The Tri-Perspective Observation Tool (3-POT): A Worthy Addition in a Comprehensive Teacher Observational System

Authors:

Adriana L. Medina
The University of North Carolina at Charlotte

Kaitlyn O. Holshouser
Gardner Webb University

ABSTRACT

The Tri-Perspective Observation Tool (3-POT) was developed with the purpose of lessening the constraints typically associated with rubric-based performance measures. This study examined how the 3-POT as an observation tool can support other observation instruments as part of a holistic teacher observational system prioritizing accountability and growth. Data was collected through interviews and focus groups from Student Teachers, Clinical Teachers, and University Supervisors. Student Teachers described how the 3-POT satisfied their desire for feedback. Supervisors reported how the 3-POT offered a new lens for conducting observations and allowed for transparency. The information captured in the 3-POT contextualized the score on the rubric-based performance measured instrument. The 3-POT captured what transpired, allowed for feedback focused on growth, was applicable in a variety of teaching contexts, and complemented a variety of measurement instruments. The 3-POT's validity, versatility, and variability make it a worthy tool for inclusion in a comprehensive teacher observational system.

Correspondence E-mail: AdrianaLMedina@Charlotte.edu

Observation is a common and critical component of teacher evaluation systems (Ross & Walsh, 2019; Steinberg & Donaldson, 2016; Steinberg & Garrett, 2016). For both pre-service and in-service teachers, observations often serve one of two purposes: accountability or growth (Gabriel, 2018; Wise et al., 1985). While on some occasions, observations are used for the purpose of teacher development, on other occasions, observations aid in making high-stakes decisions, such as whether an in-service teacher's contract will be renewed or whether a student teacher will be recommended for licensure. The student teaching experience is a time for Student Teachers to grow in their teaching practice and develop the agency needed to successfully step into the role of a classroom teacher. The student teaching observation process is a critical component of most teacher education programs.

Classroom observations begin in teacher preparation programs and are typically thought of as necessary supports for Student Teachers as they transition into their roles as practitioners (Jonsson & Panadero, 2017). These observations are often conducted by a Clinical Teacher who is in the classroom, modeling teaching practices in real-time, and a University Supervisor who serves as the liaison between the University and the school system. These two supervisors are not the sole factors influencing the quality of feedback and the degree of support provided to Student Teachers. Caughlan and Jiang (2014) note the observation instrument is also an actor influencing the focus and feedback produced from an observation. They warn observation instruments "are not neutral but reflect the values of the programs that use them through particular (and sometimes contradictory) discourses of teacher learning and student learning" (p. 375). Observation instruments have been critiqued for their lack of depth, rigid criteria, relevant feedback, and user subjectivity (Bell et al., 2015; Cohen & Goldhaber, 2016; Gabriel, 2018). While according to Gabriel (2018) the goals of observations are typically accountability and growth, given rigid criteria and numerical values associated with many observation instruments in student teaching, emphasis is placed on accountability more so than growth. This is concerning as the primary focus of student teaching should be growth. Often the observation instruments focused on accountability are rubric-like and numerical in nature, that is, there is a number or score the Student Teacher receives for the competencies under observation. These instruments often do not allow much space for the observer to take and make notes regarding what was observed, nor are they accompanied by an observation tool to support the score. Without notes regarding what was observed, the numerical feedback from these instruments might seem to be more for accountability purposes than for fostering growth in the Student Teacher's ability to demonstrate proficiency in teacher competencies. Therefore, there exists a need for an observational tool that can complement these rubric and numerical types of observation instruments. There is a need for a tool that captures what transpired during an observation, allows for feedback focusing on growth, is applicable in a variety of teaching contexts, and can complement a variety of measurement instruments.

Literature Review

Rubrics are a primary instrument employed for observations in teacher preparation programs. When the observation instruments have criteria and numeric performance levels serving as a scoring guide to measure the proficiency levels of Student Teachers, these instruments can be classified as rubrics. Rubrics, in and of themselves and

independent of the observers, primarily focus on accountability and not growth. Gabriel (2018) posits that instruments cannot serve the purpose of accountability and growth simultaneously, therefore, one must ask which purpose do rubrics prioritize? Given the quantitative nature of rubrics, the goal is to determine whether Student Teachers meet a certain criterion demonstrating proficiency; therefore, it can be argued their central purpose is one of accountability. Shortcomings of these rubrics fall into the categories of lack of depth, rigid criteria, and relevant feedback.

