The Journal of Language Teaching and Learning™

| 2025 | Volume 15/Issue 2 | Article 1 |

Do Professional Development Activities Work Effectively? Insights from English Instructors

Ertan Yazıcı, Hacettepe University, Ankara, Türkiye, ertanyazici@hacettepe.edu.tr
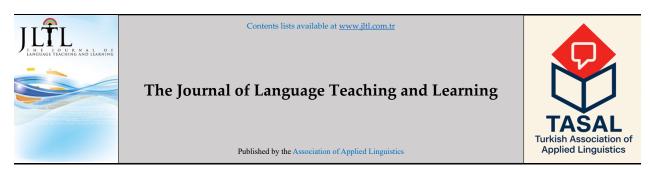İskender Hakkı Sarıgöz, Gazi University, Ankara, Türkiye, iskender@gazi.edu.tr

Do Professional Development Activities Work Effectively? Insights from English Instructors[1]

Ertan Yazıcı[2], İskender Hakkı Sarıgöz[3]

ARTICLE INFO

ABSTRACT

This paper presents a case study examining the perceptions of English instructors (*N*=91) towards the effectiveness of clinical supervision and reflective teaching (peer observation, team teaching and video coaching) at the school of foreign languages of a foundation university in Türkiye. Specifically, the study aims to compare the perceived effectiveness of these professional development activities and to gather suggestions for potential improvements. Utilizing an explanatory sequential mixed-methods design, a survey and two sets of semi-structured interviews were employed to collect the data. The results reveal that clinical supervision is generally perceived as more effective than reflective teaching, while video coaching emerges as the most highly valued reflective teaching modality. These findings suggest that language-teaching institutions can increase the return on professional development activities by prioritizing video coaching and by streamlining clinical supervision schedules to reduce redundancy. Furthermore, the research uncovers several common themes related to the perceived effectiveness of the activities, as well as actionable suggestions for enhancing their impact.

English teachers instruct students from various generations and should continually strive for professional growth. Since each generation comes with its unique demands, and regardless of their teaching experience, English teachers should make efforts to address these evolving needs. Ur (1996) categorizes teachers with 20 years of experience into two groups: those with genuine 20 years of experience and those with 20 repetitions of one year of experience. Early in their careers, most English teachers recognize the importance of professional development (PD), driven either by their institutions or their own personal desire to excel.

Darling-Hammond and Richardson (2009) stress the importance of equipping teachers with a diverse skill set to prepare students for the 21st century difficulties. With the globalization of English, technological advancements and growing demand for language education, countries and educational institutions must provide advanced language training. Therefore, engaging in professional development activities (PDAs) becomes crucial in achieving this goal. Peyton (1997) underscores the significance of PD for English teachers due to rapidly changing student demographics, national education reforms and the establishment of national standards for foreign language learning. Thus, English teachers may need to participate in various PDAs to address teaching challenges effectively. In Türkiye, all universities with schools of foreign languages must demonstrate instructor development in accreditation reviews by Council of Higher Education Quality. However, policy documents provide only broad descriptors, leaving each institution to decide which PD formats to prioritize. Foundation universities often rely on ad-hoc mixes of clinical supervision and reflective-teaching activities without empirical evidence of their comparative value.

The foundation of this study rests on the concept of teacher professional learning as articulated by Opfer and Pedder (2011), which conceptualizes professional growth as a dynamic, iterative cycle that interweaves collaborative inquiry, sustained engagement and evidence-based reflection. Within this ecology, PDAs cease to be one-off events; they become structurally embedded routines that progressively reshape teachers' beliefs, classroom practices and local cultures. To specify how this learning cycle is enacted in schools, the study adopts DuFour's (2006) professional learning community lens, which repositions teachers from isolated practitioners to co-constructors of shared expertise. Professional learning communities transform individual insight into collective competence by providing stable arenas for dialogue, joint problem-solving and peer feedback. Finally, McLaughlin and Talbert's (2006) communities of practice perspective clarifies the mechanism that sustains change over time; repeated participation in a joint enterprise nurtures self-efficacy, a sense of belonging and a commitment to continuous improvement. Taken together, teacher professional learning, professional learning community and communities of practice triad supplies an integrated, multilevel framework that links the micro-processes of individual learning to the meso-structures of departmental collaboration and, ultimately, to macro-level gains in pedagogical quality.

Teacher PD is a vital research area in English language teaching, with extensive literature covering various PD topics. Recent research has focused on clinical supervision (CS) (Amini & Gholami, 2018; Baltaci-Goktalay et al., 2014; Caughlan & Jiang, 2014; Moradi et al., 2014; Paba, 2017). Researchers have also recently investigated reflective teaching (RT) practices. Peer observation (PO) has been extensively studied (Ahmed et al., 2018; Bailey & Van Harken, 2014; Chien, 2019; Copland, 2010; Dos Santos, 2016; Dos Santos, 2017; Gonen, 2016; Hamilton, 2013; Hendry et al., 2014; Kim & Silver, 2016; Lasagabaster & Sierra, 2011; Motallebzadeh et al., 2017; Tezcan-Unal, 2018; Tosriadi et al., 2018; Tsoulou, 2016; Yürekli, 2013). Team teaching (TT) has received some attention (Barahona, 2017; Igawa, 2009; Ng, 2015; Rao & Chen, 2020; Simons et al., 2020; Thompson & Schademan, 2019; Tim, 2018), while video coaching (VC) has been relatively underexplored (Masats & Dooly, 2011; Wong & Pow, 2012; Yücel et al., 2016). Despite the importance of teacher PD, there is a dearth of data on English instructors' opinions regarding the effectiveness of these PDAs. Hence,

