Journal of the Scholarship of Teaching and Learning, Vol. 25, No. 3, September 2025, Article 1
doi: 10.14434/josotl.v25i3.35964

Undergraduate Students' Perceptions on Student Evaluation of Teaching (SET): Implications for Meaning Negotiation as Critical Pedagogy

Soha Youssef
Thomas Jefferson University
soha.youssef@jefferson.edu

Abstract: Scholarship indicates the existence of implicit biases against nonnative English-speaking teachers (NNESTs). Those biases are revealed in terms of their ethnicity impacting their perceived comprehension by US native English-speaking students (NESSs) and how students' prejudices against perceived "foreign" accents undermine NNESTs' performance on Student Evaluation of Teaching (SETs). However, little research examines students' abilities to detect implicit biases. To fill that gap, a qualitative study where 47 undergraduate NESSs from Midwestern University (pseudonym) were surveyed. On the survey, they were asked whether they have had experiences with NNESTs. Based on their answers, students were divided into 1. an Experience group who had experiences with NNESTs and 2. a No Experience group who had not had experiences with NNESTs. Both student groups were asked to interpret an existing SET question related to the English-speaking abilities of the instructor. The original hypothesis was that the Experience group would be more capable of detecting the question's implicit bias and less inclined to reveal biases against NNESTs than the No Experience group. Feminist methodologies informed the research practices, and a general inductive coding and critical pedagogy were operationalized for analysis. The findings indicated that the Experience Group were more capable of detecting implicit biases, expressing some willingness to exert what Lippi-Green calls NESSs-NNESTs communicative labor, whereas the No Experience group did not. The implications present critical pedagogy as a potential intervention in the use of SETs where students are made aware of their implicit biases prior to evaluation.

Keywords: nonnative English-speaking teachers, student evaluation of teaching (SET), meaning negotiation, implicit biases, communicative labor, critical pedagogy.

Scholarship shows that Student Evaluation of Teaching (SETs) often reveal students' implicit biases against racially minoritized instructors (Huston, 2006; Reid, 2010). Another area of research on SETs highlights how the form reveals implicit biases against nonnative English-speaking teachers (NNESTs). Those biases are often revealed in terms of how NNESTs' ethnicity impacts their perceived comprehension by native English-speaking students (NESSs) (Rubin, 1992; Liu, 2005; Mutua, 2014), and how students' prejudices against foreign accents undermine NNESTs' performance on Student Evaluation of Teaching (SETs) (Rubin, 1992; Plakans, 1997; Liu, 2005; Mutua, 2014; Roessel et al., 2018; Baratta, 2018). Thus, it is safe to infer from all this research that SETs often reveal biases that are directed against various aspects of instructors' identities.

Indeed, one's race, ethnicity, nationality, linguistic competency and/or accent are aspects of identity that are inseparable in terms of shaping one's experiences. In fact, Shuck (2006) argued that NNESTs tend to be racialized in a way where someone's comment on our accent should not be merely perceived as a comment on our language competency, but rather a way for the speaker to mark us as a racialized other. In other words, one's race and accent are closely connected. Lippi-Green, for example, argued that "when people reject an accent, they also reject the identity of the person speaking: his or her race, ethnic heritage, national origin, regional affiliation, or economic

class" (1994, p. 165). Thus, on SETs, when a student comments on a NNEST's accent or speech mannerisms (or any other instructor's speech mannerisms for that matter), that student is, indeed, racially discriminating against the instructor, since race and speech mannerisms are closely associated. By extension, SET questions that guide students to comment on their instructors' speech mannerisms are grounded in and underscore the systemic racism of the corresponding educational institution or, more specifically, the personnel who are responsible for that rhetoric on SETs. SETs, then, become a space where students are encouraged to exercise their privilege, power dynamics, and implicit biases.

To better understand the complexity of language and how it informs and is informed by privilege and power dynamics, a definition of language ideology (also see Woolard & Schieffelin, 1994) is necessary to ground this discussion. Subtirelu (2014) defined language ideology as:

> The networks of beliefs that language users hold, either tacitly or overtly, about language and its assumed relation to other aspects in their environments, especially other individuals and social groups, which stem either from explicit teaching or implicit socialization. Owing to their connections between language and language users or social groups, language ideologies are intimately connected to macro-structures of power and privilege. In particular, language ideologies are often used as an interpretive lens to explain the ways in which language varies across social groups in a society and to draw connections between specific aspects of language and the identities, moral goodness, or worth of those using particular features of language (p. 121).

That connection between "language ideology" and "macro-structures of power and privilege" (i.e., language inheriting its privilege from that of its users) presents itself in the scholarship that examines students' perceptions of their NNESTs in the context of SETs (the focus of this paper). Such scholarship reveals students' linguistic implicit biases against NNESTs (Berdie et al., 1976; Rubin, 1992; Fox & Gay, 1994; Halleck & Moder, 1995; Plakans, 1997; Fitch & Morgan, 2003; Liu, 2005; Villarreal, 2013; Mutua, 2014; Roessel et al., 2018). Little research, however, examines students' abilities to detect implicit biases in the language of the SET form itself and to what degree can their contact with NNESTs prepare them to identify implicit biases when they encounter said rhetoric on SETs. This study aims to fill that gap, offering an intervention as to how SETs are or should be used in US educational institutions. My study is, therefore, driven by the following research questions:

> RQ1. Are undergraduate NESSs able to detect implicit biases in the language of SET forms?
> RQ2. To what degree can NESSs' contact with NNESTs prepare them to identify biases on SET forms?

The purpose of this paper is to present the results of a qualitative study where undergraduate students who consider themselves native speakers of English were asked, first, to indicate whether they have had experiences with NNESTs and, second, to provide their interpretation of an existing SET question[1] that asks whether the instructor speaks the English language clearly. The choice of undergraduate native English-speaking students (NESSs) as the research participant population was informed by scholarship that shows that NESSs tend to discriminate against NNESTs more than

[1] This yes/no SET question appeared on a SET form at an academic institution where the primary investigator of this study used to assume the position of a Teaching Assistant.

their nonnative English-speaking counterparts do (Borjas, 2000; Liu, 2005; Mutua, 2014; Plakans, 1997; Rubin, 1992). It was hypothesized that students who have had experiences with NNESTs would be more capable of detecting the question's implicit bias and would be less inclined to reveal implicit biases against NNESTs than their counterparts who have not had contact with NNESTs. The rationale for this hypothesis was based on scholarship that indicated the way contact exercises between NNESTs and NESSs raised both groups' awareness of their differences in nonverbal communication and cross-cultural misunderstandings (Kang & Rubin, 2012) and fostered cross-cultural competencies (Halleck, 2008). Thus, by extension, contact time between NNESTs and NESSs might also mitigate the latter's implicit biases toward the former, making them more able to detect implicit bias on SETs.

Literature Review

Though SETs are ubiquitously used in the academy, the validity of those forms has been contested, with scholars taking strong stances on the issue. On one hand, a significant amount of research suggests that course evaluations lack validity in terms of measuring teaching quality, particularly because the forms are biased by factors such as grading leniency (Blunt, 1991; Greenwald & Gilmore, 1997a, b; Griffin, 2004; Olivares, 2001), course workload (Greenwald & Gilmore, 1997a, b; Griffin, 2004), and pre-course student expectations (Griffin, 2004). On the other hand, little research argues that course evaluations offer a valuable measure of teaching quality, with biased factors having little or no effect (Marsh & Roche, 2000; Remedios & Lieberman, 2008). Thus, scholarship is inconsistent with whether course evaluations carry validity.

