Peer Observation of Teaching in EFL Classrooms in Saudi Universities: Challenges and Applications

Najla Alghamdi
Foreign Languages Department
Taif University, Main Campus, Taif, Saudi Arabia
Corresponding Author: n.algamdi@tu.edu.sa

Haifa Al-Nofaie
Foreign Languages Department
Taif University, Main Campus, Taif, Saudi Arabia

Reem Alkhammash
English Department, University College
Taraba, Taif University, Taif, Saudi Arabia

Received:01/11/2023 Accepted:04/25/2023 Published:06/24/2023

Abstract
This paper aims to analyze the kinds of peer observation of teaching in Saudi English language institutes. The data revealed that three models of peer observation of teaching are used: evaluative, developmental and collaborative. It was apparent that the evaluative model is the dominant one. The analysis of the results showed that there is a negative attitude toward Peer Observation of Teaching (POT) among teachers. They do not welcome the idea of being monitored and judged. The authors realized that institutes could enhance the benefits of POT by providing teachers with clear guidelines and engaging them in the construction of institutional policies that welcome other POT models

Keywords: Collaborative Observation; Developmental Observation; Evaluative Observation; Peer Observation of Teaching; Saudi universities.

Introduction

The educational sectors continually seek to improve the quality of teaching and educational output. Peer Observation of Teaching (POT) is the primary instrument that institutions use to achieve their goals for teaching excellence. Nowadays, POT is applied to attain further goals, such as improved learning outcomes, the incorporation of technology in the classroom, arming teachers with the latest practice knowledge in the field, training novice teachers and achieving higher educational standards. POT is sometimes used as an obligatory procedure, if it is for an evaluative or developmental purpose, or as a non-obligatory procedure, when it is conducted for a collaborative purpose. Universities such as Cambridge, the University of Edinburgh, the University of Queensland and University of South Carolina provide pages on their websites dedicated to defining POT, describing its benefits, illustrating the steps involved and providing guidelines on how to use it effectively.

In research, many educators have investigated POT thoroughly to examine its efficacy, its role and to find ways to overcome its disadvantages and emphasize its advantages. Fletcher and Orsmond (2005) pointed out that POT is used in higher education in Britain to develop the reflective practice of lecturers; this is one of the important benefits of POT. Keig and Waggoner (1995) have suggested that POT provides three crucial advantages for teachers in higher education: a) it improves their understanding of the teaching process; b) it increases their understanding of teaching actions; c) it increases the level of collegiality in the workplace. However, other research has found that some teachers show strong resistance to accepting and implementing the changes demanded by POT (Trowler 2002). Evans and Nation (2000) gave an explanation for this resistance; they found that changes tend to overload teachers in university, who are already experiencing internal pressure to teach and publish more.

The focus of this study is on delineating all the recognized types of POT as practiced by English language teachers at Saudi universities. There are still unanswered questions regarding which model seems preferable and why. The issue of POT has begun to receive attention lately in Saudi higher education (e.g., Ali 2012; Shousha 2015; Tawalbeh 2020). The evaluation model is the most dominant model, and it has been characterized as involving power inequality, as will be explained in the literature review. The current studies have mainly focused on teachers’ attitudes and perceptions towards POT and how POT may contribute to their professional development, based on data obtained from questionnaires (e.g., Ali 2012; Shousha 2015; Tawalbeh 2020). Studies that compare the three models of observations at both international and local levels are scarce, a point that forms the significance of this study.

Another significance point is that the researchers have not found studies focusing on interview data which could lead to an in-depth understanding of the issues related to POT. Therefore, this study uses interviews as a main data collection method to provide an in-depth understanding of POT as practiced in language centers. This study answered the following questions:

1. How does POT/observer’s power influence authentic teaching practices?
2. To what extent does POT/observer’s influence limit teaching creativity/innovation?
3. Why do some teachers think that POT is not a helpful strategy for their professional development?
4. Why do other teachers believe in the importance of POT, even though they view it as a stressful strategy?
The first section of this study reviews the relevant literature on teaching observation models in language classrooms. The second section describes the methodology. The results are presented in the third section and there is a discussion in the fourth section. Finally, policy recommendations and the conclusion are given in the last section.

**Literature Review**

The literature discusses three models, namely: evaluative, developmental, and collaborative which are different based on who is doing the observation and its purpose.