research on how English teachers perceive the value of clinical supervision (CS) and three modalities of RT (PO, TT and VC) might be considered warranted. Moreover, although several studies have examined CS and RT separately, none have explored them concurrently. Crucially, no published study has investigated (a) how in-service English teachers judge the relative effectiveness of CS versus three RT modalities within the same context or (b) how those judgments translate into concrete recommendations for a school of foreign languages. Existing Turkish studies tend to evaluate a single PDA in isolation. As a result, administrators lack comparative data to decide which PDAs deserve scarce resources. The present study addresses this two-fold gap by providing the first side-by-side evaluation of CS, PO, TT and VC among 91 English instructors at a foundation university.

To address this gap, the present study simultaneously examines English instructors' perceptions towards CS and three RT modalities (PO, TT and VC) within an explanatory-sequential mixed-methods design. By combining survey data and in-depth interviews from 91 instructors in one of Türkiye's largest tertiary language schools, it offers the first head-to-head comparison of these activities in a Turkish higher-education context. This mixed-method, multi-activity focus extends prior work that treated each practice in isolation and provides practice-oriented evidence for aligning PDAs with English instructors' demonstrated preferences and perceived impact.

At tertiary level institutions, professional development units (PDUs) play a crucial role in assisting English instructors as they navigate new teaching environments. Recognizing the universal consensus among academics on the significance and attributes of PD, universities have established PDUs to support instructors in their continuous improvement efforts. It is noteworthy that the primary focus of PDUs often revolves around CS and RT practices. Despite these two PDAs being initially implemented for the benefit of English instructors, some English instructors at the school of foreign languages where this study was

conducted expressed the need for their enhancement.

Since CS and RT (PO, TT and VC) were the two PDAs practiced most frequently in in-service teacher training at the PDU of the research context of this study, it was intended to investigate how the English instructors felt about their effectiveness. Additionally, the study aimed to compare instructors' perceptions towards the effectiveness of CS and RT, and the effectiveness of PO, TT, and VC, and if necessary, provide recommendations for refining current practices. By mapping perceived effectiveness onto concrete improvement suggestions, the study offers actionable evidence for PDUs across Turkish higher education and extends international PD literature with data from an under-reported EFL context. To guide the study, the following research questions were formulated:

1. Which type of the PDA (CS or RT) do the English instructors find more effective?

2. Which type of RT (PO, TT or VC) do the English instructors find more effective?

3. What do the English instructors think of the effectiveness of the PDAs?

4. What are the suggestions of the English instructors for the improvement of the effectiveness of the PDAs, if any?

2. Literature Review

In its broadest sense, PD refers to an individual's progression in their professional career (Villegas-Reimers, 2003). Teacher PD, as defined by Glatthorn (1995), pertains to the professional growth that teachers achieve through increased knowledge and continuous reflection on their teaching practices. This suggests that teachers are in a perpetual state of PD. There are two main forms of teacher PD known as extracted and embedded. According to Flint et al. (2011), extracted PD prioritizes external experts. In this form, teachers receive training from experts who may not be familiar with the specific context of the educational setting (Fiszer, 2004). Conversely, embedded PD emphasizes learning from instructors within institutions (Desimone, 2011; Opfer & Pedder, 2011). Borko (2004) underscores the significance of PD, as it equips

teachers with the tools to navigate evolving educational standards, a task that can be demanding. This navigational aspect is facilitated through dedicated PD activities.

Richards and Farrell (2005) categorize PDAs into four distinct classifications that reflect the inherent structure of these endeavors. These classifications are individual-oriented pursuits, collaborative undertakings with peers, cooperative initiatives involving multiple colleagues and institutionally organized initiatives. Individual PDAs may include self-monitoring, maintaining reflective journals, reflecting on critical incidents and curating teaching portfolios. On the other hand, collaborative ones involve practices such as CS, PO, TT and VC. PDAs involving multiple colleagues encompass action research endeavors and the formation of teacher support groups. Workshops and conferences organized by educational institutions are examples of institutionally organized PDAs.

The concept CS, introduced by Goldhammer (1969), is a pedagogical framework that emphasizes data collection during observational activities. This formative approach to observation, involving pre-observation planning, the observation itself and subsequent feedback, aims to enhance teachers' pedagogical skills (Paba, 2017). CS is a collaborative process in which both the observer and the observee actively participate. Unlike other forms of supervision, CS prioritizes PD over evaluative judgment.

Baharom (2002) conducted a study on teachers' perceptions of instructional leadership supervision, involving 380 participants from Johor, Malaysia, revealing positive views on supervision's utility but challenges in practical implementation. Hişmanoğlu and Hişmanoğlu (2010) carried out a case study on CS with 42 English instructors in North Cyprus, indicating the strengths of CS for improvement but also some reservations among participants. Moradi et al. (2014) studied the perceptions of 34 Iranian English teachers on CS, finding positive attitudes during observations but noting a focus on identifying weaknesses. Kaneko-Marques' (2015) longitudinal study investigated the integration of video materials in supervision, highlighting the effectiveness of video compared to in-class presence.