Lack of Depth

A generic rubric might be used when observing a Student Teacher's classroom management which is a broad category and has several large facets which could take on a multitude of forms. Additionally, a generic classroom management rubric could transcend subject areas and grade levels. Although generic rubrics may not encourage criteria compliance to the same extent as specific instruments, they do not come without critique (Gabriel, 2018; Gabriel & Woulfin, 2015; McAbee, 2016). While highly specific observation instruments may provide narrow definitions of quality teaching, generic rubrics may fail to capture the complexity inherent in teaching, especially if the evaluator is not attuned to the best practices within the disciplines they are observing (McAbee, 2016). Tierney and Simon (2004) point out that although generic rubrics may afford the user more versatility, there exists a "tradeoff" in terms of rater reliability (p. 2). This is due to the fact that generic rubrics do "not contain concrete or task-specific descriptions to guide interpretation" (Tierney & Simon, 2004, p. 2). Whereas task-specific rubrics may compensate to some extent for an untrained evaluator, generic rubrics do not make such affordances.

Rigid Criteria

Currently, teacher evaluation measures "are used to make hiring, promotion, tenure, and dismissal decisions;" therefore, it comes as no surprise that observation instruments exert influence on the type of teaching occurring in schools (Harris et al., 2014, p. 74; Williams & Hebert, 2020). While this may have positive outcomes in some educational settings, McAbee (2016) argues that in some cases "the instrument is creating a situation where the teacher has to sacrifice quality instruction" (p. 179). For Student Teachers, the student teaching observations hold power over candidates' ability to obtain recommendation for licensure. Consequently, Student Teachers may feel pressure to teach to an observation instrument. This criteria compliance (Torrance, 2007), over time, could result in decreased agency (Caughlan & Jiang, 2014).

Cohen and Goldhaber (2016) discuss the fact many "constructs" of good teaching practices have emerged through educational research over time. Observation instruments that are highly specific may neglect other research-based teaching strategies, thus limiting teachers' creative freedom. Observation instruments that highlight one form of teaching over another fail to realize the intricacy of teaching to meet a given context (Connor, 2013; Cohen & Goldhaber, 2016). Therefore, time and attention must be given to the selection of appropriate observation instruments, since they often set the tone for what is valued as good teaching (Caughlan & Jiang, 2014; Connor et al., 2014).

Relevant Feedback

Feedback is two-pronged - what the user of the instrument provides as well as what the Student Teacher receives during debriefing conversations (Moskal, 2000). The nature of the instrument can influence the feedback provided (Caughlan and Jiang, 2014). A well-designed, task-specific rubric can be used for growth if additional feedback is provided beyond the score number. The language of each criteria can be used as a goal for the Student Teacher to attain. However, this is only possible if the feedback from the observation describes and shows what the Student Teacher did so the Student Teachers and the observer/evaluator can have a discussion of how to do things differently next time - not just to attain a higher score, but to work towards being a more proficient teacher. The conversations stemming from observations are just as important as the observation instrument used (Helgevold et al., 2015). Sosibo (2013) conducted a study with pre-service teachers to examine their perspective on the Student Teacher observation process. A common theme emerging from Sosibo's research was Student Teachers' desire for feedback. Student Teachers were not satisfied with "evaluators who merely made checks on the forms" (Sosibo, 2013, p. 150). Student Teachers desired detailed feedback and felt this feedback was necessary as evidence of growth in teaching.

To fill these gaps, the Tri-Perspective Observation Tool (3-POT) was developed with the purpose to lessen the constraints typically associated with rubric-based performance measures. The purpose of this study was to examine how the 3-POT as an observation tool can support a rubric-based performance measurement observation instrument to complement a teacher observational system prioritizing both accountability and growth.

Conceptual Frameworks

Ethnographic Practice

Ethnography is the study of culture and focuses on the observation and analysis of social practices and interactions in order to better understand a culture (Bloom & Green, 2018). Ethnographers examine, through a cultural lens, interactions and events within a particular community environment. A classroom serves as a community, and within the boundary of a classroom there exists a culture. Thus, using an ethnographic perspective as a lens for classroom observations can be helpful in understanding what is occurring in a classroom (Frank, 1999).

One ethnographic practice is making cultural ways visible through description. "Ethnography can be used as a tool by classroom observers to make visible what members are doing and learning in classrooms and to record, analyze, and represent the particular kind of classroom culture that is being created" (Frank, 1999, p. 3). The 3-POT is developed with this in mind (see Appendix A). The "What I Saw" and "What I Heard" columns allow for descriptions of the environment and of interactions which can make transparent the connection between what was seen and heard and what was thought. From the notes in these columns, the observed chain of events can be reconstructed. A related ethnographic practice is withholding quick judgement. These descriptive notes can later be used as evidence for interpretive notes which is also an ethnographic practice. A related ethnographic practice is taking fieldnotes, and the "What I Thought" column allows for observer fieldnotes, both in the moment and soon afterwards.