**Evaluative**

Evaluative Observation (EO) is a two-phase procedure. In particular, it is designed to fulfill institutional requirements; it has a managerial purpose. In addition, the results of an evaluation can be used to develop teachers’ performance, hence improving the institute’s outcomes. Overall, teacher observation creates a culture of observation that teachers need to become accustomed to. Once teachers are accustomed to such evaluation, it can help them to develop their interpersonal skills and their ability to benefit significantly from observers’ feedback. Sharp (1995) discussed one advantage of observation by pointing out that ‘it provides staff the opportunity to “acclimatize” themselves to observation by external assessors’ (As cited in Martin & Double, 1998, p. 161). Furthermore, some teachers find it a beneficial procedure because it pushes them to work harder, to develop a better understanding of their profession (Keig & Waggoner, 1995).

In spite of certain positive aspects of POT, many teachers have a negative perception of EO, which prevents them from benefiting fully from it. Generally, they do not welcome the idea of being judged, especially if this will threaten their teaching career. Shortland (2004) alluded to the fact that teachers in higher education may tend to avoid being observed. Teachers may accept observation just to obey their institution’s policy, rather than taking it as an opportunity to improve their teaching practice. In the same vein, Martin and Double (1998) mentioned that some of participants are under teaching observation with a sense of obligation enthusiasm. From authors’ experiences, they notice that some teachers refuse EO when it is linked to determining promotions or investigating under-performance. Cosser (1998) pointed out that there is not enough evidence to support the idea that EO can be an effective developmental tool. His study finds it ineffective. In some cases, this model is not more than perfunctory work to fulfill the university or institution requirements. Cosser (2002) added that this model can make teachers feel uncomfortable because they perceive it to be judgmental. Cosser (2002) took an extreme position with regard to the whole process of observation by claiming that no one is qualified to judge the teaching practice of other teachers.

Teaching in universities is different from teaching in Primary or Secondary schools (Berge, 1998). Teachers in universities maintain a strong belief that they are not the only source of knowledge for students; they need to be dependent and passionate about learning. Jensen (2011) noted that in higher education, teachers’ research is more valuable than the ways in which they teach: “at many institutions of higher education, research productivity is weighted significantly more heavily than teaching performance in merit pay evaluations and tenure and rank advancement decisions” (p. 33). Researchers for this paper found that Jensen’s point is in line with some of the comments of their colleagues who believe that they should be spending most of their academic hours on research rather than teaching. In addition, teachers believe that they have enough academic duties that it is unreasonable to subject them to the additional stress of being under teaching observation (Martin & Double, 1998). These attitudes can create a psychological barrier...
towards EO and result in teachers not accepting the need to be observed. Moreover, the authority that an observer has in the classroom can upset or irritate an observee, especially when the observer lacks pedagogical or content knowledge so their contributions are not valuable. Kayaoglu (2012) mentioned this crucial point; the data revealed that some teachers did not get the expected benefit from supervision, because the supervisors lacked the necessary knowledge, supervision skills and pedagogical value. EO can result in teachers becoming frustrated, lower levels of creativity, and a loss of interest in engaging with the process. The collaborative model is presented in the following section, through a discussion of its pros and cons.

**Collaborative**

In Collaborative Observation (CO) which is also called peer review (Siddiqi, Jonas-Dwyer, & Carr 2007), teachers engage in observing each other to reflect on their practices. Power relations can be less obvious in this model (Siddiqui et al., 2007). CO is believed to be a successful observation model because it provides teachers with a space for reflecting on their teaching practices, in a non-evaluative context (Daniels, Pirayoff, & Bessant 2013; Engin & Priest 2014). Kocur’s (2021) framework introduced best practices for CO. The first element is to identify with clarity the goal of the observation. Both the reviewer and reviewee should have a focused goal. The second element of best practice is that the observation notes captured during the observation should be thought about before the post-observation meeting. Thirdly, the interaction between the observer and observee should be active and focused during the following phases of observation: the pre-observation phase in which goals are set up; during the observation in which teaching practices are captured; the post-observation phases where written notes and feedback are discussed. Finally, setting criteria for selecting the observer is extremely important. The level of experience of the observer, rather than necessarily their academic background, should be prioritized. Also, the observer should maintain common knowledge and friendly relationships with colleagues.