Similarly, there are inconsistencies in how the forms reveal gender biases against instructors. Despite the abundance of scholarship that reflects students' gender bias against female instructors with females being evaluated less favorably than their male counterparts (Miller & Chamberlin, 2000; MacNell et al., 2015; Mitchell & Martin, 2018), some studies discovered that instructors' gender carries no effect on how they are rated (Blackhart et al., 2006; Reid, 2010).

Though such gender biases problematize institutional continued use and reliance on SETs for the evaluation of instructors, racial biases constitute a more relevant connection to the study at hand, given the rhetoric of the question under investigation. More specifically, asking students to answer a yes/no question about whether "the instructor speaks the English language clearly" might guide students to respond to the question in terms of the instructor's race, since it is possible that speaking the English language clearly does not mean the same thing to all students, given the diversity of dialectical differences in the US context. Thus, it is necessary to engage scholarship that examines how students' racial biases present themselves in SETs.

Research that compared course ratings of faculty who teach sections of the same course in the same departments discovered that faculty of color receive "significantly lower ratings on average" than their white counterparts (Huston, 2006, p. 599). One study examined the ways instructors' perceived race and gender inform how they are rated on RateMyProfessors.com across the US top 25 liberal arts colleges. Overall, results indicated that instructors of color were rated significantly less favorably than their white counterparts when it came to Overall Quality, Helpfulness, and Clarity, and more favorably when it came to Easiness (Reid, 2010). One can infer from these results that the credentials and capabilities of faculty of color are constantly questioned and judged by their students, which, in turn, as Reid (2010) articulated, could result in faculty creating courses that are less challenging in an attempt to counter the devalued status they often experience in the classroom.

Aspects of one's identity, such as ethnicity, citizenship, and nativeness are difficult to separate. One study examined the well-known public evaluative technology, RateMyProfessors.com, with the goal of understanding whether instructors with last names that are indexical of an Asian ethnicity, specifically Chinese and South Korean, are evaluated differently from their counterparts whose last names imply US citizenship and, by extension, nativeness. Additionally, that study sought to explore the extent of those differences, if any, as well as the ways "the ideology of nativeness" might have informed such differences (Subtirelu, 2015, p. 41). An analysis of the numeric ratings suggests that instructors with US last names have a significant advantage over their Asian counterparts, in terms of how students rated their clarity component in particular. Results of students' evaluative discourse, on the other hand, present evidence of patterns where the NNESTs' English competence was erased, or their racial otherness was highlighted. More specifically, students referred to the ideology of nativeness while either absolving themselves of the responsibility to exert communicative labor or circumventing the dominant language ideology by, first, acknowledging its presence and, then, identifying their communicative responsibility in order for meaning negotiation to occur. In their evaluation of NNESTs, students associated accentedness with intelligibility, identifying the former as an "ideological presupposition" and the latter as a "mitigating factor" that undermines such accentedness. For those students, intelligibility carries prominence, whereas accentedness is merely peripheral (p. 54). This study underscores implicit biases against NNESTs and complicates the way students perceive their NNESTs particularly in terms of notions such as accentedness, intelligibility, and willingness to exert communicative labor. However, the study oversimplifies what exactly constitutes nativeness and how complex racial and linguistic identity is in the US context. Such oversimplification points to a need for additional exploration of how students' implicit biases against NNESTs present themselves in SETs.

One study fills that gap by exploring students' implicit biases as they pertain to instructors' accents. Despite being a native speaker of English, Consolata N. Mutua (2014) had negative experiences as a Kenyan female instructor in the US because of her accent. She contended that undergraduates' perceptions are often negative "with students reporting that their foreign instructors' accents are hard to understand and subsequently evaluating them as not as knowledgeable or skilled as US born TAs" (Mutua, 2014, p. 51). Growing up in Kenya, Mutua does not identify as a NNEST. But her native speaker status does not protect her from students' criticisms of her accent. Mutua usually receives one or two student evaluations that say she was "hard to understand." Because the majority of her students do not usually comment on her accent, she deduced that it had to do with her students' ethnocentrism and attitudes, regarding them to be "either adjusting their ears and/or their prejudices toward foreigners they automatically perceive as non-English speaking" (p. 55). Some of her students' feedback on SETs, however, was blatantly xenophobic, such as this all-caps comment: "THERE SHOULD BE NO FOREIGNERS TEACHING AT (BLANK) UNIVERSITY!" Mutua elaborated saying she receives many condescending comments on SETs "implying that [she] would turn out to be a better teacher and be more understandable with a dash of "Americanism"" (p. 56). That was when Mutua realized that her English competency was negated by her foreignness. This student commentary above might be an exceptional example of student comments that should not be considered representative of all student comments on SET forms. The comment, however, reflects xenophobic student tendencies that should not be overlooked.

Most recently, Wang and Gonzalez (2020) conducted a study to test the presence of racial/ethnic and national origin bias in SETs as it pertains to professors from racially minoritized groups and to those from foreign origins. The authors collected Spring and Fall 2016 SETs of 264 business professors across three universities in Texas. The total number of SETs collected was 915 evaluations. All professors included had doctorate degrees and worked in a full-time capacity, 45.5%

of whom were white and 48.5% of foreign origin. Regression analysis (with gender, university, and course difficulty being controlled for) was used to test the role of racial/ethnic and foreign origin backgrounds on the ratings of faculty's SETs. The findings indicated the presence of bias in SETs where white American professors are rated more favorably than their non-white American or foreign counterparts. More specifically, white American professors were rated more favorably than their counterparts of African American origin and of foreign origin.

Despite the abundance of scholarship that indicates students' implicit biases against NNESTs on SETs, it is undeniable that Native English-speaking students (NESSs)' perceptions of their instructors' teaching hold validity and, thus, warrant our attention, particularly when those perceptions impact our SET results. Because those evaluation forms are essential for NNESTs' professional development and promotion, scholars examined ways NNESTs could utilize their undergraduate students' comments to improve their pedagogy and, hence, their performance on SETs. For example, Chia-Fang Hsu (2014) argued that in order for NNESTs to receive better teaching evaluations and for their students to perceive them positively, NNESTs need to keep an open and positive attitude toward students and their criticisms. Like many NNESTs in the US, Hsu experienced disrespect from students, receiving negative student evaluations that reflect both language-related and classroom management problems. Upon closer examination, Hsu noticed that students' negative attitudes towards her accent and teaching style tend to occur more frequently in lecture-oriented courses or teacher-centered classrooms rather than in courses in which she promoted a student-centered environment through group activities and student presentations. In application of that observation, Hsu started incorporating more student-centered activities in her lecture-oriented courses. Another strategy that improves student learning quality and, in turn, course evaluations is adopting confirmation behaviors. Those confirmation behavior practices, such as "responding to students' questions and comments, demonstrating interest in students, and using interactive teaching styles," result in a decrease in student apprehension and an increase in student cognition (Hsu, 2014, p. 43), hence, a better learning experience that might be consequently reflected in students' comments on SETs. That use of confirmation behaviors, however, only has a direct effect on student learning when the instructor is a NNEST (Hsu, 2012). That direct effect of confirmation behaviors on student learning was not evident in the context of the native English-speaking counterparts (Ellis, 2004). This contrast underscores the additional communicative labor NNESTs have to exert in order to positively inform student perceptions and for student learning to occur in the US context.