Validity

Validity is "an integrated evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of inferences and actions based on test scores or other modes of assessment" (Messick, 1989, p. 13). Observations and how they are conducted align with Messick's clarification of the term "score" which is "any means of observing or documenting consistent behavior or attributes" (p. 13). Messick explains "validity is an inductive summary of both the existing evidence for and the potential consequences of score interpretation and use" and that "what is to be validated is...the inferences derived from test scores or other indicators" (p. 13). Thus, for an observation tool to be useful, it should provide "scores" or information that will be used for its intended purpose. Messick posits that key to validity are these five elements: "interpretation, relevance, and the utility of scores, the import or value implications of scores as a basis for action, and the functional worth of scores in terms of social consequences of their use" (p. 13). When deciding upon instruments to use as part of a teacher observational system (Hill et al., 2012), oftentimes data is not provided regarding the validity of the instrument's uses and inferences, that is, its interpretation, relevance, utility of scores, implication of scores for action, and the social consequence of the instrument (Messick, 1989). This study helps to avoid this pitfall for the 3-POT.

Method

Tool Description

The first author based the 3-POT on the research by Frank (1999) on ethnographic classroom observations and the benefits of notetaking, a description of what is being observed, and notemaking, the observer's thinking and interpretation of what is observed. The 3-POT divides notetaking into two columns, What I Heard and What I Saw. The third column, What I Thought, is used for notemaking. During an observation, the observer attempts to capture as much of the talk and action as is feasible by writing it down in the first two columns. As the action is taking place, or afterwards, the observer can record their thoughts, questions, and comments in the third column. The notemaking becomes the basis for the post-observation conference and the notetaking serves as evidence. The notetaking and notemaking inherent in the 3-POT make it an ethnographic tool for classroom observations. Notetaking allows for descriptive notes and notemaking allows for interpretive notes. Thus, the tool and its use allow for an ethnographic perspective. The 3-POT does not yield a numerical score. The observer uses the 3-POT to provide feedback related to what transpired during the observation based on what they saw, heard, and thought. The feedback information is used to guide the post-observation conversation.

Research Questions

To examine how the 3-POT as an observation tool can support a rubric-based performance measurement instrument, evidence was gathered to address the following questions:
- What information is gained through the 3-POT and what is missed?
- How do Student Teachers experience being observed with the 3-POT?
- How is the 3-POT used by observers to assess Student Teacher performance?

Setting and Participants

This study took place at a public urban research university in the southeastern part of the United States. The student teaching experience occurs during the last semester of a teacher candidate's program. The Student Teachers are assigned placement by the Office of Field Experience and their Clinical Teachers are designated by the school administration. The University Supervisors are designated by the Director of the Office of Field Experiences at the University and assigned to a school. The sample came from a pre-established process and therefore would be deemed a convenience sample. The study had a total of 21 participants - three University Supervisors, who each supervised three Student Teachers, who were each supervised by a Clinical Teacher. One University Supervisor, three Student Teacher, and three Clinical Teachers were at a high school. Two elementary schools each had one University Supervisor, three Student Teachers, and three Clinical Teachers. The participants mainly self-identified as females except for one male high school Student Teacher.

Data Collection

There were four stages to the research design. Initially, all of the participants were interviewed individually or in a focus group, depending on their teaching and class schedules. Student Teachers were asked about their expectations of the student teaching experience in reference to the students, classroom management, delivery of content, instruction, and being observed and evaluated.

The supervisors, that is, the Clinical Teachers and University Supervisors, were asked about their evaluation practices, philosophy of teacher education, and their use of the current (rubric-based performance measurement) observation instrument, the Student Teaching Assessment Rubric ([STAR] Jaus et al., 2007) which was comprised of ten standards[a] on which PSTs were observed four times during their student teaching semester. Next, the supervisors were trained on how to use the 3-POT. An explanation of the tool and a demonstration were provided. Then, Clinical Teachers and University Supervisors were asked to conduct one observation using the 3-POT.

Finally, participants were interviewed in focus groups again. The supervisors were asked to describe the barriers and facilitators to using the 3-POT. The semi-structured interview questions were: Did you find it easy or difficult to use the 3-POT? What information did you gain? How did you find the 3-POT as a tool for observing the instructional practices of Student Teachers? The Student Teachers were asked how their expectations matched their experience and about their feelings of being observed with the 3-POT. Specifically, they were asked: What information did you expect to gain from the observations of your teaching? What are your post-observation reactions to being observed with the 3-POT? Was there any difference in the information you received when observed with the 3-POT than when observed with other instruments?

