



The Impact of Rubric Use on Peer Feedback Tasks in a TESOL Practicum

Kristen Lindahl, *University of Texas at San Antonio*

MaryAnn Christison, *University of Utah*

Zuzana Tomaš, *Eastern Michigan University*

Abstract

This sequential explanatory mixed-methods study examines the impact of analytic rubric use in peer feedback on preservice teachers' ability to recognize indicators of best practice for second language lesson planning and lesson delivery. Fifty-three preservice teachers in a university-level, semester-long TESOL practicum course received direct instruction on indicators presented in the analytic rubrics. They were then randomly assigned to control and experimental groups. The experimental group used rubrics with the indicators during peer feedback tasks, while the control group used a modified rubric without the indicators. The result from an independent samples t-test on posttest mean scores indicated a significant difference between groups for both lesson planning and lesson delivery, favoring the experimental group ($p=.012$). Qualitative data were also collected via written comments on the posttests and from focus-group interviews. From thematic analyses of qualitative data, three key themes emerged, including specific tensions that resulted from the type of feedback preservice teachers desired and the type of feedback they were willing to give to their peers. These findings provide further insight into the use of analytic rubrics in peer feedback practices in second language teacher education (SLTE).

Keywords: second language; teacher education; preservice teachers; practicum; rubrics

Introduction

The practicum course has long been a staple in programs that prepare educators to work in Teaching English to Speakers of Other Languages (TESOL) contexts, and it is typically the point in most programs wherein preservice teachers begin transferring the knowledge gained from coursework into actual teaching practice (Crookes, 2003; Richards & Crookes, 1998). The instructional goals for TESOL practicum courses are often multi-layered, complex, and vary depending on the context and the needs of the preservice teachers who are involved. Teacher educators must decide how to guide TESOL practicum students in developing basic teaching

skills, such as being able to recognize specific indicators of best practice for lesson planning and delivery, and to then implement these indicators in their own teaching. Through these activities, practicum students are meant to deepen their understandings of teaching and reflect on and learn from their practice. Determining how to carry out these tasks can be challenging for teacher educators, and the expectations for what preservice teachers should be able to achieve in one TESOL practicum course are often quite high (Canh, 2014).

Two pedagogical practices that we use in helping preservice teachers recognize and reflect on the complex interplay between knowledge of teaching and the development of pedagogical skills are the use of analytic rubrics (also referred to as “marking criteria” or “grading criteria”), which are intended to guide preservice teachers in recognizing indicators associated with best practices in second language (L2) teaching, and collaborative practices, such as peer feedback, peer observations, and guided small group discussions about teaching. While both are common practices in teacher preparation, little research exists on the potential relationship between the two. As such, operating from a sociocultural theoretical framework, the research presented in this article investigated the impact of analytic rubrics on peer feedback tasks in a TESOL practicum. Practicum students participated in microteaching demonstrations as both teachers and students and reflected on their experiences in guided discussions with peers and teachers. As researchers, we hypothesized that rubric use would in fact improve preservice teachers’ ability to identify the indicators of best practices for L2 lesson planning and lesson delivery that were the focus of instruction in the course. We also supposed that thematic analysis of focus-group interviews with participants and of written comments would yield depth and richness to the explanation of quantitative results.

Review of the Literature

Theoretical Support

As L2 teacher educators, we believe that the choices we make in guiding preservice teachers must be responsive to how we understand both the work of teaching and the role that teacher knowledge plays in the development of pedagogical skills (Freeman, McBee, Orzulak, & Morrissey, 2009). We see the theoretical underpinnings of rubric use and collaborative teaching practices as grounded in sociocultural frameworks of knowledge construction, which promote teacher learning via scaffolded social interactions (Johnson, 2006, 2009; Johnson & Golombek, 2011; Shabani, Khatib & Ebadi, 2010), and in teacher cognition (Borg, 2003, 2006; Calderhead,

1996; Freeman & Richards, 1996; Munby, Russell, & Martin, 2001). We define *teacher cognition* as “the unobservable cognitive dimension of teaching—what teachers know, believe, and think” (Borg, 2003, p. 81), and particularly in the case of preservice teachers, we are interested in how it evolves over time and in response to learning about and practicing the craft of teaching.

Sociocultural theory (Vygotsky, 1978) stresses that human learning is largely a social process; therefore, the interactions that take place between and among individuals in a given context, such as a TESOL practicum, are necessary components of learning. We designed the TESOL practicum course so that learning to teach would involve preservice teachers in regular and frequent observations of and interactions about teaching through the use of peer observations and peer feedback. We see these activities as integral mechanisms for understanding teaching because they involve teachers in focused interactions about teaching. A tenet of Vygotsky’s sociocultural theory is the Zone of Proximal Development (ZPD) (1978), which is the idea that learning and mastery are achieved when a learner operates at a level slightly above what he or she can do independently because the learning experience is *scaffolded* or supported by more knowledgeable others (in this case, either the instructors of the practicum course or one’s peers in the class). Thus, peer collaboration among practicum preservice teachers has the potential to mediate learning experiences in classroom settings and enrich it in ways that independent practice may not (Tudge, 1992).