Such imbalance of communicative labor exerted by NNESTs is evident when compared to that exerted not only by their native English-speaking counterparts, but also by their undergraduate NESSs. The latter is challenged in scholarship that calls for NNESTs-NESSs' joint preparation programs. Okim Kang and Donald L. Rubin (2012), for instance, proposed a structured contact approach that prompts interaction between NNESTs and NESSs. The intergroup contact exercise was "a mystery puzzle-solving activity," in which each member is given 8 clues and is supposed to use only verbal communications to communicate with the rest of the members in order to solve the given mystery (p. 160). In order for the exercise to reduce undergraduates' prejudices, specific conditions need to be fulfilled: casual settings, where participants maintain equal status, have no prior interactions, and are "interdependent in ways that require or encourage frequent communication across cultural boundaries" (p. 159). After completing two exercises and rotating among groups, participants debriefed, articulating their differences in nonverbal communication and common cross-cultural misunderstandings. Another study that suggested a shared responsibility between NNESTs and NESSs is Gene Halleck's (2008) constructed simulation. Those constructed simulations do not orient only new NNESTs, but also undergraduate NESSs. The objective of the simulation was for both groups to "become familiar with the issues" that might arise during

NNESTs-NESSs interactions and to "probe[ ] the cross-cultural competence of all the stakeholders involved" (p. 137). Not only do these intergroup contact exercises and constructed simulations suggest shared labor between NNESTs and NESSs, but those activities might actually inform undergraduates' perceptions of and attitudes toward NNESTs.

Native English-speaking listeners' attitudes toward nonnative English speakers inform their perception of what they actually comprehend (Lippi-Green, 2012). The notion "Perceived comprehensibility" refers to "listeners' perception of a speaker's comprehensibility;" which is distinct from "actual intelligibility" which has to do with "how well listeners actually understand the stimulus." Accentedness, on the other hand, refers to one's "degree of foreign accent" (Munro & Derwing, 1995a). Though accentedness might not necessarily inform intelligibility, it might increase native speakers' processing time, which, by extension, might inform native speakers' perceived comprehensibility of nonnative speech (Munro & Derwing, 1995b). On a sociolinguistic level, there is "a set of filters through which we hear the people we talk to," and those filters inform our decisions of whether to accept or reject the communicative burden (Lippi-Green, p. 73). Typically, participants in a conversation carry mutual responsibility to create new information. That mutual responsibility takes the shape of complex iterations of "repair, expansion, and replacement" until participants are satisfied (Lippi-Green, 2012, p.72). However, when a native English speaker communicates with someone whom they perceive to have an accent, they "feel perfectly empowered to reject their responsibility, and to demand that a person with an accent carry the majority of the burden in the communicative act" (p. 72). Contrastingly, when two native English speakers come in contact, with one of them being incoherent or unclear, the instant response is not to reject the communicative burden, but rather place additional effort until they reach an understanding. In other words, native English speakers' willingness to exert communicative labor is based on their perception of the speaker and of the speaker's accent. In the classroom context, learners' attitudes shape how they comprehend—or think they comprehend—their NNESTs. Those attitudes, by extension, reflect on not only the communications that take place between learners and NNESTs, but also how the former evaluate the efficacy of the latter on SETs and whether (and in what ways) students' implicit biases shape those evaluations.

The extensive literature that reveals students' implicit biases against NNESTs on SETs begs us for a solution—a Freirean solution that might address not only the problematics of the form, but also the systemic institutional issues from which those problematics stem. In his *Pedagogy of the Oppressed*, Paulo Freire theorized critical pedagogy by way of promoting libertarian education, distinguishing it from the banking model by explaining that libertarian education "lies in its drive towards reconciliation. Education must begin with the solution of the teacher-student contradiction, by reconciling the poles of the contradiction so that both are simultaneously teachers *and* students" (2000, p. 72). With that in mind and given that critical pedagogy "is preoccupied with social injustice and examines and promotes practices that have the potential to transform oppressive institutions or social relations, largely through educational practices" (Keesing-Styles, 2003), and that it is "a radical approach to education that seeks to transform oppressive structures in society using democratic and activist approaches to teaching and learning" (Braa & Callero, 2006, p. 357), critical pedagogy was used as an interpretive lens through which data was made sense of. Based on my knowledge, no studies have examined the problematics of asking inherently biased questions on SETs. Thus, the data was collected with the goal of unpacking the problematics of asking inherently biased questions on SETs, such as the yes/no question "The instructor speaks the English language clearly." More specifically, the goal was to discover whether those implicit biases would be detected by NESSs and whether their contact with NNESTs could enhance those capabilities to detect biases.

Materials and Methods

The Research Site

A qualitative study was conducted in which undergraduate NESSs enrolled at Midwestern University (pseudonym) during Fall 2016 and Spring 2017 semesters were surveyed. For the sake of context, MU was established in 1910 as a teacher-training school or a normal school, and it is currently ranked as a tier-one institution. Data for this research was collected following an institutional review board (IRB) approval (number 922705-3).

Participants

Participants consisted of undergraduate NESSs who were enrolled at MU during the Fall 2016 and Spring 2017 semesters, regardless of level. They were recruited through an institutional email listserv that was made available by the office of Institutional Research Data through which 1000 randomly chosen students were sent a link to the Qualtrics survey. However, due to the limited responses, a different recruitment strategy was employed: The primary investigator of this study visited several sections of introductory writing courses during Fall 2016 and Spring 2017 semesters in order to gauge students' interest in taking the survey. Students were requested to provide their email contacts to demonstrate such interest. Those email contacts were subsequently used to contact students with a link to the Qualtrics survey. The recruitment email message specifically asked for students who spoke English as a first language. Consent forms were embedded in the survey. NESSs were directed to signal consent by ticking the "next" button at the bottom of the first survey page.

Data Collection and Analysis

To distinguish between NESSs who have had contact with NNESTs and those who have not had any contact with NNESTs, participants were, first, asked if they have had experiences with NNESTs. The first survey question was as follows:

> Are you/Have you ever been taught by a nonnative English-speaking instructor (i.e., an instructor or professor who does not speak English as a first language)? If yes, how many instructors?
> Yes _____
> No

Following that question, NESSs were asked another question where they were provided space to offer their interpretation of a hypothetical yes/no question on a SET form. The survey question was as follows:

> If you receive the following yes/no question on a student evaluation of teaching form, how would you interpret it? "The instructor speaks the English language clearly"

The responses to this question together with those of the previous question were meant to identify ways students' contact with NNESTs might have, first, shaped their interpretations of what it means for NNESTs to speak the English language clearly and, second, fostered their ability to detect implicit biases in the rhetoric of SETs.

Theoretical Framework

The subject positionality of the primary investigator of this study informed her practices and how she made sense of the data. Given that the primary investigator identifies as a NNEST and that she has her own share of experiences with SETs being biased against women of color and those who are linguistically minoritized[2], feminist methodologies were found particularly helpful in pushing back against the researcher's potential prejudices, subject positionalities, and personal biases. More specifically, feminist research methodologies remind scholars that "data are always shaped, to a large extent, by researchers' values, theoretical perspectives, and personal histories" (Kirsch, 1999, p. 195) and that to enact our responsibilities as researchers, we need to maintain "[a]n (ongoing, continual) examination of our assumptions as researchers" by way of comprehending the potentials and limitations of our studies (Takayoshi et al., 2012, p. 99). In this study, feminist methodologies presented themselves in the form of a reflexive approach to data analysis. To do so, reflexive practices were adopted where the researcher not only acknowledged but continuously questioned her preexisting assumptions, subject positionalities, and biases as a NNEST. Iterative readings of the data enabled her to see how her own positionality as a NNEST could have been invoked or used to justify preexisting biases. A reflexive approach to data analysis helped her keep those biases in check as she iteratively made sense of the data. Feminist methodologies were also implemented during the data coding process which was executed with an open mind: instead of the researcher imposing her assumptions on the codes that were assigned to the data, the researcher's perceptions were viewed as merely one way of seeing the data (Charmaz, 2014, pp. 132-3). And that is where a reflexive approach to data analysis came in handy as it allowed the researcher to resist overstretched claims that the knowledge that she is making through qualitative research is meant to be seen as "truth." A reflexive approach also allowed her to consider what actions might result from such knowledge-making as well as what actions she should take to create that knowledge (May & Perry, 2014).