RT plays a pivotal role in teachers' professional growth, particularly in the context of the constructivist paradigm (Borg, 2011). It involves teachers reflecting on classroom dynamics and strategies for achieving pedagogical goals (Cruickshank & Applegate, 1981). Bartlett (1990) argues that moving beyond pedagogical methodologies is essential for cultivating a reflective pedagogue. RT is an iterative process in which teachers analyze their teaching experiences, merging theoretical knowledge with practical applications under the guidance of experts (Schon, 1983).

Martin and Double (1998) conducted a pilot study in a tertiary educational institution to enhance instructors' pedagogic skills, revealing majority favorability towards RT but also some unfavorable perceptions. Farrell (2011) studied professional role identities of senior English teachers using RT in a Canadian academic institution, resulting in 16 role identities organized into three clusters. Afshar and Farahani (2015) investigated the connection between RT and reflective thinking among 233 Iranian English instructors, finding a positive correlation between RT and reflective thinking, gender and teaching tenure discrepancies. Mathew et al. (2017) explored the effectiveness of RT for novice teachers, indicating its role in enhancing self-awareness and understanding of vocational paths.

PO, an effective method for improving pedagogical effectiveness, offers various approaches in line with Gosling's (2002) classifications. These include evaluative PO, focused on instructional quality assurance; developmental PO, aimed at enhancing pedagogical competencies; and collaborative PO, emphasizing transformative introspection to elevate pedagogical practices. The utilization of PO as a reflective conduit prompts teachers to reevaluate their instructional practices to foster a comprehensive and holistic approach to teaching (Brockbank & McGill, 2007). It fosters a reciprocal exchange of insights among teachers, encouraging a reflective discourse that encompasses both

emotional aspects of the teaching journey and evaluative feedback from peers (Bell, 2005).

Byrne et al. (2010) conducted empirical research in an English tertiary-level institution, resulting in insights that informed an innovative peer PD framework as an alternative to observation. Dos Santos (2016) examined English language teachers' perceptions of PO's pedagogical utility, finding efficacy for instructional improvement despite participant wariness.

TT, also known as pair teaching, in English education was historically introduced through collaboration with general education (Rao & Chen, 2020). It involves the cooperative participation of multiple instructors (Richards & Farrell, 2005). Solis et al. (2012) describe TT as a pedagogical approach where two instructors equitably share lesson design, presentation and assessment. This pedagogical framework affords teachers the opportunity to distribute pedagogical responsibilities harmoniously while co-delivering English lessons.

Igawa (2009) explored English teachers' perceptions of TT in Japan, emphasizing the importance of enthusiasm and pedagogical skills for effective TT. In their study, Sari et al. (2013) differentiated between professionalism-related issues and interpersonal dynamics in TT. The findings revealed that initially professional concerns shifted to more personal difficulties as collaboration extended.

VC has gained prominence in teacher training and PD due to the proliferation of cost-effective technologies (Gaudin & Chaliès, 2015). Portable digital cameras and applications tailored to PD have made the integration of digital video into teacher education more accessible (Hockly, 2018). VC empowers observers and practitioners to review pedagogical sessions and fosters pre- and post-observational reflection. Video provides a means for teachers to refine their pedagogical practices by examining their teaching methods (Richards & Lockhart, 1994).

Batlle and Miller (2017) examined teachers' beliefs using a customized VC application, highlighting its efficacy, especially within the specific technological context of the study. Yücel et al. (2016) conducted a case study on VC integration in tertiary education, indicating overall positive attitudes towards VC regardless of participants' skill levels and the obligatory nature of VC activities.

Taken together, prior studies confirm that each PDA can be beneficial under certain conditions, yet none compares CS with RT formats in a single tertiary context, and almost all overlook Turkish foundation universities. This absence of comparative data, especially from in-service English instructors whose teaching loads differ markedly from other settings, motivates the present study. Although previous research affirms the value of both CS and RT modalities, three limitations are evident. First, the empirical base is fragmented. Most CS studies draw on single-site qualitative cases or surveys with relatively small samples, while research on peer observation, team teaching and video coaching typically examines each practice in isolation, hindering cross-modal comparison. Second, methodological consistency is lacking. Few investigations adopt explanatory-sequential mixed methods that link survey results to follow-up interviews for deeper explanation. Third, the Turkish tertiary context remains under-represented, even though in-service PDUs are now a standard feature of tertiary level institutions.

Consequently, it is still not known which PDAs English instructors perceive as most valuable when the options are presented side-by-side, nor how those perceptions translate into concrete suggestions for institutional policy. The present study addresses these gaps by (a) offering the first head-to-head comparison of CS and three RT practices within a mixed-methods design, (b) focusing on one of Türkiye's largest tertiary language schools and (c) linking perceived effectiveness to instructors' policy-relevant recommendations.