Collaborative teacher development practices are quite well documented in SLTE (see, for example, Johnston, 2009), but the use of rubrics in collaborative teacher development practices, such as a peer feedback task, is not. Even though the value of rubrics has been widely recognized across educational contexts, most of the existing research on rubric use is related to the construction and design of rubrics, the construct validation of rubrics, the determination of interrater reliability, and the generalizability of rubric use for different contexts and with varied groups of learners (Hafner & Hafner, 2003). Furthermore, existing research has examined how instructors use rubrics as they evaluate student performances, rather than on how students (i.e., preservice teachers) might use rubrics for the development of both their knowledge of teaching and their pedagogical skills.

A major concern for us as L2 teacher educators is how preservice teachers’ beliefs and perceptions of teaching evolve as a result of early teaching experiences. In the TESOL



practicum, these experiences include direct instruction on teaching, opportunities for instructional planning, microteaching (i.e., short teaching demonstrations), observations of peer teaching, participation in guided discussions with peers and the instructor, as well as experiences in classrooms under the guidance of an experienced teacher. An investigation into preservice teacher cognition is an important concept for L2 teacher educators because we see practicum students as “active, thinking decision-makers who make instructional choices by drawing on complex practically-oriented, personalized, and context-sensitive networks of knowledge, thoughts, and beliefs” (Borg, 2003, p. 81), and “it is obvious that what teachers do is directed in no small measure by what they think” (National Institute of Education, 1975, p. 1).

Peer Feedback Tasks

We use the term *peer feedback* to refer to a process in which peers (in this case, TESOL practicum students) are involved in observing, providing feedback, and discussing their joint teaching endeavors. We see peer feedback as an important process because it assists preservice teachers in developing ownership over their own teaching, thereby assisting them in transitioning from university student to practicing teacher.

Research indicates that preservice teachers reap important benefits from participation in peer feedback tasks. For example, they receive more feedback and are more likely to incorporate that feedback into future performance tasks (Cartney, 2010). In addition, Ertmer et al. (2007) found that peer feedback was useful in promoting higher-order discourse during peer discussions, while Evans (2015) found that peer feedback tasks were useful in promoting teacher agency—the capacity to take actions on one’s own. Additional studies have found that peer feedback can have positive effects on the development of depth of knowledge in content-areas, for example science and math (Beaver & Beaver, 2011; Kilic & Cakan, 2007), demonstrating that peer feedback can be useful in supporting teacher development across disciplines. However, in their meta-analysis of 41 quantitative studies, Falchikov and Goldfinch (2000) found only three studies that addressed the use of peer feedback in the field of teacher education, and none addressed SLTE specifically.

Peer feedback tasks should be guided to be useful in promoting reflective conversations on teaching practices. Unguided peer feedback can create problems that could outweigh potential benefits (Lasater, 1994). Because TESOL practicum students are in the process of developing

both their knowledge of teaching and their understandings of teaching, there is high probability that unguided feedback to peers may be based purely on personal beliefs (Andrade, 2000; Hasbrouck, 1997; Vacilotto & Cummings, 2007) rather than on research-based features of L2 instruction that promote learning.

Analytic Rubrics

An *analytic rubric* is an instructional tool that can be used for guiding peer feedback; it breaks down the characteristics of a task into its composite parts, specifies the grading or marking criteria for the task, and describes levels of quality for performance (Abbott, 2013). Hack (2015) notes that the two most important benefits of analytic rubric use in teacher education are to increase teachers' abilities to self-reflect and develop greater engagement with the criteria. Still, Wöllenschläger, Hattie, Machts, Möller, & Harms (2016) argue that making grading or marking criteria for learning transparent through the use of analytic rubrics is not sufficient to guarantee that rubric use will be effective in promoting teacher development. Research suggests that preservice teachers be taught how to use a rubric to conduct peer feedback. Sluijsmans, Brand-Gruwel, and Van Merriënboer's (2002) analysis of feedback samples indicate that preservice teachers who are taught how to use rubrics are able to respond to diverse criteria, give constructive comments, and use more discipline-specific words in describing teaching behaviors. Struyen, Dochy, and Janssens (2008) conclude that familiarity with peer feedback tools positively impacts preservice teachers, thereby building self-confidence in teaching and enhancing skills in managing classrooms (Wilkins, Shin & Ainsworth, 2009).

Based on the potential merit of using analytic rubrics during peer feedback tasks, as well as the relative shortage of studies that address this practice in SLTE, the following research questions guided the current study:

1. To what extent does analytic rubric use during peer feedback tasks impact preservice teachers' ability to recognize indicators of best practice associated with L2 lesson planning?
2. To what extent does analytic rubric use during peer feedback tasks impact preservice teachers' ability to recognize indicators of best practice associated with L2 lesson delivery?



3. What are preservice teachers' beliefs, attitudes, and perceptions about the usefulness of utilizing rubrics during peer feedback tasks for their own development as teachers?

Methodology

We employed a sequential explanatory mixed methods approach in the design of the research, which Creswell, Plano-Clark, Gutmann, and Hanson (2003) characterize as analyzing quantitative data first, followed by the analysis of qualitative data in the interest of explaining, interpreting, and expanding upon the quantitative data. The sequential analysis of data in this study is a slight departure from Creswell, et al (2003) in that we needed to analyze qualitative data from the posttests and attach numeric values to the responses to obtain quantitative data. We then turned to the qualitative data from the focus groups.

To answer Research Questions 1 and 2, we employed an experimental research design to determine the extent to which analytic rubric use during peer feedback tasks impacted preservice teachers' abilities to recognize indicators of best practice associated with L2 lesson planning and delivery. The indicators were selected for the rubric because they were the focus of the direct instruction during the first four weeks of the course. To answer Research Question 3, we transcribed and analyzed data from the focus group.