Additionally, a general inductive coding approach was used to analyze the data presented in this study by way of "summariz[ing] the raw data and convey[ing] key themes and processes" with the goal of identifying the core meanings that the text being analyzed carries and, hence, describing emerging categories and developing those categories into an open network (i.e. nonhierarchical) model or framework (Creswell, 1998; Thomas, 2006, pp. 240-241). A general inductive coding approach was operationalized and manually implemented by, first, preparing the raw data into a word document and, then, closely and repetitively reading the data until upper-level and lower-level categories emerged from the reading.

Upper-level categories were the ones driven by the distinction that Munro and Derwing (1995a) drew between the three terms: **perceived comprehensibility, actual intelligibility, and accentedness**, by way of understanding what "speaking the English language clearly" means to undergraduates. In other words, the three terms were used to guide, but not restrict, the way students' interpretations of the hypothetical question were made sense of. This fluidity of categorization resulted in lower-level categories emerging from repetitively reading the data. Categories were, then, consolidated and thematically linked, with the goal of understanding ways various categories interacted with one another to form a nonhierarchical framework. Table 1 represents the themes that emerged from data analysis.

[2] The yes/no SET question that is under investigation in this study appeared on a SET form at an academic institution where the primary investigator of this study used to assume the position of a Teaching Assistant.

Table 1. All Themes that Emerged from Data Analysis.

Themes
Perceived comprehensibility
Little or no willingness of NESSs to exert effort or labor for effective communication with NNESTs
Accentedness
NESSs' detection of ambiguity or problematics of the question
NNESTs' English language proficiency and fluency

As mentioned earlier, the themes that emerged from data analysis were, then, interpreted through a critical pedagogy lens, offering a valuable approach to address the shortcomings in SETs and some of the challenges that NNESTs face in the classroom. In other words, critical pedagogy offers an intervention in the use of SETs where students are made aware of their implicit biases prior to them evaluating NNESTs' teaching.

In the following sections, I will explore how those themes emerged from the analysis of the survey results of student participants. In order to understand how often those categories emerged in student responses, measures of frequency were manually counted, and the percentage of each categorical frequency was calculated. The adoption of such descriptive statistics was appropriate, given the small size of the dataset and the goal of the study.

Results

Of the 47 students who returned usable surveys, 24 were categorized as Experience group because they had experiences with NNESTs, and 23 were categorized as No Experience group since they did not have experiences with NNESTs.

Experience Group

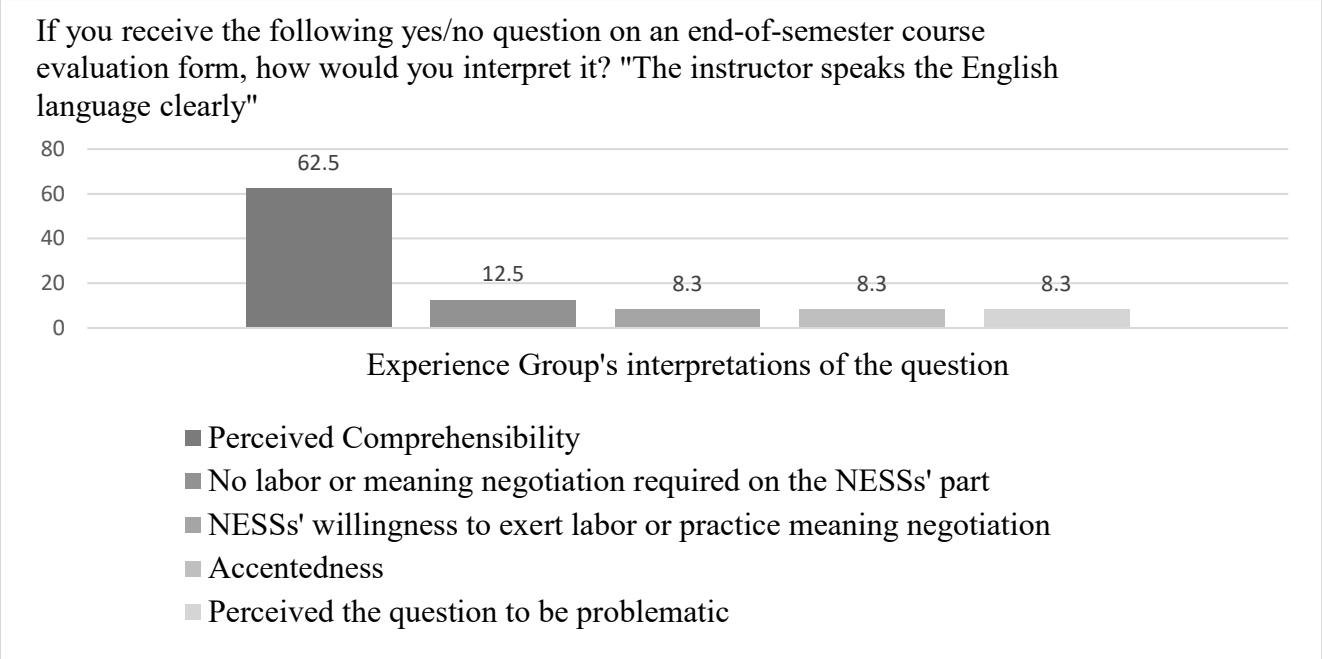

Figure 1. Code results of the Experience Group of NESSs (Figure is property of the author).

Fifteen respondents in the Experience Group (62.5%), which is the majority of the group, interpreted the yes/no question "The instructor speaks the English language clearly" in terms of their own comprehension level of NNESTs' speech or *perceived comprehensibility*. For instance, sample responses to the question were "[NNESTs] have the ability to clearly articulate what they are thinking in a manner that the majority of the class will understand," "The instructor is able to speak English and be understood clearly," and "the instructor can speak English fluently enough that all students can understand." Those statements make connections between the instructor's linguistic ability, on one hand, and their students' comprehension, on the other.

A slightly less common theme that emerged from data analysis of the survey responses of the Experience Group (i.e., five NESSs in the Experience Group or 20.8%) has to do with the amount of effort or labor required for communication. This interpretation of the question is in relation to the *little or no willingness of students to exert effort or labor for effective communication* with NNESTs to take place. For example, responses mentioned that NNESTs speaking the English language clearly meant that they, undergraduate NESSs, do not have to "think hard about what [NNESTs] might have said," or "find [themselves] guessing what [NNESTs] have said or asking others for clarification." Moreover, NNESTs speaking the English language clearly meant that NESSs are "able to make out the words as [NNESTs] talk," "without questioning what was said." For some NESSs, clarity of speech meant whether they can "easily communicate in English with the instructor." Alternatively, they would ask for clarification from their peer undergraduate students. In any case, the theme refers to undergraduate NESSs' exerted effort for communication to occur—some were more willing than others to exert effort or practice meaning negotiation. To be specific, 12.5% of the 20.8% showed no interest in exerting labor for meaning negotiation to take place, whereas only 8.3% of the 20.8% expressed some willingness to exert meaning-communication labor.

While *perceived comprehensibility* was the most common theme to emerge, *accentedness* was on the other end of the spectrum; only two NESSs in the Experience Group (8.3%) interpreted the question to be about NNESTs' accents, with interpretations of the question that presented themselves as a one-word response: accent, without any elaboration. Another theme that was equally uncommon among the Experience Group (i.e., only two NESSs in the Experience Group or 8.3%) had to do with *NESSs' detection of ambiguity or problematics of the question* "The instructor speaks the English language clearly." Participants even proposed revisions. One respondent remarked that the word "'clearly' "needs to be defined." Another was uncomfortable with the question's implied binary and, accordingly, proposed "a maybe column." This theme, though uncommon among the Experience Group, has implications on the rhetorical awareness of undergraduate NESSs who have had contact with NNESTs as well as their ability to articulate the problematics of binaries, implicit biases, and xenophobic rhetoric on SETs—forms of whom undergraduate students are one of the main audiences.