3. Methodology

3.1. Research Design

This research study employed an explanatory sequential mixed-methods research design to facilitate a multifaceted approach to data collection and analysis that aligns with the research questions addressed. An explanatory-sequential mixed-methods design was chosen for two reasons. First, the quantitative survey phase offered a broad comparison of instructors' perceived effectiveness of CS versus each RT modality (research questions 1 and 2). Second, the follow-up semi-structured interviews explained the survey patterns and elicited actionable recommendations (research questions 3 and 4), which provided the depth and contextual insight unavailable from the survey data alone (Creswell & Plano Clark, 2018). The primary objective of this study was to explore English instructors' perceptions regarding the effectiveness of CS and RT (exemplified through PO, TT and VC modalities). Moreover, the study aimed to identify potential enhancements to the current practices within the research context.

To achieve these objectives, a single-case study approach was meticulously chosen. This methodological choice was rooted in the unique context of English instructors who held reservations concerning the effectiveness of the PDAs. Notably, the research design seamlessly integrated post-positivist and social constructivist paradigms, aligning with Creswell's (2013) theoretical framework.

3.2. Context

Because foundation universities charge tuition and compete for international students, PD decisions in their school of foreign languages directly influence program quality rankings. This study focused on the school of foreign languages of a foundation university in Türkiye. The university is well-regarded for its quality education, comprising 12 faculties, a conservatory, vocational schools and a school of foreign languages. The latter offers languages like German, Chinese, Spanish, Italian and Russian, with English as a primary focus for undergraduate students. With around 130 instructors, it is among Türkiye's busiest language schools, housing three main units: English preparatory, academic English and modern languages, along with sub-units such as program development, testing and PD. The PDU encourages the English instructors' improvement through various activities. CS and RT (PO, TT and VC) are the most important PDAs there. Around 30 CSs occur annually, and the data on preferred types of RT were documented by the PDU in Table 1.

Table 1
Reflective-Teaching Activities by Term

Academic Year-Term	PO	TT	VC	Total
	(*n*)	(*n*)	(*n*)	(*N*)
2017-2018-Fall	58	8	13	79
2018-2019-Fall	54	6	15	75
2019-2020-Fall	57	4	27	88

The school's quality manual outlines that CS aims to enhance observed instructors' PD through feedback, comments, and self-reflection. Any instructor's class can be observed by the PDU members as needed, regardless of experience. RT aims to enhance collaboration, collegiality, critical reflection, and self-awareness among the instructors. All the PDAs follow three stages: pre-observation, during observation, and post-observation, using common forms across activities along with a reflection form.

3.3. Participants

This study involved two primary participant groups: the English instructors (*N*=91) completing the survey (*n*=89) and some of whom (*n*=20) were interviewed and the PDU members (*n*=2). During the 2020-2021 academic year fall term, the surveyed and interviewed instructors worked in either the academic English unit or the English preparatory unit. Those in the PDU had extra responsibilities related to PD, alongside teaching duties.

The English instructors not engaging in the PDAs (*N*=14) including the vice director, department chair, heads of the English preparatory and academic English units, testing and program development units and deputy heads were excluded to maintain participant count. Nonetheless, they participated in tool piloting to retain possible participants for actual data collection.

3.3.1. English Instructors

The initial participant group included the English instructors at the school of foreign languages. Following the pilot phase, all the instructors were contacted for survey completion. A total of 98 English instructors participated; nevertheless, nine participants not filling in two trick items correctly were excluded, leading to a final group of 89 English instructors. The demographic particulars of these participants were systematically presented, encompassing details concerning their age, years of English teaching experience, bachelor's degree and any additional responsibilities they were shouldering. These aspects were meticulously tabulated for comprehensive exposition in Table 2.

Table 2
Participant Demographics

Age	*n*	*%*	Years of English Teaching Experience	*n*	*%*	Bachelor's Degree	*n*	*%*	Additional Responsibility	*n*	*%*
21-30	15	16.8	1-5	11	12.3	English Language Teaching	34	38.2	Only Test Development Unit	4	4.5
31-40	11	12.4	6-10	10	11.2	English Language and Literature	28	31.5	Only Program Development Unit	4	4.5
41-50	50	56.2	11-15	3	3.4	American Culture and Literature	13	14.6	Both Test and Program Development Unit	29	32.6
51-60	9	10.1	16-20	8	9	English Linguistics	9	10.1	No Additional Responsibilities	49	55
61-70	4	4.5	21-25	41	46.1	English Translation and Interpreting	3	3.4	Other Responsibilities	3	3.4
			26+	16	18	Biology	1	1.1			
						Psychology	1	1.1			
Total	89	100		89	100		89	100		89	100

In relation to gender distribution, a substantial majority of 88.8% comprised female participants (*n*=79), whereas the remaining 11.2% were male (*n*=10). Regarding the educational qualifications of the participants from the survey, a notable proportion, comprising 39.3%, possessed a master's degree (*n*=35). In terms of the courses they were teaching, a majority of the participants, specifically 58.4%, were found to be involved in teaching departmental English courses (*n*=52), whereas the remaining 41.6% were dedicated to teaching classes within the English preparatory program (*n*=37).

To ensure representative interview sessions among the English instructors surveyed, 20 instructors were selected by making use of quota sampling, nearly a quarter of respondents. This aimed for a demographically balanced group aligned closely with the survey participants. The quotas were initially based on the courses they were teaching, followed by gender distribution. Cross-multiplication determined the interviewees from subgroup instructors teaching departmental English (*n*=12) and preparatory courses (*n*=8). Likewise, representation for the male instructors (*n*=10) guided appropriate the female participants through subtraction. Other potential variables were beyond this study's scope due to complexity and practical constraints.