However, this mode of observation is questionable if both the observer and observee are friends who want to avoid confrontation and maintain harmony (Schuck, Aubusson, & Buchanan, 2008). Studies on CO have found that cultural values in some Asian countries push people towards maintaining friendly relationships with others by avoiding criticism and face loss (e.g., Nguyen, Terlouw, & Pilot, 2005; Nguyen & Pham, 2020). Other cultural values that affect the quality and credibility of feedback have been identified, such as considering people’s social position and institution (Truong, 2013). In addition, some studies have found that teachers may provide superficial comments on their reflective practices because they want to save face by avoiding self-criticism (e.g., Zhan & Wan, 2016). In the context of Saudi Arabia, there is a lack of professional training program that focus the attention of language teachers on the value of reflective practice (e.g., Sibahi, 2015). Available studies on peer observation in the context of higher education in Saudi Arabia revealed that the most dominant model of peer observation is the evaluative model (e.g., Shousha, 2015; Tawalbeh, 2020). The following section discusses the developmental model.

**Developmental**

One of the earliest works on Developmental Observation (DO) was carried out by Williams (1989). The paper laid out seven important principles that any teacher observation framework, based on a developmental perspective, should implement. The first principle aims to develop teachers’ own judgements about events occurring in their classrooms. The second principle is that
the developmental framework should have limited and focused content that caters for the needs of teachers. The third principle is to link the visits to the course. The fourth principle relates to the need for teachers to take responsibility by discussing the purposes of visits with teachers. The fifth principle is about providing the teacher with an instrument for self-development. The sixth principle is to give positive and helpful feedback to teachers. The seventh principle is that the observer should be flexible in order to be able to respond effectively to the teacher in the post-observation discussion. These seven principles represent the first attempt to remove the stigma linked to the evaluative nature of teacher observation and although it represents a good start, the approach is not well-formulated into frameworks and practices.

Recently, the developmental model has been shaped into an applicable framework in some Australian universities through a framework entitled the Peer Assisted Teaching Scheme (henceforth PATS) (Drew et al. 2016). PATS was created by Angela Carbone and her team and has two main objectives: to enhance teaching practices and to align teachers’ developmental goals with observation. Since the design of the framework focuses on goal-oriented observations, teachers are required to state their developmental goals before the start of the academic term. The period before the observation aims to allow collegial stress-free relations to develop with the observer as well as to identify weaknesses and strengths and set SMART goals. During the academic term, the observer then takes notes during the observation process about whether the stated goals have been realized and how. Observable features of teaching are then noted by the observer using checklists. The design of instruments is based on the developmental nature of observations and therefore, instruments included focus on providing opportunities to give feedback to the instructor using freenote notes. Feedback from the observer can involve providing development ideas, as well as highlighting strength and weakness points. After the term finishes, teachers reflect on how effective their strategies have been in terms of achieving their goals, what they have learned and how they will employ the new knowledge to develop themselves professionally.

Although the PATS framework was designed to give the observer ample opportunity to provide feedback, goal-oriented observations were lacking. Drew et al. (2016) analyzed 42 observation notes in the Australian context and found that only half of them contained developmental ideas and only five of the sets of notes were aligned with the teacher’s developmental goals. It is possible that although developmental models of POT also face less resistance from instructors, it is necessary to pay more attention to observers and how they can optimize their observation skills to serve teachers’ aspirations.

**Peer Observation in the Saudi Context**

Recently, peer observation has gained significant attention in the higher education sector of Saudi Arabia, especially in contexts similar to our study (see Deraney & AlGhamdi, 2020; Alkhatnai, 2021; Deraney, 2022). This has been evidenced by the growing literature on the topic, including works by Ali (2012), Shousha (2015), and Tawalbeh (2020). While peer observation models have been advocated as a valuable tool for evaluating teaching practices, the literature shows that evaluation remains a critical aspect of the process. In some cases, the qualifications of peer observers have been called into question, and the role of management in the process has not been well defined. Ali’s (2012) study, for instance, examined whether there were assessment criteria for the observer and the observed. While peer observation is used as an evaluation tool, it remains unclear whether the observed has a say in selecting the observer and whether the
observation forms are widely distributed or approved by senior administration. Ali adapted a checklist from the TLA Centre at the University of Edinburgh, but it is unclear whether the original form was strictly designed for evaluation purposes. The studies generally focused on teachers’ attitudes and perceptions toward peer observation and its potential impact on their professional development, as evidenced by questionnaire responses. The sole study to consider the benefits of combining evaluative and collaborative observations is Tawalbeh’s (2020), which was limited to a case study of Taif University. Consequently, there remain unanswered questions regarding peer observation practices at Saudi universities. Therefore, this study aims to bridge a gap in the literature by investigating teachers' views regarding teaching observation.