No Experience Group

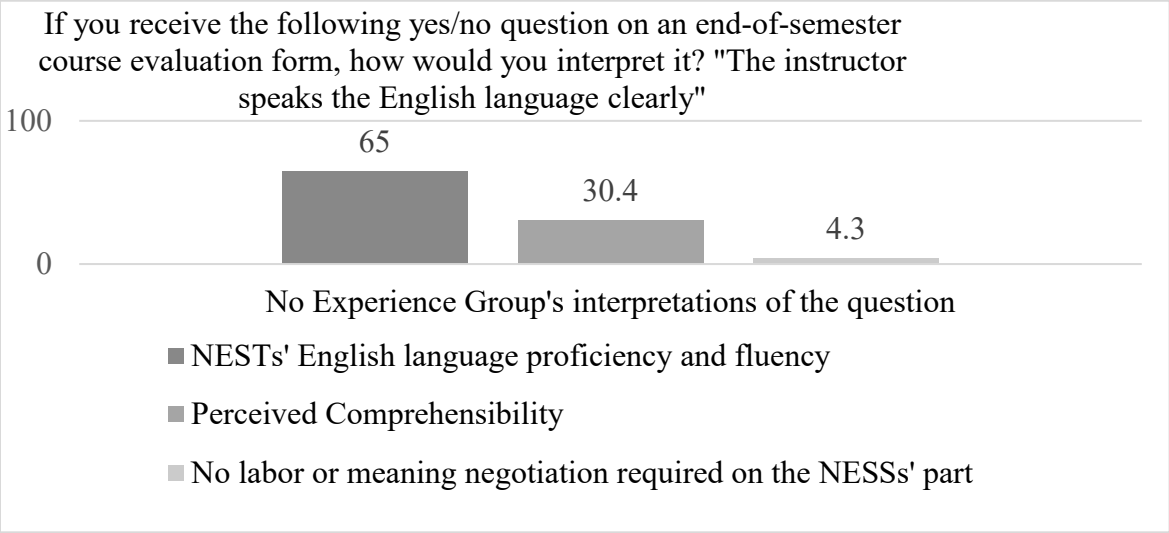

Figure 2. Code results of No Experience Group of NESSs (Figure is property of the author).

The theme *NNESTs' English language proficiency and fluency* was the most common theme (65.0% of the No Experience Group) to emerge from the analysis of the No Experience Group responses to the survey. Sample responses were "[NNESTs] have a main understanding of the language" and "Someone who is fluent in the English language and can pronounce everything clearly."

Another theme that emerged from the analysis of the survey results of the No Experience Group was in terms of their own comprehension level of NNESTs' speech (i.e., *perceived comprehensibility*). Seven of the No Experience Group (30.4%) interpreted the question in terms of perceived comprehensibility. Sample responses were "If the points that are trying (sic) to be made are coming across, then accents are easier to pick up on," "Being able to understand [NNESTs] enough to interpret their words," "That [NNESTs] know how to communicate effectively with little or no difficulty," and "If the student could understand the teacher."

For the sake of transparency, it is worth noting that only one respondent of the No Experience Group (4.3%) interpreted the question in relation to the amount of effort or labor required of them for effective communication with NNESTs to take place, stating "I would have no problem communicating with that instructor." Though this comment implies willingness to communicate with the instructor, it reflects no tolerance for communication breakdowns or willingness for meaning negotiation efforts on the part of the student, meaning that the onus is placed entirely on NNESTs for communication to take place. Now, this response, though important to consider and contemplate as it represents the interpretation of one of the student participants, it cannot be deemed a theme.

Experience Group vs No Experience Group

Table 2. Thematic Results of the Experience Group's vs the No Experience Group's Responses.

Themes that emerged from analyzing the responses of both Groups	
Experience Group	No Experience Group
Perceived comprehensibility	NNESTs' English language proficiency and fluency
Little or no willingness of NESSs to exert effort or labor for effective communication with NNESTs	Perceived comprehensibility
Accentedness	
NESSs' detection of ambiguity or problematics of the question	

Note. These are the thematic results of the Experience Group versus the No Experience Group of NESSs (Table is property of the author)

While the themes that emerged from the qualitative analysis of such a small dataset were distinct across the two groups of participants (see Table 2), one theme appeared in the responses of both groups as they contemplated what NNESTs' clarity of speech might mean: *perceived comprehensibility*. The rest of the themes that emerged from data analysis distinguished the two groups from each other. On one hand, the Experience Group perceived NNESTs' clarity of speech as a reference to *accentedness*, or *little or no willingness of NESSs to exert effort for effective communication with NESTs* (i.e., meaning-negotiation strategies as they pertain to instructor-student division of communicative labor), or *NESSs' detection of ambiguity or problematics of the question.* While the Experience Group were divided in terms of the degree of their willingness to exert some labor or to practice meaning negotiation with NNESTs for successful communication to take place, the No Experience Group did not bring this idea up in a way that deems it a theme. Instead, the No Experience Group interpreted NNESTs' clarity of speech in terms of *NNESTs' English language proficiency and fluency*.

Discussion and Implications

The goal of this study was to answer the following research questions:

> RQ1. Are undergraduate NESSs able to detect implicit biases in the language of SET forms?
> RQ2. To what degree can NESSs' contact with NNESTs prepare them to identify biases on SET forms?

Based on the themes that emerged from data analysis (See Table 2), the answer to RQ1 is in the affirmative. One can see evidence of undergraduate NESSs' ability to detect implicit biases in the language of SET forms; however, one group of participants was more able to do so: The Experience Group. The Experience Group were more able to detect implicit biases in the language of SET forms than their No Experience Group counterparts. More specifically, the theme *NESSs' detection of ambiguity or problematics of the question* emerged from the analysis of the Experience Group's responses—not from the responses of the No Experience Group counterparts. One potential implication here is that NESSs' contact with NNESTs could have prepared them—to some

degree—to identify biases on SET forms, which could provide an answer to RQ2. In using critical pedagogy as an analytical lens, one could leverage students' capabilities to identify biases on SET forms in ways that could expand even beyond the rhetoric of the SET form. And there is one theme that emerged from analysis and that could direct us to an area where students' capabilities to detect biases could be utilized: *accentedness*.