3.3.2. PDU Members

Following the interviews with the English instructors, two PDU members were interviewed for in-depth exploration. One of them, aged 39, was an experienced English instructor with 17 years of teaching experience. She had a bachelor's degree in English translation and interpreting from a state university, as well as a master's degree in curriculum and instruction. The other PDU member, aged 32, was also an experienced one with a decade of experience. She majored in English linguistics and held a master's degree in curriculum and instruction, further enhancing her expertise.

Together, both PDU members were instrumental in fostering a positive and enriching atmosphere within the school of foreign languages. Their primary objective was to elevate the quality of education by equipping their colleagues with essential skills and techniques. Notably, they played pivotal roles in overseeing CS initiatives and mentoring, all while advocating for the adoption of RT practices among the English instructors.

3.4. Data Collection

In the study, an amalgamation of both quantitative and qualitative data collection methodologies was employed to comprehensively address the research questions. Accordingly, the research design encompassed a multifaceted approach, incorporating a survey, semi-structured interview questions tailored for the English instructors, as well as an additional set of semi-structured interview questions designated for the PDU members. All three data collection instruments underwent a meticulous piloting phase prior to the commencement of the actual data collection for the study. This pilot testing procedure was conducted with the English instructors affiliated with the school of foreign languages, the very setting within which the study was carried out. All the necessary consents were taken from the board of the university and the school of foreign languages in addition to the informed consents from the participants.

The survey adhered closely to the quality manual established for the accreditation procedures of the school of foreign languages at the foundation university. This comprehensive manual details the PDU, its activities and other institution units. The survey was distributed using an online platform, accessible through instant messaging apps or email based on participants' preference. The informed consents were obtained via electronic signatures within the survey platform. To ensure content validity, a panel of six experts was consulted during survey development, including four experts with PhDs in English language teaching and two experienced statisticians. The survey underwent five meticulous revisions, involving item addition, elimination and modification. After these iterative refinements, a pilot phase was conducted to assess survey reliability. The survey prepared in English featured four sections. The first section contained

eight items capturing participants' demographic details. Following this, a nine-item segment employed a five-point Likert scale to compare participants' perceptions towards the effectiveness of CS and RT. The next section consisted of 13 items on a five-point Likert scale, collecting data on their perceptions of PO, TT and VC. The final section included four items prompting the participants to reflect on whether the discussed PDAs should remain the same, be modified, or be removed, with brief explanations. Two additional items designed as "trick items" were included alongside the items in the second and third sections. These were strategically crafted to gauge participants' attention and thoroughness in survey completion. The participants deviating from the expected response were excluded from the piloting phase. Importantly, all the participants involved in piloting (N=30) correctly responded to these specific items as intended. Exploratory factor analysis with Varimax Rotation Method confirmed the expected two-factor structure of the section two (Kaiser-Meyer-Olkin Test Value=.82; Bartlett's Test of Sphericity=χ^2(28)=337.17, p<.001). The two factors (eigenvalues>1) explained 91.8% of the variance, and all eight items loaded strongly on their intended factor (χ=.91–.96). The section three also produced a stable three-factor solution (Kaiser-Meyer-Olkin Test Value=.86; Bartlett's Test of Sphericity=χ^2(66) = 584.76, p < .001) that accounted for 92.5 % of the variance, with all 13 items loading ≥ .73. Following the factor analysis, in accordance with the calculation of Cronbach's Alpha, the survey was found to be highly reliable (α=.95) during the piloting. The same statistical tests were also employed for the actual data collection, and the two-factor model of the section two was replicated (Kaiser-Meyer-Olkin Test Value=.71; Bartlett's Test of Sphericity=χ^2(28)=484.69, p<.001), explaining 73.8% of the variance; item loadings ranged from .72 to .92. The section three again yielded the three expected factors (Kaiser-Meyer-Olkin Test Value=.82; Bartlett's Test of Sphericity=χ^2(66)=1125.96, p<.001) that together explained 83.9 % of the variance, with loadings between .81 and .93. All factors therefore met conventional criteria for construct validity and reliability. Upon the

completion of data collection from the cohort of English instructor survey participants, a subsequent verification of the survey's reliability was conducted using Cronbach's Alpha, and it was found to be highly reliable (α=.89) once again.

In addition to the survey, two sets of semi-structured interview questions were developed: one for the English instructors and the other for the PDU members. These sets were crafted in Turkish, the participants' native language, to ensure comfort and minimize communication issues. The interviews were also conducted in Turkish for ease and to avoid misunderstandings, with certain English concepts retained for familiarity. Online pseudonyms were assigned to the English instructor interviewees to ensure confidentiality. The first set of questions were prepared based on the survey findings, while the second set stemmed from the interviews with the English instructors. Before the interviews, both sets were piloted, as well. The first interview process began by asking participants to discuss the effectiveness of the PDAs using a visual summary of the findings from the survey. They then compared CS and RT in terms of effectiveness. The survey's fourth section findings, evaluating the PDAs, were presented, leading to participant comments and recommendations to improve their effectiveness. The second interview process commenced by gathering the participants' perceptions on the effectiveness of the PDAs, proceeding sequentially. The conversations included their reflections on insights from the interviews with the English instructors. The interviews ended by sharing the English instructors' suggestions, prompting the PDU members' insightful comments on them.