[a] The 10 Standards were: Content Pedagogy, Student Development, Diverse Learners, Instructional Strategies, Motivation and Management, Communication and Technology, Planning, Assessment, Professional Growth, School and Community Relationships

Table 1. *Categories, Themes, Descriptions, and Codes*

Categories	Themes	Description	Codes (Frequency)
Positive Aspects	General Positives	General positive aspects of the 3-POT	Positive Aspects (11)
	Reflection	Note-taking on what was seen and heard allowed for deeper reflection and note-making afterwards	Allows for Reflection (17)
	Transparency	The 3-POT allows for transparency	Transcript (21) Depth of Observation (20) New Lens (6) Visual (6)
Limitations	Tool Limitation	3-POT limitations or limitations arising from observer influence over the 3-POT	Writing is Time Intensive (34) Multiple Interactions (5) Length of Observation (4) Learning a New Method (4)
Uses	Teaching Points	Using the 3-POT feedback for post-observation discussions and as evidence of demonstrated teaching skills and competence	Specific Issues (34) Snapshot (11) Teaching (9)
	Documentation	Using the 3-POT to document concerns	Document issues (11)
	Suggested Modifications	Suggestions for modifications to further the 3-POT's use as an observation tool	Timing (19) Modify Layout (10) Supplemental (7)
	Comparisons	Comparing/contrasting the 3-POT with other instruments	Broad vs Specific (24) Words vs Numbers (15) Filling it Out (8) Helpful/Useful (7) Deficit-Focused vs Asset-Focused (5) Level of Intrusiveness (1)

Data Analysis

Interviews were recorded and transcribed and de-identified of individual information. NVivo (QSR International, 2019) was used for data management and organization during the qualitative analysis process. The authors coded all the data

together using a constant-comparison method (Strauss & Corbin, 1998). Inductive coding was accomplished first through iterative coding of the transcriptions and deductive coding was employed to pull out the emerging themes within the codes (see Table 1). Data saturation was achieved through the transcripts. The supervisors' observation documents were collected and examined. They confirmed codes but did not add new information.

Results

Findings are presented with regards to 1) potential uses of the 3-POT, 2) validity elements, namely, interpretability, relevance, the utility of results, the implications of the results as a basis for action, and the functional worth of the results with regards to the social consequences of the use of the results, 3) positive aspects of the 3-POT, 4) limitations of the tool, and 5) suggested modifications. Direct quotes are italicized and codes are in bold.

Potential Uses

One Clinical Teacher suggested using the 3-POT to point out specific issues, "if a [Student Teacher] was having difficulties" with some aspect of teaching. One Student Teacher suggested the focus on specific issues as well, however with the idea that the Student Teacher could indicate what issue they "would like [the supervisor] to pay attention to" when observing them.

It was suggested the 3-POT could be used to capture a snapshot of the teaching. A University Supervisor said, "I do think it's nice to have a snapshot. I think it would be great to… transcribe 20 minutes of all your classes. By the time you got to the end of the semester, you would have a nice snapshot of what was going on in that teacher's classroom. …It could be 20 minutes from the beginning one time, 20 minutes in the middle the next time, etc." It was also suggested the 3-POT made for a "better teaching tool for [the Student Teacher] to see how she could handle things." Additionally, using the tool as a point from which to teach "could spark a discussion" during the debrief about what went well and what could be done differently next time.

Another suggested use of the 3-POT was as a method to document issues. One Clinical Teacher felt the transparency afforded by the 3-POT could be used for a Student Teacher "who maybe could not take criticism well." She felt the transparency might have the Student Teacher saying, "Oh yeah, [my Clinical Teacher] is not just giving me a difficult time. That actually did happen. I see that now." A University Supervisor indicated she could use the 3-POT observation as data for a Student Teacher who is struggling "to let [him/her] know what issues [he/she is] having."

On the other hand, one Student Teacher felt the observation with the 3-POT could help her justify her actions and show evidence of her being a competent teacher. She gave the example of when her Clinical Teacher wrote: "I re-write the problem after the student reads it and, it's just that I like repeating things." The Student Teacher defended herself by pointing out how "it's a lengthy word problem. I don't think it's a bad thing repeating because when you have a kid talking right here (points next to her), I think it's good that I repeat the problem for other kids to hear it."

Several stakeholders felt the 3-POT would be useful as a supplement to other existing tools. Many suggestions were related to timing. Some supervisors preferred to use the 3-POT early on in the student teaching experience. A University Supervisor agreed it

could serve as a "baseline observation." Other supervisors thought the 3-POT might be useful at mid-term. Another thought it could be used over time to "add to the big picture."

Interpretability

The stakeholders interpreted the 3-POT to be used as intended, for observation. The supervisors indicated this tool reminded them of other observation tools they had used. For example, one University Supervisor said, "...when I was Principal...I used the same format [scripting]." The Student Teachers interpreted the 3-POT to be used for observation. One Student Teacher noted, "on this one, you can see exactly where [the supervisors] saw those things; what you were doing when this happened."