**Method**

**Participants**

The data for this study were collected via email from language centers in Saudi universities in March 2019, using structured interview questions (see Appendix A). As female researchers, convenience sampling was used to choose female participants for this work because they represented a source of data that was easy to access by the researchers (Etikan, 2016). Female language instructors from diverse levels of teaching experience and age groups were targeted. Therefore, the respondents were 35 female English language teachers.

**Instrument**

There were 11 questions in the interview. The researchers designed the questions based on their initial conversations with some language teachers regarding their views on POT, prior to conducting this study. The questions were open-ended, which reflected the variation of the participants' social and educational backgrounds and their levels of experience.

**Coding**

Cohen’s kappa calculation was implemented to identify the levels of agreement and disagreement between the two raters (Bujang & Baharum, 2017). We divided the interview questions into three main focus areas (see Appendix B). The first focus area was descriptive interview questions which focused on letting teachers elaborate on their teacher observation experiences, namely in terms of what kind of POT model was adopted in their institution and the feedback provided by the observer. In deciding which teacher observation model each participant had experienced, we asked teachers about the relationship between the observers and observed as it is the main indicator of the overall model used in different institutions. The other question related to whether the teacher had ever learned anything from observing others. If the teacher answered no, it meant they had never observed others and they had only experienced the evaluative model.

The second focus area in the interview questions was the outcome of the observation experience which involved reflective questions about altered methods of teaching as a consequence of the observation and another two questions about ‘teacher’s cognition’ (Borg forthcoming) – lessons learned and encountering different understandings of teaching. Regarding coding for evidence of professional development, the researchers decided that the teacher should demonstrate examples of modifying their practice in the classroom for something to be counted as evidence of professional development. In coding the difference between which teaching observation model, relationship, did she learn anything from observing others? The third focus area was on teachers’ willingness to be observed and attitude.


Inter-rater Reliability

The data were coded by two researchers and the third one revised both sets of codes to check the congruence between them. Following the coding step, two authors compared their coding for 50% of the data and the inter-rater reliability check was applied. The total number of coded items was 135. The total number of items in disagreement was 10%. Thus, the percentage of agreement = 92.59%. The percentage of disagreement = 13.50%. In order to contribute to research originality, some of the participants’ responses are quoted in the results section.

Results

The participants’ answers of the two of the questions assisted the authors in identifying which type of observation they had been exposed to, the nature of the relationship (e.g., staff member, head of dep…etc.) between the participant and the observer and, whether participants learned anything from observing others’ classrooms. If so, mention them. The data show that the three observation models were used: evaluative, collaborative, and developmental. However, the majority of experiences could be classified under the first model followed by the collaborative model and only five participants had experienced the developmental model.

The Evaluative Model

Looking at the responses of participants who were exposed to the evaluative model, the data reveal critical psychological and pedagogical-managerial issues.

Psychological Issues

Saving face was one of the psychological and cultural issues. Teachers in higher education believe that they are qualified and professional enough that they should not be observed and judged; they interpret observation as an insulting and disrespectful procedure. They expressed their negative sentiments towards the evaluative model. For instance, Participant 14 said, ‘A 26-year-experience teacher who is might brimful with methods of teaching, may feel it is a wasting of time! This works better in schools with new teacher NOT with teachers in higher education’; she continued, confirming her opinion, by saying, ‘Truthfully, observation for teachers in higher education is a kind of insult.’ Participant 8 responded with displeasure, stating, ‘I feel it is teachers’ insult to be tested in front of students.’ She also said, ‘a teacher is a respectable Dictation he was person, and no one has the right to degrade or insult or harass him/her.’

A part of face-saving is being defensive, showing a positive self-image and complaining. For instance, Participant 26 was very defensive in her response. She was asked ‘Did observation increase or decrease your confidence?’ and her reply was, ‘I am professional before and after the observation.’ Being defensive is a very natural reaction, especially among experienced teachers, and they can even become hostile to suggestions (Cosh, 2002).