Participants

The participants were 53 preservice teachers in a semester-long L2 teaching practicum course at a large research university in the United States. They were enrolled in either an English as a Second Language (ESL) endorsement program (a component of K-12 teaching licensure in this particular state) or an undergraduate TESOL Certificate. Participants were aged 25 to 50 and represented six different first language (L1) backgrounds—Bulgarian, Japanese, Korean, Mandarin, Slovak, and Spanish. Thirty percent of participants were L2 speakers of English. All L2 speakers of English had satisfied at least the minimum Test of English as a Foreign Language (TOEFL) score required for admission to the university (TOEFL iBT minimum score of 80). A needs analysis survey conducted prior to instruction confirmed that the preservice teachers in the course had less than one year of teaching experience. None of the participants had taught full-time and most were in the process of completing the 60-hour field experience requirement (i.e., hours spent in ESL classrooms or in content classrooms with English



Learners [ELs] under the supervision of a qualified teacher) that was required for the TESOL Practicum.

Course

The TESOL practicum course provided students with opportunities to: (1) observe teachers in classrooms with ELs, (2) plan lessons for ELs or for courses that include ELs, (3) deliver lessons in short teaching demonstrations to the practicum instructor and their peers (hereafter referred to as microteaching demonstrations), (4) receive constructive feedback on microteaching demonstrations and lesson plans, (5) receive instruction on rubric use, and (6) participate in guided discussions about teaching. The course met 180 minutes weekly for 16 consecutive weeks.

Materials

Creating rubrics for lesson planning and delivery involves selecting a limited number of quality indicators, which is a challenging task for L2 teacher educators because there are so many possible indicators that could be included in a rubric for preservice teachers, many of which are context-dependent. For example, Lemov (2010) proposed 49 indicators of effectiveness, and Marzano (2005) proposed nine strategies and 34 specific indicators. In the Sheltered Instructional Observation Protocol (SIOP) Model (Echevarria, Vogt, & Short, 2018), which is specifically focused on indicators of effective sheltered content-area instruction for teachers of ELs, there are 30 indicators of effective instruction organized under eight components. Because preservice teachers tend to have little formal experience in teaching and are at early stages of developing their knowledge about teaching, they could likely benefit from a focus on any of the indicators in these frameworks. However, in our experience as L2 teacher educators, preservice teachers can be easily overwhelmed with the amount of information they are given about teaching and the number of tasks they are asked to manage. Therefore, it was important to limit the number of indicators if the rubrics were to be useful for the purposes of identifying quality indicators.

The rubrics used during the peer feedback tasks for this study were originally designed by a senior faculty member who had taught and supervised the L2 practicum for many years, and they were refined over time with input from other practicum instructors. The rubrics were also used as templates for planning. There is some evidence to suggest that for an analytic rubric to



be useful in peer feedback, it is important to attach a numeric value to each level of quality (Andrade, 2000; Reddy & Andrade, 2010), so the rubrics used in the practicum course were constructed with numeric values from 0 to 2 (i.e., 0 = ineffective, 1 = moderately effective, and 2 = effective) to rate each construct and accompanying sets of indicators.

The lesson-planning rubric included 20 indicators embedded in five general constructs—describing context, creating performance objectives, identifying stages in lesson planning (i.e., warm-up, introduction, presentation, practice, evaluation, application), recognizing learner differences, and effective organization and presentation of materials. For example, indicators for the construct “creating performance objectives” were (1) clearly stating what learners will be able to do and (2) relating objectives to the content of the lesson.

The constructs for lesson delivery focused primarily on concepts that are central to the process of second language acquisition (SLA)—making input comprehensible to learners and providing opportunities for interaction (Gass, Mackey, & Pica, 1998). The rubric consisted of 16 indicators embedded in six constructs—L2 teacher talk, giving instructions, selecting instructional activities and tasks, evidence of preparedness, teacher attitude, and organization of the materials.¹ Some example indicators for the construct of L2 teacher talk might include (1) using a rate of speech appropriate to the proficiency level of the learners, (2) pausing at appropriate thought junctures, and (3) placing stress on important information.

Procedures

All participants received four weeks (180 minutes a week) of direct instruction on lesson planning and delivery, which specifically focused on the constructs and indicators presented in the rubrics. Modeling was an important feature during the direct instruction phase. In addition, preservice teachers critiqued videotaped teaching demonstrations using the rubrics, reviewed example lesson plans, and participated in guided discussions with their peers. After four weeks of direct instruction, participants were randomly assigned to either an experimental group ($n = 26$) or a control group ($n = 27$). Each group met for 90 minutes for 11 weeks. Each student developed two lesson plans and presented a portion of each lesson to peers in two microteaching demonstrations.

In the experimental group, participants used rubrics with the indicators during peer feedback tasks, and they were asked to assign a numeric value to the performance on each indicator. They were also encouraged to write comments on the rubric in addition to assigning a numeric value. In the control group participants used a modified rubric without the indicators and wrote comments to their peers based on the constructs. Participants in both groups also provided feedback on drafts of the lesson plans, with the experimental group using the rubrics with the indicators and the control group using the modified rubrics with only the constructs. Participants in both groups had access to the rubrics during the creation of lesson plans and the preparation of their microteaching demonstrations.