Relevance

The stakeholders indicated the 3-POT was relevant within the student teaching context. The supervisors indicated the tool allowed them to capture the lesson and gave them the opportunity to provide specific, constructive feedback. One Clinical Teacher said, "if I would have done the observation just in my nature, ... I probably would have not written down some things." Using the 3-POT, this Clinical Teacher identified relevant aspects of teaching for her Student Teacher to focus on she might not have found otherwise with another tool. She summarized, "I think there were a couple of small things to work on that probably would have not come out if I was doing it the other way." A University Supervisor felt the 3-POT was relevant in helping her remember more about the observations. "I remembered a lot more about it. ... I had a lot more detailed."

The Student Teachers felt the 3-POT was relevant in providing them growth-oriented feedback. One Student Teacher noted, "... I could really see exactly what [my University Supervisor] was saying when she told me 'You need to do more of this' and I could see how I progressed." Another agreed, "It's not so cut and dry/black and white.... They can elaborate on some things." Another Student Teacher indicated the 3-POT helped the supervisor's comments be more relevant. "I like how you can see what she thought, based on what she heard and she saw how it corresponds with each other in rows." Another agreed, "I felt this only captured the lesson. It didn't go into all the community stuff as the STAR does. You only saw and heard what was in the classroom at that time."

Utility of Results

The stakeholders noted the usefulness of the results of the 3-POT. One Clinical Teacher felt the 3-POT allowed her to "give more anecdotal kind of support than [she] would probably put in a general observation." A University Supervisor agreed. She said, "I think I would focus using this instrument so I can go back and show the student teacher exactly what they were saying, what I was seeing and what the children were doing." Another University Supervisor stressed the 3-POT was useful because of its potential for transparency: "What it was, there it was. This is what I thought, this is what the children were doing, this is what you were doing, this is what was going on in the classroom that I saw. There aren't negatives or positives. It's not negatively or positively stated. It's what the observation was. This is what I saw."

The Student Teachers felt it was helpful to know what the supervisors were seeing and thinking: "it helps to point out [they've] been thinking this when you we were doing this." The Student Teachers felt the 3-POT provided more relevant information than other

tools used in the program. For example, one Student Teacher stated, "[it's] more detailed feedback on some of the things that's not on the InTASC , like, on actual teaching, instead of just concepts, and [it's] actually showing different strategies we used, ... actually how we did it."

Implications for Action

The stakeholders were well aware of the importance of observation results for successful completion of the student teaching experience. One Clinical Teacher reflected, "I would have preferred to use [the 3-POT] for an earlier observation because this is very, very specific as to what I'm seeing and the things that I wrote down that I saw and I thought would be helpful in the next observation." Another agreed, "I think it would have been helpful to have it in the beginning." A third Clinical Teacher noted, "If you do it earlier, maybe on the 2nd or 3rd observation, it gives [the Student Teachers] a chance to use [the information] to try to improve areas where they might need a little more work to do." A University Supervisor thought the 3-POT might have more valuable implications for action for when "you have a struggling student and they don't see it. When you have a student to whom you say 'this is what my impression is of what you are doing' and they go (makes a blank face), it's like they don't get it, they don't see that, they can't step outside of themselves and see how they are perceived." One Clinical Teacher found the 3-POT to be so useful and valuable, she mentioned she'd like to use it as a mentor teacher at her school to observe a new teacher.

Worth and Social Consequence of Use

Student Teachers agreed the 3-POT results were worthwhile because it gave them information "that's in their [Supervisors'] head." One Student Teacher indicated "the most important column for [her] was the one about what [her] teacher thought about the actions [she] was doing." For some Student Teachers, "it's most important just getting that [supervisor's] feedback and the suggestions and the criticisms." One Clinical Teacher noted the worth of the 3-POT was "you can't argue with it, ..., this is what you said and this is what the student did or, this is what you did. This is where you walked; this is where you stood most of the time. It is un-debatable; it's unquestionable."

Positive Aspects

The 3-POT allowed for depth of observation. One University Supervisor indicated that the column, "'What are you seeing?' made [her] more cognizant of looking." A Clinical Teacher indicated "this tool had me analyze some things more deeply. Taking a look at what I heard and then what I saw and then what I thought. I normally do those things ... but probably not to the same depth." The 3-POT served as a new lens for seeing things not seen before. As one Clinical Teacher said, "there were a couple of small things to work on that probably would have not come out if I was [observing] the other way [with the other tools]."

