org/_/File/_/eis2010/usa.pdf

Ferris, Dana R., & Hedgcock, John. (2005). *Teaching ESL composition: Purposes, process and practice* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

"General Information." (2008) *Project Implicit.* Retrieved from http://www.projectimplicit.net/generalinfo.php

Huang, Jinyan. (2009). Factors affecting the assessment of ESL students' writing. *International Journal of Applied Educational Studies, 5*(1), 1-17.

"International Student Enrollments in United States Rose in 2009-2010." Nov 15 2010. http://www.state.gov/r/pa/prs/ps/2010/11/150933.htm

Kobayashi, Hiroe, & Rennert, Carol. (1996). Factors affecting composition evaluation in an EFL context: Cultural rhetorical pattern and readers' background, *Language Learning*. *46*(3), 397-437.

Kobayashi, Toshihiko. (1992). Native and nonnative reactions to ESL compositions. *TESOL Quarterly, 26*(1), 81-112.

Lane, Janet, & Lange, Ellen. (1999). *Writing clearly: An editing guide*. 2nd ed. Boston, MA: Heinle & Heinle.

Lee, Hee-Kyung. (2009). Native and nonnative rater behavior in grading Korean students' English essays. *Asia Pacific Education Review, 10*, 398-397.

Leki, Ilona. (2006). Negotiating socioacademic relations: English learners' reception by and reaction to college faculty. *Journal of English for Academic Purposes*, *5*, 136-152.

Ludwig, Jeanette. (1982). Native-speaker judgments of second-language learners' efforts at communication: A review. *Modern Language Journal*, *66*, 274-83.

Mendelsohn, David, & Cumming, Alister. (1987). Professors' ratings of language use and rhetorical organization in ESL compositions. *TESL Canada Journal*, *5*, 9-26.

Milnes, Terry, & Cheng, Liying. (2008). Teachers' assessments of ESL students in mainstream classes: Challenges, strategies, and decision-making. *TESL Canada Journal*, *25*(2), 49-65.

Nairn, Lyndall. (2003, March). Faculty response to grammar errors in the writing of ESL students. *TESOL Newsletter for the ESL in Higher Education E-Section*. Retrieved from http://www.sfu.ca/heis/archive/22-1_nairn.pdf

Ndura, Elavie. (2004). ESL and cultural bias: An analysis of elementary through high school textbooks in the Western United States of America. *Language, Culture & Curriculum*, *17*(2) 143-153.

 "Open Doors 2010 Fast Facts." (2010, November 15). *Institute of International Education*. Retrieved from http://www.iie.org/en/Research-and-Publications/Open-Doors/Data/Fast-Facts

Roberge, Mark. (2009). A teacher's perspective on Generation 1.5. In Mark Roberge, Meryl Siegal, & Linda Harklau (Eds.), *Generation 1.5 in college composition: Teaching academic writing to U.S.-educated learners of ESL* (pp. 3-24). New York: Routledge.

Roberts, Felicia, & Cimasko, Tony. (2008). Evaluating ESL: Making sense of university professors' responses to second language writing. *Journal of Second Language Writing, 17*, 125-143.

Rubin, Donald L. & Williams-James, Melanie. (1997). The impact of writer nationality on mainstream teachers' judgments of composition quality. *Journal of Second Language Writing, 6*(2), 139-153.

Salem, Lori, & Jones, Peter. (2010). Undaunted, self-critical and resentful: Investigating faculty attitudes toward teaching writing in a large university writing-intensive course program. *Journal of the Council of Writing Program Administrators*, *34*(1), 60-83.

Santos, Terry. (1988). Professors' reactions to the writing of nonnative-speaking students. *TESOL Quarterly*, *22*(1), 69-90.

Schreyer Institute for Teaching Excellence. (2007). "The Basics of Rubrics." Retrieved from http://www.schreyerinstitute.psu.edu/pdf/rubricbasics.pdf.

Siegmund, John. (2009, September). Higher education shows a big trade surplus for the United States. *International Trade Update*. Retrieved from http://trade.gov/press/publications/newsletters/ita_0909/higher_0909.asp

Silva, Tony. (1997). On the ethical treatment of ESL writers. *TESOL Quarterly*, *31*(2), 359-363.

Silva, Tony. (1993). Toward an understanding of the distinct nature of L2 writing: The ESL research and its implications. *TESOL Quarterly*, *27*(4), 657-677.

Song, Bailin, & Caruso, Isabella. (1996). Do English and ESL faculty differ in evaluating the essays of native English-speaking and ESL students? *Journal of Second Language Writing*, *5*(2), 163-182.

Tobin, Kenneth, & Gallagher, James. (1987). The role of target students in the science classroom. *Journal of Research in Science Teaching*, *24*(1), 61–75.

Tyler, Kenneth M., Boykin, A. Wade, & Walton, Tia R. (2006). Cultural considerations in teachers' perceptions of student classroom behavior and achievement. *Teaching and Teacher Education, 22*, 998–1005.

Tyler, Kenneth M., Stevens, Ruby, & Uqdah, Aesha L. (2009). *Cultural bias in teaching*. In Eric Anderman & Lynley Anderman (Eds.), *Psychology of Classroom Learning: An Encyclopedia*. (pp. 292-296). Farmington Hills, MI: Thomson Gale Publishing.

Valdez-Pierce, Lorraine. (2000, October 27). *Using assessment to inform instruction: Learning by doing*. Keynote presented at the Ohio TESOL Conference, Columbus, OH.

Vann, Roberta J., Meyer, Daisy E., & Lorenz, Frederick O. (1984). Error gravity: A study of faculty opinion of ESL errors. *TESOL Quarterly*, *18*(3), 427-440.

Weigle, Sara C., Boldt, Heather, & Valsecchi, Maria Ines. (2003). Effects of task and rater background on the evaluation of ESL writing: A pilot study. *TESOL Quarterly*, *37*(2), 345-354.

Zamel, Vivian. (2004). Strangers in academia: The experiences of faculty and ESOL students across the curriculum. In Vivian Zamel & Ruth Spack (Eds.), *Crossing the Curriculum: Multilingual Learners in College Classrooms* (pp. 3-17)*.* Mahwah, NJ: Lawrence Erlbaum.

## Notes

[1] The Schreyer Institute for Teaching Excellence at Penn State offers useful definitions for understanding holistic vs. analytic scoring: "Holistic rubrics provide a single score based on an overall impression of a student's performance on a task…. Analytic rubrics provide specific feedback along several dimensions."

## Contact Information

Peggy Lindsey
Department of English
300 College Park
University of Dayton
Dayton OH 45469
Email: plindsey1@dayton.edu

Deborah Crusan
Department of English Language and Literatures
470 Millett Hall
Wright State University
Dayton, OH  45435
Email: deborah.crusan@wright.edu
Phone: 937-775-2846

## Complete APA Citation

Lindsey, Peggy, & Crusan, Deborah. (2011, December 21). How faculty attitudes and expectations toward student nationality affect writing assessment. *Across the Disciplines, 8*(4). Retrieved from https://wac.colostate.edu/docs/atd/ell/lindsey-crusan.pdf