At the end of each class period, participants engaged in guided discussions. The discussions were integral to the development of a reflective practice and were intended to build a climate of trust and support among peers in which they could talk openly about teaching and share their successes, challenges, and concerns. To accomplish this goal, the instructor framed the discussions using the following three prompts: (1) tell the group something you liked about _____'s demonstration; (2) give _____ one suggestion for improvement; and (3) tell us something you learned about teaching today by watching your peers teach. After the guided discussions the rubrics were given to peers who had presented so they could benefit from both oral and written feedback.

The posttest for lesson planning. The posttest for lesson planning was given at the end of 16 weeks of instruction. The participants in both groups read a model L2 lesson plan in which the 20 quality indicators were present and had been identified by the researchers. Participants were asked to provide written feedback, identifying what they believed to be the quality indicators of lesson planning. Prompts were provided that corresponded to each of the constructs (e.g., comment on the performance objectives in the lesson plan).

The posttest for lesson delivery. The posttest for lesson delivery was also given at the end of 16 weeks of instruction. All participants watched a 15-minute video of teaching in an ESL classroom in which the 16 quality indicators were present and had been identified by the researchers. Participants were asked to provide written feedback on the video teacher's performance, identifying what they believed to be the quality indicators of lesson delivery.



Prompts were provided that corresponded to each of the constructs (e.g., comment on the language the teacher used to talk to the students).

Focus groups. To obtain qualitative data about preservice teachers' beliefs and perceptions about the usefulness of peer feedback for their own development as teachers, a subset of participants were asked to meet with at least one of the researchers in small focus groups at the end of the 16 weeks. The interviews were semi-structured, consisting of five main questions (see Appendix A). Researchers asked additional questions of the participants, depending on the responses to the initial questions. Focus groups lasted approximately 30 minutes and 13 students participated. All interviews were recorded digitally and transcribed.

Analyses

Trained raters first analyzed the participants' written responses to the posttests on lesson planning and delivery by classifying preservice teachers' statements into *a priori* categories (10 constructs on lesson planning and six on lesson delivery). Raters further analyzed the data to determine how many of the quality indicators participants were able to identify for each construct, and this number served as the score. Because the data were nominal, we calculated interrater reliability (i.e., the degree of agreement among raters), by using joint probability of agreement. Ninety-three percent agreement was achieved for lesson planning and 94% for lesson delivery. Then, an Independent Samples t-test was calculated to determine the effect of rubric use during peer assessment tasks on participants' ability to identify indicators of effective L2 lesson planning and delivery.

Qualitative data from the focus group interviews were transcribed and analyzed thematically (Miles & Huberman, 1994). The researchers then employed a constant comparison method (Glaser, 1965) (i.e., a process in which newly collected data are compared with previous data or with data from previous studies) to the analysis of the initial concepts until stable categories for the concepts were established. Finally, the researchers identified how the concepts were united thematically and labels were attached to represent themes (Patton, 2002).

Results

Quantitative Data

The results from quantitative analyses on lesson planning and lesson delivery tasks,

respectively, are described below. In addition to the Independent Samples t-test, we used Cohen's d to calculate effect size to determine the meaningfulness of the difference between groups. In Cohen's d , two groups must differ by at least 0.2 standard deviations; otherwise, the difference between the means is trivial and not meaningful, even if it is significant.

Lesson planning. There was a significant difference in the scores between the experimental group, ($M = 9.04$, $SD = 5.65$) and the control group, ($M = 5.61$, $SD = 3.71$), $t(51) = 9.746$, $p = .012$. The effect size for lesson planning is medium (Cohen's $d = 0.5103$). There were 20 indicators in total for lesson planning, and the mean scores for the groups were 5.61 for the control group and 9.04 for the experimental group.

Lesson delivery. There was a significant difference in the scores between the experimental group, ($M = 8.331$, $SD = 4.98$) and the control group, ($M = 6.42$, $SD = 3.82$), $t(51) = 8.956$, $p = .022$. In addition, we calculated effect size for lesson delivery (Cohen's $d = 0.430$). The effect size for lesson delivery is between small and medium.

Qualitative Data

We conducted an analysis of qualitative data from the comments made during the interviews during the focus groups. Five themes emerged from these data: (1) perceived usefulness of rubrics during peer feedback tasks for the development of pedagogical reasoning, (2) potential for the use of rubrics in the development of future pedagogical tasks for language learners, (3) the limitations of rubric use during peer feedback, (4) giving versus receiving feedback, and (5) oral versus written feedback. All participants in the focus groups were assigned pseudonyms.

Pedagogical reasoning. Participants noted that giving and receiving feedback from peers influenced their thinking and developed their pedagogical reasoning during the process of lesson planning.

You have worked with your lesson so much that you just can't see anything else to do with it; it's perfect. So, it's nice to have a fresh perspective, to know what you can do to change your lesson. Because a lot of what people have suggested I had not thought of. (Anna)



Doing it [peer feedback] helps me solidify in my mind when I like something. I like to detail that in writing so that it stays with me, and I can remember it better and maybe incorporate it in my own teaching. (Alan)

These excerpts from the data support the notion that the process of lesson planning is important in the development of pedagogical reasoning skills (Pang, 2016), but the process of peer feedback may also be a critical piece. Preservice teachers may benefit from exposure to perspectives that are different from their own (Koc, 2011), as well as from the cognitive demands of assessing a peer's work in a written form.