While the Experience Group perceived NNESTs' clarity of speech to be related to accentedness, none of the No Experience Group brought up the notion altogether. Instead, the No Experience Group saw speech clarity to have to do with *NNESTs' English language proficiency and fluency*. If the distinction between the Experience and the No Experience Groups lies in students having had the opportunity to interact with NNESTs, one might infer that having interactions with NNESTs effects a change in undergraduate NESSs' perceptions on NNESTs' speech clarity. Those interactions, in light of the results, seem to have raised undergraduate NESSs' awareness of accentedness in a way that is connected to speech clarity—such awareness is not evident in the perceptions of those who reported not having interactions with NNESTs yet. From a critical pedagogy lens, however, those interactions should "seek[ ] more equitable and liberating educational experiences" (Keesing-Styles, 2003) with the possibility of creating teachable moments where both parties can unpack the problematics of accentedness as a notion (as it compares to a matter-of-fact theme such as *English language proficiency and fluency*). Those problematics of accentedness present themselves in comments on a person's accent as a way to mark a racialized other (Shuck, 2006), particularly since "when people reject an accent, they also reject the identity of the person speaking: his or her race, ethnic heritage, national origin, regional affiliation, or economic class" (Lippi-Green, 1994, p. 165). So, in light of critical pedagogy, how can those problematics with accentedness create teachable moments? Examples of those teachable moments can present themselves in the form of exposing students to what scholars, such as Lippi-Green and Shuck (see above), have said about accents by way of debunking the myth that native speakers of English do not have accents and/or exploring why some nonnative accents are typically perceived in a more positive light than others (Lippi-Green, 2012) and/or how rejecting someone's accent implies a rejection of the speaker of the accent (Shuck, 2006). Instead of having the instructor's nonnative accent as the elephant in the room, our classrooms could be a space for an open, transparent conversation about effective communication as a mutual responsibility shared by the instructor and their students. That shared responsibility between the instructor and their students is one of the core values of critical pedagogy. Freire (2000) calls for solidarity and true communication between the teacher and their students, emphasizing that:

> One must seek to live *with* others in solidarity. One cannot impose oneself, nor even merely co-exist with one's students. Solidarity requires true communication…only through communication can human life hold meaning. The teacher's thinking is authenticated only by the authenticity of the students' thinking. The teacher cannot think for her students, nor can she impose her thoughts on them. Authentic thinking, thinking that is concerned about *reality,* does not take place in ivory towers of isolation, but only in communication (emphasis in original, pp. 76-77).

One example of such "reality" that our students and we can be concerned about is the accentedness that everyone brings to the classroom as an inseparable part and a reflection of their and our identity. In dialogue with their instructor, students can make sense of scholarship on accents and how those theories apply in their daily lives and encounters. Another example of such "reality" involves raising students' consciousness of the responsibility they have for effective communication to take place in the classroom. Encouraging students to reflect on the role they play and the labor

they exert in their knowledge making is in line with Freire's conception of teacher-student solidarity, pushing back against the banking model where students do not play a role in their knowledge making. Students' reality of knowledge-making should equitably involve all members of the learning community. That way, adopting critical pedagogy can prove helpful in creating a socially just, equitable environment where the labor of effective communication is shared among the classroom community.

Students' willingness, however, to exert effort for effective communication to take place was another distinction between the Experience Group's and the No Experience Group's interpretation of the question, with the former expressing either *little or no willingness to negotiate meaning* with NNESTs, while the latter hardly bringing up the idea altogether (See Table 2). Through a critical pedagogy lens, such distinction between the two groups brings attention to communicative labor inequalities in the US classroom when it comes to NNESTs-NESSs communications (in line with Rosina Lippi-Green's uneven distribution of communicative labor, 2012). The different reaction of the two groups might mean that actual interactions with NNESTs in the classroom could change how undergraduates do or perform meaning-negotiation, with a more labor-balanced perception of what it takes for successful communication to occur, especially with someone who does not necessarily sound like the self. With critical pedagogy being used as an intervention, students should feel comfortable asking their NNEST (or any instructor for that matter) or "asking others for clarification" (See Results section, Experience Group), such as classmates, whenever they experience confusion. After all, like Freire (2000) reminds us, "The teacher cannot think for her students, nor can she impose her thoughts on them. Authentic thinking, thinking that is concerned about *reality,* does not take place in ivory towers of isolation, but only in communication" (emphasis in original, pp. 76-77). In other words, meaning negotiation is central to critical pedagogy in a way that can inform question-posing, allowing it to be multidimensional rather than unidimensional in order for instructor-student solidarity to take form.

Though the results of this study are not generalizable, it could be inferred that actual interactions with NNESTs raise NESSs' awareness about meaning-negotiation, even when the latter express no willingness to exert effort for meaning-negotiation to take place with NNESTs. However, through a critical pedagogy lens, one might feel the urge to push back against such unidirectional knowledge-making settings, where NESSs are on the receiving end, and where the onus of meaning-negotiation is placed almost entirely on NNESTs' shoulders—most scholarship tends to focus on the preparation of NNESTs only, rather than involving NESSs in that preparation (Nemtchinova et al., 2010; Wu et al., 2010). With these findings in mind, one can propose critical pedagogy as a potential intervention to help North American institutions (re)envision the traditional preparation programs that NNESTs typically go through before they are released to teach without supervision. Those traditional preparation programs often treat NNESTs as a remedial population, placing them in accent-reduction and cultural adjustment courses, which often have an adverse impact on their confidence, sense of loneliness, and perceived linguistic deficiency (Gareis & Williams, 2004; Seloni, 2012). In line with previous studies (such as Subtirelu's 2015), this study offers implications on the lack of comfort NESSs might feel communicating with NNESTs whether to request clarification or to ask questions. What distinguishes this study, however, is that it highlights the distinction between the group of students who have had experiences with NNESTs (i.e., Experience Group) and those who have not (i.e., No Experience Group), particularly in terms of their perception of what it means to speak the English language clearly. The Experience Group either identified their own communicative responsibility in order for meaning negotiation to occur or acknowledged the communicative labor while absolving themselves of the responsibility to exert such labor. Since that willingness to negotiate meaning happens only after interacting with NNESTs, then those interactions would be additionally beneficial if they took place early on, before the

semester starts or during first-year student orientations—along the line of Kang and Rubin (2012) or Halleck (2008). Those early interactions might, then, positively inform NESSs' meaning negotiation skills as well as NNESTs' perceived comprehensibility. And, if those early opportunities for meaning negotiation are not feasible *before* the semester starts, then they should be adopted as a way to implement critical pedagogy as a strategy *during* the semester. Those strategies can take the shape of instructors explicitly guiding students or modeling for them as they learn and practice ways to request repetition and/or paraphrasing and to ask questions.

Given that critical pedagogy "is preoccupied with social injustice and examines and promotes practices that have the potential to transform oppressive institutions or social relations, largely through educational practices" (Keesing-Styles, 2003), critical pedagogy should not only inform the way SETs are designed by eliminating injustices and oppressive language used in the form, but also shape the rhetorical preparation that NNESTs and their NESSs need to communicate effectively. Not only that, but "[critical pedagogy] is primarily an educational response to oppressive power relations and inequalities existing in educational institutions. It focuses on issues related to opportunity, voice and dominant discourses of education and seeks more equitable and liberating educational experiences" (Keesing-Styles, 2003). With this definition in mind, critical pedagogy can be helpful in two ways: Top down and bottom up. From a top-down approach, the application of critical pedagogy in any institutional setting informs how the administrative governing bodies conceive of the SET forms, how its questions are articulated, and how the data is analyzed and for what purposes. In other words, a true embracement and integration of critical pedagogy by institutional governing bodies would shape the policies and practices of the institution in a way that cultivates a DEI mindset in driving the SET questions—rather than the current move by institutions to ostensibly add a few DEI questions to the existing form in order to reflect a surface-level effort at being inclusive and culturally sensitive.

Conversely, critical pedagogy can be helpful from a bottom-up approach, starting in our classrooms. To illustrate, an articulation of what pedagogy means is necessary. The traditional definition of pedagogy is "a deliberate attempt to influence how and what knowledge and identities are produced within and among particular sets of social relations" (Giroux & Simon, 1989, p. 239; cited in Keesing-Styles, 2003). In order for us to "seek [] more equitable and liberating educational experiences," the how and what of the knowledge-making that takes place in our classrooms need to be grounded in equitable practices, with the instructor and students collaborating on knowledge-making. If the labor that is placed in knowledge-making is skewed one way or the other, then we are not *truly* liberating either party. It would be hypocritical of us to promote an equitable and liberating mindset for knowledge-making and simultaneously argue that NNESTs need to be trained to communicate effectively in isolation from their students, while maintaining that their students do not share the same need. Preparing instructors in isolation from students can only be seen as a missed opportunity where both parties could have benefited from sharing the labor of practicing effective communication skills together.