3.5. Data Analysis

During the initial stage of evaluating the two scales in the second and third sections of the survey, several statistical analyses were performed using SPSS 23. These analyses consisted of the Kaiser-Meyer-Olkin Test, Bartlett's Test of Sphericity, factor analysis with Varimax Rotation Method and reliability analysis via Cronbach's Alpha calculation. The scales' reliability was verified, and

the same tests were applied to the main dataset. After confirming scale reliability, the quantitative data were subjected to frequency analysis using SPSS 23.

However, the survey's final section contained open-ended questions. Since the qualitative data from the survey were relatively limited, they were analyzed using content analysis following the framework outlined by Miles and Huberman (1994) on NVIVO 10. The survey was originally in English; thus, the participants responded to the open-ended questions in English, as well. Subsequently, the researcher translated these English responses into Turkish. This translation was done to integrate the responses into the semi-structured interview questions for the English instructors. To ensure validity, another expert translated the Turkish versions back into English. The data from the semi-structured interviews with the English instructors and the PDU members underwent content analysis

using Yin's (2014) method, employing again NVIVO 10, for the qualitative data were much richer compared to the data gathered from the survey. In order to implement internal validation measures and enhance the reliability of the results, an intercoder agreement approach was employed, and another researcher assumed the role of the second coder and analyzed 20% of both qualitative datasets, utilizing the same content analysis frameworks.

4. Findings and Discussion

4.1. Comparing the Effectiveness of Clinical Supervision vs. Reflective Teaching

The frequency analysis showed that CS (*M*=3.84) was perceived as more effective than RT (*M*=3.35). These results were further elaborated in Table 3.

Table 3
Perceived Effectiveness of Clinical Supervision vs. Reflective Teaching

	CS								RS							
	M	*SD*	1* %	*n*	2* %	*n*	3* %	*n*	*M*	*SD*	1* %	*n*	2* %	*n*	3* %	*n*
pre-observation	3.58	1.02	13.4	12	33.8	30	52.8	47	3.06	1.07	31.5	28	30.3	27	38.2	34
observation	3.6	1.09	16.9	15	26.9	24	56.2	50	3.46	1.22	21.4	19	26.9	24	51.7	46
post-observation	3.98	1.08	9	8	21.4	19	69.6	62	3.4	1.21	24.8	22	22.5	20	52.8	47
feedback	4.19	.86	3.3	3	15.8	14	80.9	72	3.47	1.12	20.2	18	25.9	23	53.9	48

(CS overall *M* = 3.84; RS overall *M* = 3.35)

*1=Total of "Completely disagree" and "Disagree", 2= "Somewhat agree", 3=Total of "Agree" and "Completely agree"

The surveyed English instructors in this study expressed the perspective that CS yielded greater effectiveness in comparison to RT. This particular finding aligns with the assertions posited by Beach and Reinhartz (2000), who had previously noted that a substantial portion of teachers held a favorable view of CS, which corroborates the findings of the current study. On the other hand, the findings of the present study stand in contrast to the conclusions drawn from Inyamah's (2011) research, which had advocated a perspective of the

ineffectiveness of CS. It is noteworthy that this disparity may be attributed to several factors. For instance, in this study, CS might have been perceived as more effective than RT due to its potential for providing clearer and more immediate feedback (Phelps, 2013). Additionally, the structured approach and immediate outcomes associated with CS might have contributed to its perceived effectiveness, as indicated by Marwati et al. (2019).

4.2. Comparing the Effectiveness of Peer Observation, Team Teaching and Video Coaching

Among the three types of RT, VC garnered the highest perceived effectiveness among the English instructors (*M*=3.64). Subsequently, TT (*M*=3.46) and PO (*M*=3.45) trailed closely, displaying marginal differentials between them in terms of perceived effectiveness. Further elucidation regarding the particulars of each RT type was provided in Table 4.

Table 4
Perceived Effectiveness of Reflective Teaching Modalities

			Collaboration	Collegiality	Critical Reflection	Self-awareness
Peer Observation	*M*		3.21	3.25	3.4	3.94
			3.45			
	SD		1.18	1.19	1.13	1.06
	1*	%	23.6	23.6	18	9
		n	21	21	16	8
	2*	%	31.4	28	29.2	14.6
		n	28	25	26	13
	3*	%	45	48.4	52.8	76.4
		n	40	43	47	68
Team Teaching	*M*		3.54	3.39	3.43	3.47
			3.46			
	SD		1.26	1.15	1.13	1.2
	1*	%	20.2	22.5	24.7	22.5
		n	18	20	22	20
	2*	%	21.4	30.3	23.6	23.6
		n	19	27	21	21
	3*	%	58.4	47.2	51.7	54
		n	52	42	46	48
Video Coaching	*M*		3.17	3.19	3.94	4.26
			3.64			
	SD		1.34	1.29	1.13	1.07
	1*	%	34.9	33.7	13.5	10.1
		n	31	30	12	9
	2*	%	23.6	23.6	15.8	10.1
		n	21	21	14	9
	3*	%	41.6	42.7	70.7	79.8
		n	37	38	63	71

*1=Total of "Completely disagree" and "Disagree", 2="Somewhat agree", 3=Total of "Agree" and "Completely agree"

The English instructors in this study collectively held the viewpoint that VC emerged as the preeminent form of RT, with TT and PO following in succession. This notable finding is consistent with the research conducted by Mathew et al. (2017), which underscores the potential of VC as a viable avenue for effective RT. Their study provides empirical evidence supporting the notion that VC holds promise as a modality for enhancing pedagogical expertise. Specifically, teachers can harness video recordings, whether self-generated or featuring peer demonstrations, to advance their instructional skills.