Another positive aspect noted was the 3-POT allowed for transparency. As one Student Teacher remarked, "I don't think it was as subjective. It was there. It was all written down for me." Another positive aspect was the 3-POT functioned as a transcript and provided "word by word by word" what transpired during the observed time. All stakeholders indicated the transcript helped them create a mental picture of the

observation where they could "visualize this is what she said, this is what the kids are doing." While to some it felt the observation "was almost like you took a videotape and you wrote the videotape," to others it felt like "it kind of gives you a transcript of what you did without being on video, because on video you change." Another positive aspect was the 3-POT helped eliminate forgetting. One Clinical Teacher said: "what I found myself doing was trying to get as much of what I heard and saw because I felt that would be parts I would easily forget." This was important because Student Teachers indicated they felt there was a disconnect with the University Supervisor who was not always in the classroom. One Student Teacher said it this way: "when the University Supervisor is observing me, she doesn't see that I do give feedback to each child individually, but my Clinical Teacher sees it. And, so [the University Supervisor] will point out: 'you should have done more individual feedback' and I've done it before in another setting, it's just that in this certain setting that you are observing me on, it's not...I feel like my ratings are lower because she only sees a sip of what I've been doing; she doesn't see my full day or exactly everything I do. So, the University Supervisor was saying she didn't see this, she didn't see that, yet my Clinical Teacher was saying 'I see it.'" The transcript nature of the 3-POT could serve as counterpoint evidence for this type of discrepancy.

The stakeholders noted another positive aspect of the 3-POT – it allowed for reflection. Clinical Teachers indicated the 3-POT gave them a "chance to reflect back on what I heard and saw" and "a second chance to analyze."

Limitations of the Tool

The stakeholders felt the main limitation of the 3-POT was the writing. All of the supervisors wrote their observations by hand. The focus on "transcribing" and "getting everything down" was seen as a limitation. Some Clinical Teachers didn't mind the writing, but "did mind the typing it up part" afterwards. One Clinical Teacher stated, "This is 14 pages of typing. And it's really hard to type Math so, it took forever." Another agreed, "it did take much longer to transcribe and type." This concern is valid as the utility of the 3-POT could be compromised over time if observers reduce what they capture in writing. Capturing multiple interactions such as small groups and "all the chatter and questions of preschoolers" in a Kindergarten classroom was "difficult because there's so much going on." Naturally, it is not expected or possible for one observer to capture all that transpires during an observation; thus, there is selectivity and subjectivity inherent in observations when using any tool. A third limitation noted was this was a new method to learn and they already had methods they preferred. One University Supervisor confessed, "the first one I did I kept having this urge to write down numbers next to each of my comments. ... That's a 2.4."

Suggested Modifications

Several stakeholders suggested modifying the layout of the 3-POT. One stakeholder suggested adding columns for the Student Teachers to weigh in on the observation. One Student Teacher felt this would have been useful for her observation where her Clinical Teacher noted: "You gave such-and-such the eye." During the interview, the Student Teacher countered: "That's what [the Clinical Teacher] saw. [She] didn't hear me say anything to go with it..." The Student Teacher did not feel comfortable confronting

her Clinical Teacher. With an additional column, she felt she could have responded and explained: "Well, they were tapping another kid."

One Clinical Teacher suggested changing the order of the columns and having "what I saw" to be the first column, then "what I heard" and "what I thought." A couple of Clinical Teachers wanted to change the orientation of the page from portrait to landscape. Another suggested adding specific categories of focus, for example, "How did the teacher open the class?" And then, what you heard, what you saw, what you thought about the opening of the class."

Limitations to the Study

It might be seen as a limitation that the supervisors' level of observation experience was not captured. Yet, two of the University Supervisors were part of an author team that developed observation instruments (some of which were used at this institution), and the Clinical Teachers were chosen by the Office of Field Experiences which works with district school administrators in the identification of qualified teachers to serve as Clinical Teachers. Secondly, the use of focus groups might be a limitation with regards to threats of social desirability. While the first author was not involved in the student teaching program, the Student Teachers were aware of her role as a faculty member and this might have influenced their critiques of the system they were currently participating in towards recommendation for licensure. Thirdly, it could also be seen as a limitation that the first author developed the 3-POT and she conducted the interviews. Future studies could employ an independent interviewer; however, the second author joined the study at the analysis stage and was beneficial in counteracting this limitation. Fourth, it should be noted that the instruments and tools influence the activity, so it is possible the supervisors saw differences in them because of what the instruments and tools ask them to do. Future studies might compare similar types of instruments or tools. The 3-POT was compared to a low inference instrument; the findings might be different if the 3-POT were compared to another ethnographic tool. Finally, the fact that the 3-POT was not used with middle school Student Teachers and was only used with pre-service teachers could be seen as a limitation. Future studies should include a more varied sample population.