Other participants preferred to be positive and avoided mentioning negative points. Participant seven stated, ‘They appreciated my teaching style; It was up to the mark Alhamdulilah.’ With the same attitude, Participant 19 replied, ‘I only got positive feedback.’ She also commented on the post-observation discussion by saying, ‘The discussion was about the good and negatives if any.’

Other responses were loaded with complaints. Participant eight said, ‘Sometimes observers are not fair enough’ and also ‘observation is a killing device for education department.’ Another
point related to participants’ anxiety when they are observed. Participant eight generally commented on the observation process by saying, ‘It makes me stressed’; ‘feel uncomfortable.’

**Managerial and Pedagogical Issues**

The EO exists because of a departmental and institutional requirement. The observers usually play an authoritative role in which they have the power to evaluate, judge and act for or against the observees (teachers).

The data reveals the teachers’ thoughts about EO. Participant three admitted that she changed her teaching style during observation, when she was asked, ‘Did you alter your style/way of teaching because of the observer’s presence? Explain in both cases!’ She said, ‘because I want to do my best.’ In the same vein, Participant 19 changed her teaching in front of the observer; she described the observation as pushing her to do better: ‘Observation pushes you to raise your standards.’

On the other hand, Participant 35 divulged that she altered her teaching practice because of the pressure that EO creates, ‘because I was a bit anxious on being observed by the evaluation committee. It’s very normal that in front of others we always want to show our good face.’

**The Collaborative Model**

This model was not reported as a common observation type in Saudi Arabia. A few participants had experienced this model. This model seems to be more favourable than the evaluative model due to a number of advantages; teachers who were observed by colleagues in a non-evaluative context seem to develop the confidence to share their teaching practices. For example, Participant 27 said, ‘I strongly support the idea of sharing the way different teachers teach. It helps us boost our teaching style.’

Another participant said, ‘Yes, this is a very good way to exchange ideas and techniques to learn and we can be notified where all we have to develop ourselves in teaching. Observations should be done frequently.’ This view goes hand in hand with Siddiqui, Jonas-Dwyer, and Carr’s view (2007) that power does not dominate this model of observation. However, some teachers find CO stressful and being observed by colleagues might not always be positive. ‘I was anxious to get the feedback. I focused in our discussion about my teaching (the positive and negative remarks)’ (Participant 33). This means that feedback that emerges from CO is not always positive, a finding that differs from other findings that claim the influence of relationships on the credibility of CO feedback (e.g., Nguyen, Terlouw, and Pilot 2005; Nguyen and Pham 2020). Another interesting result was that teachers might change their teaching style in CO, and this is not related to the presence of the observer. It may be in order to try out certain creative practices: ‘I try to create something new and practical’ (Participant 21). It seems that this participant is enthusiastic about trying her new teaching practices for self-development. This result seems to differ from the results found in one other study which says that teachers avoid self-criticism by avoiding reflection (e.g., Zhan & Wan 2016).

**The Developmental Model**

In general, the developmental model was less practiced than both the evaluative and the collaborative model; only five participants reported experiencing observation which was developmental in nature. It is noted that in all of the cases reported, the circle of development was incomplete; all participants only reported receiving feedback that made them think about
knowledge that was new to them, whereas none of them touched upon whether their goals for the observation were aligned with DO and whether the objectives of the institutions were the main purpose of the observation. It is also worth mentioning that the developmental model was only talked about directly by one of the participants; the researchers inferred that the model in question was developmental with other participants, based on the narratives of the participants who reported the ‘good’ feedback they received, including ‘the suggestions given by the observer develop my teaching and I want to improve myself’ (Participant 16). It is also noticeable from the previous quote that such feedback affected the careers of teachers by raising their professional standards.

When the observation was developmental, one participant noted that their confidence increased: ‘It increased my confidence; I saw it as an opportunity to assess myself and to validate what I do in the classroom. It was a way to expel any doubt of myself and that of my employer’ (Participant 12) and ‘My confidence was boosted because this gave me invaluable insight into how I can improve my teaching’ (Participant 28). Part of the confidence gained related to the experience of DO in terms of how it increased some teachers’ awareness of many aspects of lesson preparation and delivery – some focused on their lesson plan: ‘My lesson plan was enhanced to provide the observer with clear teaching objectives’ (Participant 35) and others used the experience to verify their teaching techniques, for example by using humor or developing a good rapport with the students: ‘I was told I have a very good relationship with my students’ and ‘a great sense of humor that’s unique to me’ (Participant 35).