Future pedagogical tasks. Beyond seeing the value of peer feedback in their immediate coursework, preservice teachers also began to consider the use of rubrics and peer feedback as they imagined working with their future language learners. Of the 13 preservice teachers who were interviewed, 11 said they planned to use rubrics to provide feedback to their future students, and three participants noted that they were also thinking about how students might provide feedback to one another. Participants also commented further on the specific value of rubrics for learning. They indicated that they viewed a rubric as an important tool because it served as more than a guide. It was “a checklist that helped identify important features of lesson planning and delivery,” (Tina) but it also provided a breakdown of established criteria that could help teachers “evaluate student performance fairly and communicate a teacher's expectations” (Lee) to the students. These types of comments indicate that preservice teachers both recognized the utility of rubrics and also noticed that rubrics could be used pedagogically to further their own practice, similarly to the teachers in Wilkins, Shin, and Ainsworth (2009).

Limitations of analytic rubric use. Some participants expressed concerns related to limitations of peer feedback and rubric use relative to the age and proficiency level of learners.

If rubrics are adapted, I think they can be used with beginning level language learners, and with encouragement can help each other (Larisse)

I think rubrics are great for kids because then they know what is expected of them. I'm planning to work with younger elementary kids, so I'd have to simplify the rubric and



perhaps not [ask learners] give a rating. I would try to use pictures, like a smiley face.
(Arielle)

Although they were identifying potential obstacles to the efficacy of using rubrics and peer feedback, they were also thinking about how to adapt the process.

Giving versus receiving peer feedback. Although participants made overall positive claims about their experiences in using peer feedback, an analysis of focus group data suggests that tensions existed between participants' perspectives as givers and receivers of peer feedback. There was an underlying reluctance to give critical feedback to peers, with about a third of the interviewees admitting that they avoided giving critical feedback to their peers even when they noticed a problem.

The person I evaluated said she felt she didn't get enough feedback because I was trying to be nice. So, it's kind of hard... I think it's because of my culture and personality.
(Ikuko)

I never felt that confident doing my own lesson plan, so I don't feel good giving comments on anyone else's. (Nick)

Sometimes it's hard to give feedback to someone from the same cultural background. Because we are friends, I know them. So, it's hard if I give them like 0 points. (Kim)

It seems that the cultural backgrounds of the participants, their confidence relative to their own skills, and personal friendships played a role in participants' uneasiness in providing feedback to peers.

Despite their reluctance to provide feedback to their peers, participants tended to value only constructive, critical, and specific feedback on their own work. Some participants expressed dissatisfaction, even frustration, with inadequate or superficial feedback from peers.

It's very interesting to see others' comments, but, sometimes, it does not help at all because, honestly, [peers] did not put much effort and time in giving feedback. Just "good", "nice," and stuff like that. It's not helpful if we want to be good teachers. (Sujin)

When I read a comment and it says, “This is so wonderful,” I always think: What worked? What touched you? Give me something to work off of! (Alan)

While participants clearly acknowledged that it was not easy to be critical of peers’ work, most preservice teachers who were interviewed agreed that they had made progress in giving feedback and stated that their skills had improved by the end of the course, indicating that the culture of resisting critical or constructive feedback was changing.

At first I was giving a lot of compliments, but now I kind of know how to do it. I now know that it’s not going to help them if I just say “good, good, good.” Now I have my point. (Kim)

Personally, when I used rubrics to evaluate, I felt awkward because I am a peer. I did not want to offend my fellow students, but I had to realize that as a teacher I have to evaluate students too. So, it’s a good thing to learn. (John)

Data also show that preservice teachers’ confidence in giving feedback evolved over time (Grainger & Adie, 2014) and that they also offered more constructive feedback to their peers on both lesson planning and delivery as the course progressed. Table 1 exemplifies the increase in the degree of sophistication in feedback between the first and the second lesson plans.

Table 1: *Differences in Qualitative Feedback on Lesson Plans*

Student/Stage	Feedback on Feedback from Lesson Plan 1	Feedback on Feedback from Lesson Plan 2
Ikuko on the warm-up/ review stage	2 points	Maybe you can create OHT for the K-W-L Chart and write down some examples in each column as you explain the instruction. Since there is no instruction OHT, demonstration/giving examples will be very important.
Dana on the Practice Stage	Very engaging activity	I think that maybe groups of 3-4 would work better in this case, because they would be able to come up with more examples.
Arielle on the Application stage	words→ sentences	Maybe they could do something further with the nouns they write down?

These example data illustrate how preservice teachers evolved from offering general non-specific feedback to providing pedagogical suggestions on performance. The evolution of the comments aligns with Wöllenschläger, et al's (2016) claim that analytic rubrics not only provide information about the performance itself, but also help users develop skills on future tasks.

Oral versus written feedback. Another theme that emerged from the qualitative data was participants' preference for the oral language feedback that resulted from participation in the guided discussions.

In the class, you can speak directly with the person, and they can't misinterpret what you're saying ... I feel that when my peers write it out, perhaps you are more inclined to be meaner. (Sam)



I think it's better if I talk with the person and then tell the person what she can do better. But when I'm writing, maybe she doesn't know the tone of my voice, and I tried to write in a nice way and that's not really giving good feedback. (Ikuko)

More than half of the participants in the focus groups asserted that receiving written peer feedback in class and receiving comments in *situ* were useful, but they valued the informal face-to-face interaction with their classmates. They also acknowledged that, with oral feedback, they did not feel nervous or worry about receiving overly critical comments from peers.