Video recordings afford teachers the opportunity to autonomously revisit their instructional interactions, fostering a culture of self-reflection and facilitating a more profound understanding of their pedagogical methodologies (Suhrheinrich & Chan, 2017). Furthermore, the integration of video as a pedagogical tool offers

tangible instances of instructional practices, which in turn expedites deliberations and enables constructive evaluations during reflective discussions (Setyaningrahayu et al., 2019).

4.3. The Effectiveness of the Professional Development Activities

In accordance with the findings gathered from the last part of the survey, the fraction marginally below half of the participants (44.9%; *n*=40) thought that CS needed to be modified as it needed an overall change (*f*=12), was repetitive (*f*=5), inefficient (*f*=5), stressful (*f*=4) and was considered as a burden (*f*=4). A proportion exceeding one-third of the English instructors (36%; *n*=32) believed that PO needed to be removed. This perspective was substantiated by their characterization of PO as being characterized ineffective (*f*=10), useless (*f*=9) and most of the English instructors pretended to conduct it (*f*=5). It was also mentioned that PO was unreliable (*f*=4) and superficial (*f*=4) in addition to feeling insecure during the process (*f*=4).

Table 5
Recommendations for Professional Development Activities

	Clinical Supervision		Reflective Teaching					
			Peer Observation		Team Teaching		Video Coaching	
	%	n	%	n	%	n	%	n
could remain the same	40.4	36	34.8	31	36	32	65.2	58
could be modified	44.9	40	29.2	26	27	24	21.3	19
could be removed	14.6	13	36	32	37.1	33	13.5	12

A proportion surpassing one-third of the English instructors (37.1%; *n*=33) indicated that TT needed to be removed because it was complicated (*f*=14), ineffective (*f*=12), useless (*f*=10) and unnecessary (*f*=7). A proportion slightly below two-thirds of the participants (65.2%; *n*=58) believed that VC needed to be remained the same because it was effective (*f*=19), gave them an opportunity for self-observation (*f*=16) and self-reflection (*f*=16), increased their self-awareness (*f*=14), was beneficial (*f*=10), could be conducted in the natural classroom atmosphere (*f*=6), and they felt comfortable during the process (*f*=5).

The surveyed English instructors said that CS increased their stress and therefore needs adjustment. This concern, which resonated during the interviews with the English instructors, shed light on elevated stress levels experienced by both instructors and students due to the supervision process. Notably, this alignment with the stress aspect of CS corresponds to Glickman and Tamashiro's (1980) contention that unclear objectives can induce stress among teachers. When subjected to observation, the participants may have felt increased pressure, which could have influenced their teaching methods and causing stress (Abbott & Carter, 1985). Beyond the issue of stress, the surveyed English instructors also expressed the belief that CS could benefit from modifications due to its repetitive nature. They observed that the effectiveness of CS diminished over time, especially for long-serving individuals who were frequently observed by the PDU. This finding is in accordance with Moradi et al.'s (2014) research, which suggested that although valuable, CS can become noticeably repetitive with prolonged use. This repetitiveness may have led to diminishing returns in terms of gaining new insights, and as a result, its perceived value among the participants could have decreased (Ghapanchi & Baradaran, 2012). The survey results revealed that the perception of CS as burdensome stemmed from stress and repetition, exacerbated by substantial work responsibilities, particularly for those in the English preparatory unit. This finding parallels Hişmanoğlu and Hişmanoğlu's (2010) observations

of similar workload-related hindrances affecting PD and attitudes towards supervision. The participants may have viewed CS as burdensome rather than a worthwhile process if they had not perceived immediate benefits or changes thanks to it (Paba, 2017). Lastly, the interviews with English instructors and PDU members revealed a theme related to the effectiveness of CS. The English instructors noted that CS was most effective when conducted by proficient, objective and organized PDU members. This finding aligns with existing literature that indicates the superior observational skills of such members (Centra, 1993; Hammersley-Fletcher & Orsmond, 2004; Keig & Waggoner, 1994; Manning, 1986). However, this contrasts with Baharom's (2002) findings, which occasionally identified inadequacies in CS execution by supervisors. Additionally, the English instructors stressed the effectiveness of CS in recognizing their personal strengths and areas for improvement, consistent with literature emphasizing supervision's role in enhancing teachers' professionalism (Glickman et al., 1995; Thomas, 2008). Similarly, the PDU members emphasized the effectiveness of CS, particularly when adopting a developmental rather than judgmental approach (Glickman et al., 2004), reinforcing its role in pedagogical growth. The practical application of criticism and suggestions from CS may have been linked to its effectiveness (Hişmanoğlu & Hişmanoğlu, 2010).