Another expected result, associated with the merits of DO, is that other participants believe that the model provided them with new learning opportunities as in the case of learning a new teaching strategy for one of the participants: ‘I have learned to involve my students more in pair or group work where I can make groups of students on my own’ (Participant 13) and another participant learned to focus attention on low performing students: ‘I should pay attention to some weak students’ (Participant 16). Although the developmental model did not go beyond providing feedback, some participants found that the feedback given to them had a strong impact on them.

Discussion

The results in the previous section reveal that all three types of observation exist in the context of Saudi Arabian higher education; however, they are practiced with different degrees of power. In addition, the attitudes of teachers towards these models varies, depending on the type of observation model, level of teaching experience, training, and the degree of observer and institution power.

The first research question explores how POT/observer’s power influences the authentic teaching practice. The results show that this influence can be positive or negative depending on the observation model that teachers are exposed to. Mainly, the influence is positive when the teachers are observed for developmental purpose and involved in CO and DO models. Two common comments provided by some teachers are that they are learning more about the teaching process via POT and that their teaching has been enhanced through sharing different ideas and strategies. This goes along with Donnlley’s (2007) participants’ (apprentice teachers) comments who admitted that POT was a great opportunity for them to learn about how to teach (p. 124). The negative influence appears in EO context where teachers claim that they are not themselves in the classroom, and they alter their teaching style just to satisfy the observer. The first research question explores how POT/observer’s power influences the authentic teaching practice. The results show that this influence can be positive or negative depending on the observation model that teachers
are exposed to. Mainly, the influence is positive when the teachers are observed for developmental purpose and involved in CO and DO models. Two common comments provided by some teachers are that they are learning more about the teaching process via POT and that their teaching has been enhanced through sharing different ideas and strategies. This goes along with Donnlley’s (2007) participants’ (apprentice teachers) comments who admitted that POT was a great opportunity for them to learn about how to teach (p. 124). The negative influence appears in EO context where teachers claim that they are not themselves in the classroom, and they alter their teaching style just to satisfy the observer.

Another important point this research is to what extent does the POT procedure limit creativity and innovation in teaching practices. Again, the results vary here based on the POT model that the institute uses with its teachers. Teachers who have been observed for evaluative purpose, find this model a stressful one which frustrates them and stops them from thinking out of the box. This result is consistent with Millis’ (1992) finding that teachers experience a distinctive nervousness during observation, while teachers who have been involved in CO say that this type of observation is a chance for them to try creative practice.

In the third research question, the researchers attempt to uncover some of the reasons why teachers think that POT is not a helpful process to develop their profession. The teachers’ answers for this question reveal a very negative attitude towards the EO and DO models, while other teachers welcome the idea of being observed within the framework of the CO model. Regarding the last research question which investigates teachers’ belief of the importance of POT, only teachers who have been involved in CO find POT an important tool to develop their profession. One teacher even said that POT should be done frequently. Below, these results are discussed thoroughly based on their occurrence in the three observation contexts.

The results indicate that although the evaluative model is the dominant one in Saudi universities, it is not the preferred model among teachers, as shown by data in the current research. In the evaluative model, the results, in general, reveal that this model is not welcomed by university teachers. This, indeed, is due to a variety of reasons. The main reason being that teachers in English language centers come from different cultural, social, and educational backgrounds. This diversity creates a competitive atmosphere where teachers work very hard to prove themselves. Teachers in this institutional environment have their own worries about promotions and sometimes about contract renewals. Therefore, the idea of being judged and criticized is loaded with negative sentiment. Additionally, being a teacher at university is another crucial reason to not accept being observed for evaluative purposes. University-level teachers believe that they have a high content knowledge which, in turn, translates into a great sense of self-efficacy. Hence, there is an argument that they should not be assessed under EO since they are above this insulting procedure (as the teachers stated in the data).

Based on the above, the concept of face-saving is raised in the data. In such an environment, face denotes honor, respect, and dignity (Nguyen & Pham, 2020). Certain responses to the interview questions by participants were face-saving in manner; they gave the impression that EO made them feel the need to protect their dignity and respect. This manifested in different ways: some participants complained, others were defensive, and some adopted a deliberately positive attitude.