Closing Thoughts and Limitations

This study was approached with the hope of understanding NESSs' ability to identify implicit biases towards NNESTs in the rhetoric of SETs and exploring the ways that their interactions with NNESTs might inform their abilities to do so. By drawing a comparison between the Experience and No Experience groups, the Experience Group was more capable than their No Experience counterpart at identifying the implicit bias in the rhetoric of the SET question at hand. Though both groups did not see themselves as responsible for effective communication to take place, the Experience Group expressed willingness to exert communicative labor that the No Experience

group hardly expressed. These results underscore the value of NNESTs-NESSs interactions in a way that, I argue, should inform how NNEST preparation programs should be conceived through a critical pedagogy lens, with students being an active, engaged party in that preparation. Those NNESTs-NESSs preparation programs should take the form of meaning negotiation, which can, of course, be taught explicitly or modelled by instructors for students to emulate. The absence of such meaning-negotiation on the part of students was shockingly evident in the results when students would "find [themselves] guessing what [NNESTs] have said or asking others for clarification." In other words, despite the proven value of peer learning and peer relationships (a topic that expands beyond the limits of this paper), students need to feel comfortable maneuvering discussions with all instructors, regardless of whether English is their first or additional language. Instructors also need to be reminded about the value of frequent comprehension checks. For instance, instructors should periodically ask students *what* questions they might have rather than *whether* they have questions. This way, students would find themselves more compelled to ask questions. Additionally, instructors need to be reminded about the value of creating rapport with students, which, by extension, enhances the approachability of instructors, so students feel comfortable to turn to them with questions.

Critical pedagogy can also prove helpful in preventing biases from presenting themselves in the rhetoric of SETs. And that is why I propose critical pedagogy as an intervention. In addition to its benefits at the classroom level (See the Discussion and Implications section and earlier in this Closing Thoughts and Limitations section), critical pedagogy can prove helpful on an institutional level, regardless of geographic location. With the fundamentals of critical pedagogy in mind, the administrative governing bodies can better conceive of the SET forms, articulate its questions, and reach a socially just understanding of how the data should be analyzed and for what purposes. Efforts to adopt diversity, equity, and inclusion (DEI) in (re)designing SETs should not be an afterthought, but a driving force that shapes not only the questions on the form, but the process of creating the form and who gets to sit at the table when such decisions are being made. Governing bodies should open a seat at the table for students to engage in the creation and/or revision of SETs. For example, students should have a representation on committees that are tasked with (re)designing SETs. Additionally, faculty who serve on those committees should remind themselves and educate others about the grounding principles of critical pedagogy in ways that can facilitate a deep engagement with and application of DEI.

However, like any study, this one comes with limitations. In addition to the small sample size being a limitation, additional facts about the nature of the recruitment of participants could have informed the results of this study: they all self-selected to participate, some of them were enrolled in introductory writing classes at the time, and all were recruited by the primary investigator of the study who also identifies as and linguistically presents as a NNEST. These limitations could have skewed the results of the study and could have been avoided by, for example, having a volunteer do the recruitment of participants and having that volunteer seek out participants from across the university. Despite these limitations, there is definitely value in implementing critical pedagogy in creating opportunities where NNESTs interact with their prospective students *before* the semester starts or in modeling meaning negotiation skills for our students *during* the semester. If anything, creating such interactive opportunities can only foster students' communicative abilities with any professor—not just NNESTs—and act as an intervention in raising student awareness about their responsibility in sharing communicative labor with others who might not look or sound like the self. Not only that, but critical pedagogy can create a space where undergraduates have a spot at the table where SET forms are being created or re-envisioned.

Acknowledgements

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of interest statement

The author declares that there are no relevant financial or non-financial competing interests to report.

Data availability statement

The data that support the findings of this study are openly available in Mendeley Data.

References

Baratta, A. (2018). "I speak how I speak:" A discussion of accent and identity within teachers of ELT. In B. Yazan & N. Rudolph (Eds.), *Criticality, teacher identity, and (in)equity in English language teaching: Issues and implications* (pp. 163-178). Springer International Publishing. https://doi.org/10.1007/978-3-319-72920-6_9

Berdie, D. R., Anderson, J. F., Wenberg, M. S., & Price, C. S. (1976). Improving the effectiveness of teaching assistants: Undergraduates speak out. *Improving College and University Teaching, 24*(3), 169–171. https://doi.org/10.1080/00193089.1976.9927343

Blackhart, G. C., Peruche, B. M., DeWall, C. N., & Joiner, T. E. Jr. (2006). Factors influencing teaching evaluations in higher education. *Teaching of Psychology, 33*(1), 37–39. https://doi.org/10.1207/s15328023top3301_9

Blunt, A. (1991). The effects of anonymity and manipulated grades on student ratings of instructors. *Community College Review*, *18*(4), 48–54. https://doi.org/10.1177/009155219101800409

Borjas, G. J. (2000). Foreign-born teaching assistants and the academic performance of undergraduates. *The American Economic Review*, *90*(2), 355-359. https://doi.org/10.1257/aer.90.2.355

Braa, D., & Callero, P. (2006). Critical Pedagogy and Classroom Praxis. *Teaching Sociology*, *34*(4), 357–369. https://doi.org/10.1177/0092055X0603400403

Charmaz, K. (2014). *Constructing grounded theory* (2nd ed.). London: Sage Publication Ltd

Creswell, J. W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Sage Publishing.

Ellis, K. (2004). The impact of perceived teacher confirmation on receiver apprehension, motivation, and learning. *Communication Education*, 53, 1-20. https://doi.org/10.1080/0363452032000135742

Fitch, F., & Morgan S. E. (2003). "Not a lick of English": Constructing the ITA identity through student narratives. *Communication Education 52*, 297–310. https://doi.org/10.1080/0363452032000156262

Fox, S., & Gay, G. (1994). Functions and effects of international teaching assistants. The *Review of Higher Education, 18*(1), 1–24. doi:10.1353/rhe.1994.0000

Freire, P. (2000). *Pedagogy of the Oppressed: 30th Anniversary Edition* (M. Bergman Ramos, Trans.). The Continuum International Public Group Inc. (Original work published 1970)

Gareis, E., & Williams, L. (2004). International faculty development for full-time and adjunct faculty: A program description. *Journal of Faculty Development*, *20*(1), 45-56.

Giroux, H., & Simon, R. (1989). Popular culture and critical pedagogy: Everyday life as a basis for curriculum knowledge. In H. Giroux & P. McLaren (Eds.), *Critical pedagogy, the state, and cultural struggle*. New York: SUNY.