Discussion

The 3-POT was developed to lessen the constraints typically associated with rubric-based performance measures and complement those rubrics by providing qualitative feedback. The absence of numeric proficiency levels of the 3-POT positions Student Teachers to receive feedback supporting their growth. A number score "tells" while the 3-POT allows for the observer to "show" what transpired during an observation to support the score earned. In essence, the 3-POT is simply the detailed notetaking and notemaking of one's person's observation. Yet, the information captured in the 3-POT is highly valuable because it contextualizes the score on the measurement instrument. There are many observation protocols that allow for observation notes, but it was beyond the scope of this study to identify those or compare them to the 3-POT. However, for scenarios where this is not the case, the 3-POT could be a simple and useful complementary tool. Measurement tools would benefit by being paired with the 3-POT as its validity, versatility, and variability make it a worthy tool for inclusion in a comprehensive teacher observational system.

Validity

The 3-POT was interpreted as an observation tool by all stakeholders. The tool had relevance within the student teaching observation experience with regards to the qualitative nature of the feedback. Unlike a score rubric with a specific criteria, the open-ended and ethnographic nature of the 3-POT allowed for growth-oriented and context-specific feedback which spurred conversation among supervisors and student teachers proving the utility of the 3-POT as an observation tool within the student teaching observation process.. This highlights the use of assessment as and for learning rather than solely using assessment of learning. Supervisors felt the 3-POT had implications for action because they could see themselves using it at various points throughout the student teaching experience and for various purposes within the observational system. The results indicated the 3-POT had worth for student teacher observation. Yet "validation is a continuing process" (Messick, 1989, p. 13); therefore, further research on variations of the 3-POT and its uses in different contexts will be informative. Generally speaking, observation instruments demand research attention because the consequential validity of many observation instruments is not documented in the literature.

Versatility

The 3-POT is versatile enough to stand alone or to be used in conjunction with other methods, instruments, and tools. The 3-POT can be used for observing a specific area in which a teacher needs support (e.g., classroom management). Additionally, the 3-POT is applicable to a variety of observation formats, subject areas, and grade levels. It can be used for remote observations, virtual teaching observations, and as a method of observing teachers on video (e.g., via YouTube) or videos uploaded to a video capture platform (e.g., GoReact). In light of the findings which suggested the 3POT was writing intensive, an appropriate modification could be converting it into a digital format allowing the observer to type thereby enhancing the versatility of the instrument.

The 3-POT could also be used by preservice teachers during field placements throughout their program of study and up to student teaching. Using the 3-POT as a classroom observation tool would allow preservice teachers to become familiar with the 3-POT before they are observed with it. The 3-POT could help preservice teachers understand the field placement classroom culture they are observing and entering.

The 3-POTs versatility has the potential to complement other tools used as part of a teacher observational system to help with teacher growth and accountability. While observational systems are in need of more tools that are growth-oriented rather than focused solely on accountability, it should be noted that the observer is the sieve through which the observation is captured. If the observer is not aware of their biases and does not approach the observation with a growth mindset, then the 3-POT will inadvertently serve the purpose of accountability. Thus reflection is needed when utilizing the 3-POT or any observation tool.

Variety

The 3-POT can be used within an observational system. Given the range of placements Student Teachers may find themselves in, there exists a need for a set of observation instruments for a variety of teaching contexts, without losing specificity as this aids in providing detailed feedback. While the 3-POT is versatile, as with many other

instruments, one tool cannot provide a well-rounded picture of teacher quality. Different instruments should be used for different purposes and in combination with each other in order to create a high-functioning teacher observational system (Hill et al., 2012). In the larger in-service teacher observational system, the 3-POT could be used informally, for peer-teacher observations, and formally within the teacher evaluation process. While the 3-POT was studied within the context of pre-service teacher education, its uses could extend to a variety of other settings.

Appendices

Appendix A: Tri-Perspective Observation Tool (3-POT)

References

Bell, C. A., Qi, Y., Croft, A. J., Leusner, D., McCaffrey, D. F., Gitomer, D. H., & Pianta, R. C. (2015). Improving observational score quality: Challenges in observer thinking. In T. J. Kane, K. A. Kerr, & R. C. Pianta (Eds.), *Designing teacher evaluation systems: New guidance from the measures of effective teaching project* (pp. 50–97). Jossey-Bass. https://doi.org/10.1002/9781119210856.ch3

Bloom, D., & Green, J. L. (2018). Ethnography. In B. B. Frey (Ed.), *The SAGE encyclopedia of educational research, measurement, and evaluation*. SAGE Publications. https://doi.org/10.4135/9781506326139.n239

Caughlan, S., & Jiang, H. (2014). Observation and teacher quality: Critical analysis of observational instruments in preservice teacher performance assessments. *Journal of Teacher Education, 65*(5), 375–388. https://doi.org/10.1177/0022487114541546

Cohen, J., & Goldhaber, D. (2016). Building a more complete understanding of teacher evaluation using classroom observations. *Educational Researcher, 45*(6), 378–387. https://doi.org/10.3102/0013189X16659442