Another important reason for not welcoming EO is the pressure that this model puts on teachers. It causes them to feel stressed, which in turn, limits their practice and decreases their readiness to think out of the box and be creative (Gosling, 2002, as cited in Peel, 2005).
Regardless of the frequent use of the evaluative model, the outcomes of this managerial procedure are not always reliable. This is because of the authority that the observers have during the observation. Some of the teachers in the current study admitted that they are not themselves during an observation. In fact, some stated that they alter their teaching style just to satisfy the observers and fulfill their institution’s requirements and standards (Peel, 2005).

The results also demonstrated that teachers who had experienced this model did not undertake any reflection, which constitutes a very important step in improving the teaching profession. Wlodarsky (2020) concludes that reflection helps individuals and groups to change and improve their practice and to avoid repeating past mistakes. If a teacher is unable to reflect as a result of an observational model, there is a weakness in the process for two reasons. Firstly, this model is managerially judgmental and thus, teachers are not interested in engaging with it except to simply fulfill their perfunctory duty. Secondly, teachers are not well trained to professionally reflect on their teaching. For them, reflection simply involves talking about what they have learned from an experience, but they do not discuss the implementation and application of improved practice.

The collaborative model seems to be preferred by those teachers who have had experience of this type of observation. This model of observation is stress-free (Daniels, Pirayoff, & Bessant 2013; Engin, & Priest, 2014), but it is not commonly used at Saudi universities, where EO seems to be dominant, as reported by the participants in this study. One main concern, as seen in the literature, is that this type of observation might be affected by institutional values (e.g., complements due to the status of the observees). If teachers are well-trained on how to reflect on their teaching practices, the influence of such factors will be minimized (Sibahi, 2015). It should be noted that friendly relationships might not represent a barrier to reliable feedback from observers. One of the participants reported that CO provided her with a space for using her creativity by altering her normal teaching style: ‘I’m implementing new ways to integrate technology in the classroom, new strategies to increase interaction and collaborative learning’ (Participant 31). A similar finding has been reported in Bell’s (2002) study where some teachers said that after POT, they make immediate changes to their teaching practice; they even improve their implementation of teaching and learning activities (p. 33).

The developmental model seems to be at an early stage of implementation in the Saudi context. The experience is still new to both the institutions and the teachers; however, it seems that many teachers prefer the DO that provides them with feedback that is centered around their development. This result confirms the existing evidence in Bell and Mladenovic (2008), where they found that teachers value the process of peer observation when it is done for a university developmental program. Many of the participants in Donnelly’s (2007) study reported that POT increases their confidence and self-assurance which is a finding that has been found in this study under the DO model. The strong dislike of the EO and consequently the increasing popularity of the DO are due to teachers’ trust in the institution that the scrutiny and judgment that some teachers experience will not be directed toward their evaluation and possible de-recruitment from the institution. This result is in line with studies that adopt the developmental model since judgment cannot be avoided sometimes. The model can be employed when prizes for teaching are announced or promotions are given in a higher education context (Drew et al., 2016).

Both the evaluative and the developmental models are inherently judgmental even though they serve different purposes. The attested power differential between the observer and the observed is noted in Gosling (2002). The findings showed that some language instructors who
preferred the developmental model are happier to be evaluated when the outcome of the evaluation is geared towards their professional growth. In addition, judgment is not welcomed when it is for evaluative purposes. This shows that language instructors in Saudi universities do not mind being evaluated or judged but mind very much how they are being evaluated and whether the purpose of any evaluation is in line with their professional agenda.

**Suggestions and Implications for Policy**

POT is helpful for developing teaching practices; however, institutional authorities need to rethink of the implementation of EO. There should be a place for teaching innovation. Institutions need to become flexible about accepting and including new practices and not treating teaching as a fixed or prescribed process (it is not a medical procedure. Teaching is affected by various factors (context, learners, policies, etc). Institutions need to play down their power and adapt the double-learning loops (Argyris, 1976). This means that teaching practices should inform educational policy and the challenge to change the institutions towards this move.

One framework for classroom observation could involve a multi-staged purposeful activity in language institutes. We propose that any teaching observation framework should include three stages: the first stage should be informative; the second stage should be interventionist; and the third should be evaluative. The informative stage should involve getting to know the teacher, including their teaching aspirations, their goals and their ideals. The interventionist stage should involve linking what is observable on the ground with teachers’ aspirations, goals and ideals and giving advice on how to achieve them if they are not realized. The third stage then opens a door for evaluation of teachers’ performance.