Greenwald, A. G. & Gillmore G. M. (1997a). Grading leniency is a removable contaminant of student ratings. *American Psychologist*, *52*(11), 1209–1217. https://doi.org/10.1037/0003-066X.52.11.1209

Greenwald, A. G. & Gillmore, G. M. (1997b). No pain, no gain? The importance of measuring course workload in students' ratings of instruction. *Journal of Educational Psychology*, *89*(4), 743–751. https://doi.org/10.1037/0022-0663.89.4.743

Griffin, B. W. (2004). Grading leniency, grade discrepancy and student ratings of instruction. *Contemporary Educational Psychology*, *29*(4), 410–425. https://doi.org/10.1016/j.cedpsych.2003.11.001

Halleck, G. B. (2008). The ITA problem: A ready-to-use simulation. *Simulation & Gaming*, *39*(1), 137-146. https://doi.org/10.1177/1046878107308060

Halleck, G. B., & Moder, C. L. (1995). Testing language and teaching skills of international teaching assistants: The limits of compensatory strategies. *TESOL Quarterly*, *29*(4), 733–758. https://doi.org/10.2307/3588172

Hsu, C.-F. (2012). The influence of vocal qualities and confirmation of nonnative English-speaking teachers on student receiver apprehension, affective learning, and cognitive learning. *Communication Education*, *61*(1), 4-16. https://doi.org/10.1080/03634523.2011.615410

Hsu, C-F. (2014). Open and positive attitudes toward teaching. *New Directions for Teaching and Learning*, 138, 41-49. https://doi.org/10.1002/tl.20095

Huston, T. A. (2006). Race and gender bias in higher education: Could faculty course evaluations impede further progress toward parity? *Seattle Journal for Social Justice, 4*(2), 590-611. https://digitalcommons.law.seattleu.edu/sjsj/vol4/iss2/34/

Kang, O., & Rubin L. D. (2012). Intergroup contact exercises as a tool for mitigating undergraduates' attitudes toward nonnative English-speaking teaching assistants. *Journal on Excellence in College Teaching*, *23*(3), 159-166.

Keesing-Styles, L. (2003). The relationship between critical pedagogy and assessment in teacher education. *Radical Pedagogy, 5*(1), 1-19. https://hdl.handle.net/10652/1931

Kirsch, G. (1999). *Ethical dilemmas in feminist research: The politics of location, interpretation, and publication.* State University of New York Press.

Lippi-Green, R. (1994). Accent, standard language ideology, and discriminatory pretext in the courts. *Language in Society, 23*, 163–198. https://www.jstor.org/stable/4168513

Lippi-Green, R. (2012). *English with an Accent: Language, Ideology and Discrimination in the United States.* (2nd ed.). Routledge. https://doi.org/10.4324/9780203348802

Liu, J. (2005). Chinese graduate teaching assistants teaching freshman composition to native speaking students. In E. Llurda (Ed.), *Nonnative language teachers: Perceptions, challenges, and contributions to the profession* (pp. 155-77). Springer. https://doi.org/10.1007/0-387-24565-0_9

MacNell, L., Driscoll, A. & Hunt, A.N. (2015). What's in a name: Exposing gender bias in student ratings of teaching. *Journal of Collective Bargaining in the Academy, 0*(52)**,** 1-13. https://doi.org/10.1007/s10755-014-9313-4

Marsh, H. W., & A. Roche L. A. (2000). Effects of grading leniency and low workload on students' evaluations of teaching: Popular myth, bias, validity, or innocent bystanders. *Journal of Educational Psychology*, *92*(1), 202–228. https://doi.org/10.1037/0022-0663.92.1.202

May, T., & Perry, B. (2014). Reflexivity and the practice of qualitative research. In *The SAGE Handbook of Qualitative Data Analysis* (pp. 109-122). SAGE Publications, Inc., https://doi.org/10.4135/9781446282243

Miller, J., & Chamberlin M. (2000). Women are teachers, men are professors: A study of student perceptions. *Teaching Sociology*, *28*(4), 283–298. https://doi.org/10.2307/1318580

Mitchell, K. M. W, & Martin J. (2018). Gender bias in student evaluations. *PS: Political Science & Politics, 51*(3), 648-652. doi:10.1017/S104909651800001X

Munro, M. J., & Derwing T. M. (1995a). Foreign accent, comprehensibility in the speech of second language learners. *Language Learning*, *45*(1), 73-97. https://doi.org/10.1111/j.1467-1770.1995.tb00963.x

Munro, M. J., & Derwing, T. M. (1995b). Processing time, accent, and comprehensibility in the perception of native and foreign accented speech. *Language and Speech*, *38*(3), 289–306. https://doi.org/10.1177/002383099503800305

Mutua, C. N. (2014). Opposite worlds, singular mission: Teaching as an ITA. *New Directions for Teaching and Learning*, 138, 51-60. https://doi.org/10.1002/tl.20096

Nemtchinova, E., Mahboob A., Eslami Z., & Dogancay-Aktuna, S. (2010). Training non-native English speaking TESOL professionals. In A. Mahboob (Ed.), *The NNEST lens: Nonnative English speakers in TESOL* (pp. 222-238). Cambridge Scholars Publishing. https://doi.org/10.1111/j.1467-971X.2012.01782.x

Olivares, O. J. (2001). Student interest, grading leniency and teacher ratings: A conceptual analysis. *Contemporary Educational Psychology*, *26*(3), 382–399. https://doi.org/10.1006/ceps.2000.1070

Plakans, B. S. (1997). Undergraduates' experiences with and attitudes toward international teaching assistants. *TESOL Quarterly*, *61*(1), 95-119. https://doi.org/10.2307/3587976

Reid, L. D. (2010). The role of perceived race and gender in the evaluation of college teaching on RateMyProfessors.com. *Journal of Diversity in Higher Education*, *3*(3), 137–152. https://doi.org/10.1037/a0019865

Remedios, R., & Lieberman D. A. (2008). I liked your course because you taught me well: The influence of grades, workload, expectations and goals on students' evaluations of teaching. *British Educational Research Journal*, *34*(1), 91-115. https://doi.org/10.1080/01411920701492043

Roessel, J., Schoel, C., & Stahlberg, D. (2018). What's in an accent? General spontaneous biases against nonnative accents: An investigation with conceptual and auditory IATs. *European Journal of Social Psychology, 48*(4), 535–550. https://doi.org/10.1002/ejsp.2339

Rubin, D. L. (1992). Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. *Research in Higher Education*, *33*(4), 511-531. https://doi.org/10.1007/BF00973770

Seloni, L. (2012). Going beyond the native-nonnative English speaker divide in college courses: The role of nonnative English-speaking educators in promoting critical multiculturalism. *Journal on Excellence in College Teaching*, *23*(3), 129-155.

Shuck, G. (2006). Racializing the nonnative English speaker. J*ournal of Language, Identity & Education, 5*(4), 259–276. https://doi.org/10.1207/s15327701jlie0504_1

Subtirelu, N. C. (2014). A language ideological perspective on willingness to communicate. *System*, 42, 120-132. https://doi.org/10.1016/j.system.2013.11.004

Subtirelu, N. C. (2015). "She does have an accent but…": Race and language ideology in students' evaluations of mathematics instructors on RateMyProfessors.com. *Language in Society*, 44, 35-62. https://doi.org/10.1017/S0047404514000736

Takayoshi, P., Tomlinson E., & Castillo J. (2012). The construction of research problems and methods. In K. M. Powell & P. Takayoshi (Eds.), *Practicing research in writing studies: Reflexive and ethically responsible research* (pp. 97-122). Hampton Press Inc. http://www.worldcat.org/oclc/828736036

Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American Journal of Evaluation*, *27*(2), 237–246. https://doi.org/10.1177/1098214005283748

Villarreal, D. (2013). Closing the communication gap between undergraduates and international faculty. *CATESOL Journal, 24*(1), 8–28. https://eric.ed.gov/?id=EJ1111863

Wang, L., & Gonzalez, J. A. (2020). Racial/ethnic and national origin bias in SET. *International Journal of Organizational Analysis*, *28*(4), 843–855. https://doi.org/10.1108/IJOA-06-2019-1793

Woolard, K. A., & Schieffelin, B. B. (1994). Language Ideology. *Annual Review of Anthropology*, *23*, 55–82. https://doi.org/10.1146/annurev.an.23.100194.000415

Wu, A., Liang J., & Csepelyi, T. (2010). Coping strategies for NNES' teacher development. In A. Mahboob (Ed.), *The NNEST lens: Nonnative English speakers in TESOL* (pp. 202-221). Cambridge Scholars Publishing. https://doi.org/10.1111/j.1467-971X.2012.01782.x