Discussion

Our first and second research questions explored the extent to which analytic rubric use during peer feedback tasks impacted preservice teachers' ability to recognize indicators of best practice associated with L2 lesson planning and delivery. The quantitative data analyses yielded confirmatory results about the use of analytic rubrics during peer feedback in a TESOL practicum. As a result of rubric use during the peer feedback process for both lesson planning and delivery, participants in the experimental group were able to recognize significantly more indicators of best practice than the control group. While these results are encouraging, it is also important to consider the results in context. It seems that the use of analytic rubrics made the indicators more transparent for the experimental group; nevertheless, participants were only able to identify about half of the indicators (i.e., for the experimental group 9/20 indicators for lesson planning and 8.3/16 for lesson delivery) they had been working with during direct instruction and through the peer observations.

As teacher educators, we were initially disappointed with this result until we asked ourselves whether the expectations we had for the development of preservice teachers' skills were overly ambitious. The quantitative data confirm that preservice teachers need more time and experience observing and participating in teaching to develop their skills in recognizing indicators of best practice, and, certainly, it is reasonable to conclude that teachers will not learn everything they need from their initial education (Hodgson, 2013). Further studies should analyze experienced teachers' skills in identifying indicators of best practice to determine a reasonable trajectory for the development of preservice teachers' skills.

Teaching skills develop in incremental steps; there will always be a need for ongoing professional development as learning about teaching continues throughout one's professional life. These data provide support for current models of teacher development that provide room for teachers to eventually become “reflective sense-makers” (Borg, 2006, p.15), learning from experience and ongoing reflection on their teaching and for research on teacher education that attributes teacher expertise to the development of both theoretical and practical knowledge (Shulman, 1986).

As teacher educators, we are also cognizant of the fact that preservice teachers' abilities to identify indicators of best practice as a result of the peer feedback process may also be dependent on the practicum instructor because it is this individual who is responsible for the initial modeling of the targeted skills and demonstrating explicitly how to apply them in particular contexts (Pleogh, Tilema & Segers, 2009). The results of the current study support studies in the existing literature that the use of rubrics promotes learning by making performance criteria explicit, as seen in Hack (2015) and Jonsson and Svingby (2007). While rubrics may make learning goals transparent, that benefit alone might not be sufficient to consider them universally effective across contexts (Wöllenschläger et al, 2016). We also wanted to examine the usefulness of rubrics beyond the benefit of the transparency.

Our third research question regarded preservice teachers' beliefs, attitudes, and perceptions about the usefulness of utilizing rubrics during peer feedback tasks for their own development as teachers. Qualitative data analysis provided additional insights into how the use of analytic rubrics in peer feedback tasks influenced preservice teachers' perceptions. The themes derived from the qualitative data indicate that preservice teachers demonstrated a positive orientation towards the collaborative teacher practices that were embedded in the TESOL practicum—peer observations, guided discussions about teaching, and the use of analytic rubrics in the peer feedback process. Data analyses also revealed that there were tensions inherent in the process of providing peer feedback, such as the fact that preservice teachers were reluctant to give constructive or critical feedback to their peers, while at the same time they wanted this type of feedback from their peers. Reluctance to provide feedback may be attributed to the possibility that participants may still be seeing their practicum instructor in the traditional role of “sage

on the stage.” A similar tendency emerged in a study of Irish preservice teachers, who expressed an appreciation for peer feedback activities but continued to prefer the instructor’s feedback to other types (Lynch, McNamara, & Seery, 2012). To resolve this tension, teacher educators must work with preservice teachers to involve them in discussions about teacher development and the rationale behind the use of peer feedback, as well as provide modeling for the practices in which they want preservice teachers to participate (Brew, 2009).

Perhaps it is not surprising that in the context of a university TESOL practicum, preservice teachers’ experiences with peer feedback are better captured when juxtaposed as a series of tensions. For many preservice teachers enrolled in a practicum course, providing and receiving peer feedback on lesson planning and delivery may be their first encounter with performance tasks in the role of a teacher and the critical instructional feedback that tends to follow such performances. Even though studies have established that peer feedback is reliable in comparison to the instructor’s feedback (Kilic & Cakan, 2006), preservice teachers may lack confidence as providers of feedback (Li, Liu, & Steckelberg, 2010) and may not have confidence in their peers. L2 teacher educators can facilitate the process of building confidence among preservice teachers by clearly delineating the role and purpose of a rubric-based peer feedback task, modeling its use, and encouraging reflectivity on preservice teachers’ performances through guided discussion. As such, future research on rubric use and/or peer feedback tasks could investigate how preservice teachers implement [indicators of best practice](#) and use analytic rubrics and peer feedback as they move into real classrooms with their own students.

Conclusion

The results of the current study indicate that analytic rubric use in peer feedback tasks on L2 lesson planning and delivery positively impacts preservice teachers’ abilities to identify quality indicators. Although this study does not provide support for the notion that knowledge of the quality indicators results in their implementation of these indicators in actual teaching practices, it is true that for teachers to transfer observed practices to their own classrooms, the first step is “seeing” and recognizing these practices in observations of teaching. It is important for L2 teacher educators who work with preservice teachers to incorporate activities in practicum courses that promote the noticing of indicators of best practice, as well as provide opportunities for preservice teachers to reflect on the pedagogical behaviors in their own teaching and as they



observe their peers. We suggest that the use of such practices in TESOL practicum courses may result in more robust connections between the information presented during practicum coursework and what teachers actually do in the field as practicing teachers. The current study shows that when preservice teachers participate in the process of identifying indicators of best practice in peer feedback tasks using analytic rubrics, they exit the TESOL practicum course with a heightened awareness of what constitutes effective L2 instruction.