Any institutional observation cycle should start with several COs where teachers can reflect on teaching practices and align these practices with institutional policies. This allows an assessment to be made of whether new implemented practices meet the institutions’ policies and allows for changes in practice or policy to be suggested. Shared discussions can be translated into goals used to identify criteria for informing evaluative checklists. Another value of the constant CO is preparing teachers for the EO and thus reducing the levels of stress that are experienced by teachers during EO.

The most popular alignment to date in Saudi higher education relates to learning outcomes and assessment. However, it is missing from many institutional efforts to focus on another equally important alignment between learning outcomes and teaching activities, that aims to achieve those outcomes. As Biggs (2014) has pointed out teaching should be “designed to engage students in learning activities that optimize their chances of achieving those outcomes, and assessment tasks are designed to enable clear judgments as to how well those outcomes have been attained” (pp. 5-6). Evidence in this study indicates that institutions need to support teachers with teaching to better the experiences of students, in line with Biggs (2014) who has stressed that policies and procedures need to be supportive. The institutions need to support the observer and the observee by providing a clear alignment between learning outcomes and teaching activities to make the whole POT process more systematic and productive in a way which would allow the institution to measure the POT experience at the end of the academic year. In other words, improvements in the teaching process should be reflected in students’ cognitive and linguistic abilities and skills. If these improvements are achieved then teachers will recognize the importance of POT and understand that POT is not just a matter of institutional procedure, rather, it is a tool to improve both teachers as professionals and productions. Higher education institutes will be able to transition from quality
assurance (alignment between LO and assessment) to quality enhancement (alignment between LO and teaching) if such improvements are put in place. A notable example is the UTAS assessment policy created by the Tasmanian Institute for Learning and Teaching mentioned by Biggs (2014) as one of the pioneering institutions supporting teaching.

- Purpose: using this as a guidance for both
- Observer and observee justified feedback based on the alignment

Conclusion

This study aims to understand, in depth, observation experiences as practiced at many language institutes in Saudi universities. As language instructors are one of the most important pillars in the process, data were gathered from them through interviews. The qualitative analysis indicated that most language instructors had experienced the evaluative model of observation rather than any other model, including the collaborative or the developmental models. Most teachers agreed with the shortcomings normally associated with the evaluative type of observation and they seemed to be eager to switch to models that are centered around personal and professional growth. One of the most prominent implications of this study is that frameworks for POT models need to be accompanied by a willingness in institutes to welcome other models of observation and initiate forward-thinking and pro-growth policies, procedures and guidelines that support language instructors and set clear expectations of the desired outcomes of POT.

About the authors

**Dr. Najla Alghamdi**, is an assistant professor of sociolinguistics at Taif university at the Department of Foreign Languages. Najla is interested in dialect variation and change, and she published number of articles about Saudi dialects. She is intersected in language teaching as well. Najla is a member in a linguistic research group at the University of Essex, UK, and a member of the scientific committee at Forum for Arabic Linguistics, UK. ORCID: https://orcid.org/0000-0003-3451-0167

**Dr. Haifa Al-Nofaie**, is an associate professor at the Department of Foreign Languages at Taif University, Saudi Arabia. She holds a Ph.D. in Applied Linguistics from Newcastle University, UK Her research interests include bilingualism, CALL and language pedagogies. She was a series cultural advisor to ELT Oxford University Press between 2014 and 2015. She has published some research studies in international journals. ORCID: https://orcid.org/0000-0001-8984-8777

**Dr. Reem Alkhammash**, is an associate professor of linguistics at Taif University. She also the Director of the English Language Centre. She obtained her Ph.D. from Queen Mary University of London in Linguistics. Her work published in reputable journals such as Discourse & Communication, Journal of Psycholinguistic Research, Scientific Studies in Reading and other international peer-reviewed journals. ORCID: https://orcid.org/0000-0003-1663-5185

References


http://doi.org/10.1177/105960117600100312


https://www.westga.edu/~distance/Berge12.html


http://doi.org/10.2427/12267

http://doi.org/10.1080/1355800980350211

http://doi.org/10.1023/A:1003155703281

http://doi.org/10.13189/ujer.2013.010318


http://doi.org/10.1080/02602938.2016.1209733

http://doi.org/10.14297/jpaap.v2i2.90

http://doi.org/10.11648/j.ajtas.20160501.11