Connor, C. (2013). Commentary on two classroom observation systems: Moving toward a shared understanding of effective teaching. *School Psychology Quarterly, 28*(4), 342–346. https://doi.org/10.1037/spq0000045

Connor, C. M., Spencer, M., Day, S. L., Giuliani, S., Ingebrand, S. W., McLean, L., & Morrison, F. J. (2014). Capturing the complexity: Content, type, and amount of instruction and quality of the classroom learning environment synergistically predict third graders' vocabulary and reading comprehension outcomes. *Journal of Educational Psychology, 106*(3), 762–778. https://doi.org/10.1037/a0035921

Frank, C. (1999). *Ethnographic eyes: A teacher's guide to classroom observation*. Heinemann.

Gabriel, R. (2018). Reframing observation. *The Learning Professional, 39*(4), 46–49. https://learningforward.org/wp-content/uploads/2018/08/reframing-observation.pdf

Gabriel, R. E., & Woulfin, S. (2015). Evaluating the structure and content of observation instruments. In R. Gabriel & R. Allington (Eds.), *Evaluating literacy instruction: Principles and promising practices* (pp. xx–xx). Routledge.

Harris, D. N., Ingle, W. K., & Rutledge, S. A. (2014). How teacher evaluation methods matter for accountability: A comparative analysis of teacher effectiveness ratings by principals and teacher value-added measures. *American Educational Research Journal, 51*(1), 73–112. https://doi.org/10.3102/0002831213517130

Helgevold, N., Naeshiem-Bjørkvik, G., & Østrem, S. (2015). Key focus areas and use of tools in mentoring conversations during internship in initial teacher education. *Teaching and Teacher Education, 49*, 128–137. https://doi.org/10.1016/j.tate.2015.03.005

Hill, H. C., Charalambous, C. Y., & Kraft, M. A. (2012). When rater reliability is not enough: Teacher observation systems and a case for the generalizability study. *Educational Researcher, 41*(2), 56–64. https://doi.org/10.3102/0013189X12437203

Jaus, V. P., Cockman, N. R., Frazier, J. W., Hopper, C. J., & Rebich, S. K. (2007). *Student teaching assessment rubric*. Kendall/Hunt.

Jonsson, A., & Panadero, E. (2017). The use and design of rubrics to support assessment for learning. In D. Carless et al. (Eds.), *Scaling up assessment for learning in higher education* (pp. xx–xx). https://doi.org/10.1007/978-981-10-3045-1_7

McAbee, S. T. (2016). When leadership skills are not enough: The role of principals in high-stakes observations. In R. E. Gabriel & R. L. Allington (Eds.), *Evaluating literacy instruction: Principles and promising practices* (pp. 176–187). Routledge.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 13–103). Macmillan.

Moskal, B. M. (2000). Scoring rubrics: What, when and how? *Practical Assessment, Research, and Evaluation, 7*(3). https://doi.org/10.7275/a5vq-7q66

QSR International Pty Ltd. (2019). *NVivo* (Version 12) [Computer software]. https://www.qsrinternational.com/nvivo-qualitative-data-analysis-software/home

Ross, E., & Walsh, K. (2019). *State of the states 2019: Teacher and principal evaluation policy.* National Council on Teacher Quality.

Sosibo, L. (2013). Views from below: Students' perceptions of teaching practice evaluations and stakeholder roles. *Perspectives in Education, 31*(4), 141–154.

Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). SAGE.

Steinberg, M. P., & Donaldson, M. L. (2016). The new educational accountability: Understanding the landscape of teacher evaluation in the post-NCLB era. *Education Finance and Policy, 11*(3), 340–359. https://doi.org/10.1162/edfp_a_00186

Steinberg, M. P., & Garrett, R. (2016). Classroom composition and measured teacher performance: What do teacher observation scores really measure? *Educational Evaluation and Policy Analysis, 38*(2), 293–317. https://doi.org/10.3102/0162373715616249

Tierney, R., & Simon, M. (2004). What's still wrong with rubrics: Focusing on the consistency of performance criteria across scale levels. *Research & Evaluation, 9*(2). https://doi.org/10.7275/jtvt-wg68

Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria, and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education, 14*(3), 281–294. https://doi.org/10.1080/09695940701591867

Wise, A. E., Darling-Hammond, L., McLaughlin, M. W., & Bernstein, H. T. (1985). Teacher evaluation: A study of effective practices. *The Elementary School Journal, 86*(1), 60–121. https://www.jstor.org/stable/1001217

Williams, K., & Hebert, D. (2020). Teacher evaluation systems: A literature review on issues and impact. *Research Issues in Contemporary Education, 5*(1), 42–50.