Biodata

Kristen Lindahl, Ph.D. is Assistant Professor of Applied Linguistics/TESL in the Department of Bicultural-Bilingual Studies at the University of Texas at San Antonio, USA. She conducts research on teacher language awareness and how it relates to issues of identity and ideology in teacher education and ongoing career development.

MaryAnn Christison, Ph.D. is Professor of Linguistics in the Department of Linguistics and the Urban Institute of Education at the University of Utah, USA. Her research program draws on a wide range of theoretical and methodological approaches from various disciplines to understand the role that language plays in society, with her current research focusing on the use of online technologies in language teacher education.

Zuzana Tomas, Ph.D. is Associate Professor of TESOL in the Department of World Languages at Eastern Michigan University, USA. She conducts research in the areas of second language writing, supporting international student success in higher education contexts, and the role of community-engaged learning in teacher preparation.

References

- Abbott, S. (2013). Rubric. *The glossary of education reform*. Retrieved from <http://edglossary.org/rubric/>.
- Andrade, H. G. (2000). Teaching with rubrics: The good, the bad, and the ugly. *College Teaching*, 53(1), 27-30.
- Beaver, C. & Beaver, S. (2011). The effect of peer-assessment on the attitudes of preservice elementary and middle school teachers about writing and assessing mathematics. *Issues in the Undergraduate Mathematics Preparation of School Teachers*, 5, 1-14.
- Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching Research*, 36(2), 81-109.
- Borg, S. (2006). *Teacher cognition and language education: Research and practice*. London, England: Continuum.
- Brew, A. (2009). *Academic research and researchers*. London, England: McGraw-Hill Education (UK).
- Calderhead, J. (1996). Teachers: beliefs and knowledge. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of Educational Psychology*, (pp. 709–25). New York, NY: Macmillan.
- Canh, L. V. (2014). Great expectations: The TESOL practicum as a professional learning experience. *TESOL Journal*, 5(2), 199-224.
- Cartney, P. (2010). Exploring the use of peer assessment as a vehicle for closing the gap between feedback given and feedback used. *Assessment & Evaluation in Higher Education*, 35(5), 551-564. doi:10.1080/02602931003632381
- Creswell, J. W., Plano Clark, V. L., Gutmann, M. L., & Hanson, W. E. (2003). Advanced mixed methods research designs. *Handbook of mixed methods in social and behavioral research*, (pp. 209-240). Thousand Oaks, CA: Sage.
- Crookes, G. (2003). *A practicum in TESOL: Professional development through teaching practice*. Cambridge, England: Cambridge University Press.
- Echevarria, J. J., Vogt, M. E., & Short, D. (2018). *Making content comprehensible for English learners: The SIOP® model*. White Plains, NY: Pearson Education.
- Ertmer, P. A., Richardson, J. C., Belland, B., Camin, D., Connolly, P., Coulthard, G., & Mong, C. (2007). Using peer assessment to enhance the quality of



- student online postings: An exploratory study. *Journal of Computer-Mediated Communication*, 12(2), 412-433. doi:10.1111/j.1083-6101.2007.00331.x
- Evans, C. (2015). Students' perspectives on the role of peer feedback in supporting learning. *Journal of Cognitive Education and Psychology*, 14(1), 110.
- Falchikov, N., & Goldfinch, J. (2000). Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks. *Review of Educational Research*, 70(3), 287-322.
- Freeman, D., & Richards, J. C. (Eds.). (1996). *Teacher learning in language teaching*. Cambridge, England: Cambridge University Press
- Freeman, D., McBee Orzulak, M. & Morrissey, G. (2009). Assessment in second language teacher education. In A. Burns and J. C. Richards (Eds.), *The Cambridge guide to second language teacher education*, (pp. 77-90). Cambridge, England: Cambridge University Press.
- Gass, S., Mackey, A., & Pica, T. (1998). The role of input and interaction in second language acquisition. Introduction to the special issue. *Modern Language Journal* 82(3), 299-307.
- Glaser, B. G. (1965). The constant comparative method of qualitative analysis. *Social problems*, 12(4), 436-445.
- Grainger, P. R., & Adie, L. (2014). How do preservice teacher education students move from novice to expert assessors?. *Australian Journal of Teacher Education*, 39(7), 89-105.
- Hack, C. (2015). Analytical rubrics in higher education: A repository of empirical data. *British Journal of Educational Technology*, 46(5), 924-927. doi: 10.1111/bjet.12304
- Hafner, J. C., & Hafner, P. M. (2003). Quantitative analysis of the rubric as an assessment tool: An empirical study of student peer-group rating. *International Journal of Science Education*, 25(12), 1509-1528.
- Hasbrouck, J. E. (1997). Mediated peer coaching for training Preservice teachers. *Journal of Special Education*, 31(2).
- Johnson, K. E. (2006). The sociocultural turn and its challenges for second language teacher education. *TESOL Quarterly*, 40(1), 235-257.
- Johnson, K. E. (2009). *Second language teacher education: A sociocultural perspective*. New York, NY: Routledge.
- Johnson, K. E., & Golombek, P. R. (2011). *Research on second language teacher education: A sociocultural perspective on professional development*. New York, NY: Routledge.



- Johnston, B. (2009). Collaborative teacher development. In A. Burns and J. C. Richards (Eds.), *The Cambridge guide to second language teacher education*, (pp. 241-249). Cambridge, England: Cambridge University Press.
- Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, 2(2), 130-144.
doi:10.1016/j.edurev.2007.05.002
- Kilic, G. B., & Cakan, M. (2006). The analysis of the impact of individual weighting factor on individual scores. *Assessment & Evaluation in Higher Education*, 31(6), 639-654.
- Kilic, G. B., & Cakan, M. (2007). Peer assessment of elementary science teaching skills. *Journal of Science Teacher Education*, 18, 91-107.
- Koc, M. (2011). The Views of Prospective Class Teachers about Peer Assessment in Teaching Practice. *Educational Sciences: Theory and Practice*, 11(4), 1979-1989,
- Lasater, C. A. (1994). Observation feedback and analysis of teaching practice: Case studies of early childhood student teachers as peer tutors during a preservice teaching practicum. *Dissertation Abstracts International*, 55(1916), 66-71.
- Lemov, D. (2010). *Teach like a champion: 49 techniques that put students on the path to college* San Francisco, CA: Josey-Bass. .
- Li, L., Liu, X., & Steckelberg, A. L. (2010). Assessor or assessee: How student learning improves by giving and receiving peer assessment. *British Journal of Educational Technology*, 41(3), 525-536. doi:10.1111/j.1467-8535.2009.00968.x
- Lynch, R., McNamara, P. M., & Seery, N. (2012). Promoting deep learning in a teacher education programme through self-and peer-assessment and feedback. *European Journal of Teacher Education*, 35(2), 179-197.
- Marzano, R. (2007). *The art and science of teaching: A comprehensive framework for effective instruction*. Alexandria, VA: ASCD.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: A sourcebook*. Thousand Oaks, CA: Sage Publishers.
- Munby, H., Russell, T., & Martin, A. K. (2001). Teachers' knowledge and how it develops. In V. Richardson (Ed.), *Handbook of research on teaching* (pp. 877-904). Washington, D.C.: American Educational Research Association.



- National Institute of Education (1975). *Panel 6, Teaching as clinical information processing*. National Conference on Studies in Teaching. Washington, D.C.: National Institute of Education. February, 1975. Retrieved from <http://files.eric.ed.gov/fulltext/ED111807.pdf>
- Patton, M. Q. (2002). *Qualitative evaluation and research methods* (3rd ed.). Thousand Oaks, CA: Sage Publications, Inc.
- Pang, M. (2016). Pedagogical reasoning in EFL/ESL teaching: Revisiting the importance of teaching lesson planning in second language teacher education. *TESOL Quarterly*, 50(1), 246-263.
- Ploegh, K., Tillema, H. H., & Segers, M. S. (2009). In search of quality criteria in peer assessment practices. *Studies in Educational Evaluation*, 35(2), 102-109.
- Reddy, Y. M., & Andrade, H. (2010). A review of rubric use in higher education. *Assessment & Evaluation in Higher Education*, 35(4), 435-448. doi: 10.1080/02602930902862859
- Richards, J. C., & Crookes, G. (1988). The practicum in TESOL. *TESOL Quarterly*, 22(1), 9-27.
- Shabani, K., Khatib, M., & Ebadi, S. (2010). Vygotsky's zone of proximal development: Instructional implications and teachers' professional development. *English language teaching*, 3(4), 237-248.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4-14.
- Sluismans, D. M. A., Brand-Gruwel, S., & Van Merriënboer, J. J. G. (2002). Peer assessment training in teacher education: Effects of performance and perceptions. *Assessment & Evaluation in Higher Education*, 27(5), 443-454.
- Struyven, K., Dochy, F., & Janssens, S. (2008). The effects of hands-on experience on students' preferences for assessment methods. *Journal of Teacher Education*, 59(1), 69-88.
- Tudge, J. (1992). Vygotsky, the zone of proximal development, and peer collaboration: Implications for classroom practice. In L. Moll (Ed.), *Vygotsky and education: Instructional implications and applications of sociohistorical psychology* (pp. 155-172). New York, NY: Cambridge University Press
- Vacilotto, S., & Cummings, R. (2007). Peer coaching in TEFL/TESL programmes. *English Language Teaching Journal*, 61(2), 153-160.
- Vygotsky, L. (1978). Interaction between learning and development. *Readings on the development of children*, 23(3), 34-41.



Wilkins, E. A., Shin, E., & Ainsworth, J. (2009). The effects of peer assessment practices with elementary education teacher candidates. *Teacher Education Quarterly*, 36(2), 79-93.

Wollenschläger, M., Hattie, J., Machts, N., Möller, J., & Harms, U. (2016). What makes rubrics effective in teacher-feedback? Transparency of learning goals is not enough. *Contemporary Educational Psychology*, 44-45, 1-11. doi:10.1016/j.cedpsych.2015.11.003



Appendix A

Focus Group Interview Questions

1. Comment on your general experience in the practicum course.
2. Throughout the practicum course you were engaged in providing feedback to your peers. Please comment on your experiences in receiving and providing feedback to your peers.
3. How would you define a rubric? What do you think is the purpose of the rubric? Did you find using a rubric helpful? Do you think you might use a rubric in your future teaching? If so, how or for what purpose?
4. Please comment on your general experience with the guided discussion.
5. Do you think you improved your teaching skills in the practicum course? Why? Why not?
6. Is there anything else you would like me to know about your experiences in the practicum course?

¹ For information about the rubrics used in the current study, please contact the second author, [ma.christison@utah.edu].