Conducting RT carelessly was apparent in both groups. The insights gleaned from the interviews exposed a recurring trend of RT protocols not being rigorously adhered to, particularly by the senior instructors. Among the English instructors, a prevalent inclination was observed towards opting for less demanding RT type. Furthermore, instances of simulated engagement in RT were evident, particularly when paired with less familiar or less preferred peers. This is in line with the findings of Richards and Lockhart (1994), who identified teachers' reluctance stemming from the perceived evaluative nature of reflective practices. In addition, Martin and Double's (1998) research underscores how certain teachers perceive reflection as an obligatory task, potentially leading to lackluster execution. Echoing these sentiments, Florez (2001)

posits that teachers may find themselves unprepared to confront uncertainties within reflective practices, which might lead them to avoid or engaging superficially in such endeavors. It is conceivable that the participants' initial enthusiasm for RT may have waned over time, resulting in less consistent efforts, as suggested by Kayaoğlu (2012).

The surveyed English instructors predominantly leaned towards the perspective that PO needed to be removed because most of the peers were pretending to conduct it. The insights garnered from the interviews with the English instructors further highlighted the existence of individuals who approached PO as a procedural requirement, rendering it ineffective in terms of reliability and perceived utility. Byrne et al. (2010) likewise found that some teachers view reflective practice as a compliance exercise. The PDU members also acknowledged the presence of instructors who simulated participation during PO while also recognizing those who genuinely adhered to observation guidelines. They attributed this discrepancy to the instructors' individual approaches to the process. This perspective resonates with Brockbank and McGill's (2007) conceptualization of PO as a developmental opportunity for mutual enrichment. It is plausible that the participants may have been less motivated to engage in PO seriously if they had not perceived clear advantages or incentives associated with diligent participation (Shortland, 2004). The findings from the survey, coupled with the insights from the interviews conducted with the English instructors and the PDU members, echoed a prevailing sentiment among the instructors advocating for the removal of PO due to its perceived superficiality. This sentiment emerged in a context of heightened stress levels for both instructors and students, leading to a proclivity towards superficial teaching practices. Therefore, PO mirrored this superficial teaching atmosphere. This corresponds with Race and Fellows' (2009) study, which addresses the issue of superficiality in PO and suggests mitigation through the accumulation of iterative observation instances. In the absence of clearly articulated goals and incentives, the participants may have primarily focused on completing the procedural aspects of PO

rather than engaging with its intended purpose (Hammersley-Fletcher & Orsmond, 2005).

Based on the findings of the survey, the English instructors perceived the removal of TT as a warranted step, primarily due to its complicated nature. This viewpoint was substantiated through the interviews with the instructors who further scrutinized the effectiveness of TT, indicating the intricate nature of its reflective pedagogy and the demands associated with collaborative teaching roles. Notably, a pronounced reluctance towards TT emerged, primarily attributed to the heightened preparation requirements when compared to PO and VC. This perception aligns with Igawa's (2009) research, which delved into English teachers' perspectives on the complexities of TT. Sari et al. (2013) similarly shed light on the intricacies associated with TT in their research. The intricacies encompassed aspects such as coordinating between instructors, harmonizing instructional strategies and managing varying teaching philosophies, all contributing to the perceived complexity (Jo & Woo, 2010). Furthermore, the PDU members acknowledged the complexities inherent in TT but proposed that diligent inquiry and accessible digital resources could render it more tenable. A pivotal aspect of effective TT, as both the English instructors and PDU members concurred, lay in the meticulous planning of the teaching collaboration. This finding is in accordance with Tim's (2018) study, which underscores the challenges of organizing productive TT, requiring extensive discourse and groundwork. Bailey (1996) similarly asserts that while teachers cannot anticipate every classroom scenario, a comprehensive blueprint remains pivotal in TT to navigate unforeseen situations effectively during co-teaching. The absence of proper planning could potentially result in inconsistencies in subject coverage, leading to student perplexity (Ken-Maduako & Oyatogun, 2015).

Given the demonstrated effectiveness of VC, as evidenced in the survey results, the English instructors were inclined to advocate for the continued utilization of this practice. In addition, beyond classroom observation, both the interviewed English instructors and the PDU members reached a consensus regarding the multifaceted effectiveness of VC, particularly in its capacity to illuminate instructional strengths and weaknesses. This finding is consistent with the research conducted by Gaudin and Chaliès (2015), which affirms the potential of VC for fostering reflective practices among teachers, enabling them to critically assess their instructional methods. It is worth noting that the utilization of video recordings within VC may have served to mitigate potential biases that can arise during the evaluation of educational procedures, as suggested by Borer et al. (2018). Moreover, the effectiveness of VC can be attributed to the heightened comfort level of English instructors when it comes to both providing and receiving feedback, supported by tangible video evidence. This is in line with the insights gleaned from Batlle and Miller's (2017) study, which stresses the credibility enhancement of feedback through the use of video recordings as tangible evidence during post-observation phases and reinforces its effectiveness. A recurring theme that emerged consistently in both the survey responses and the interviews with the PDU members revolved around the concept of a natural classroom atmosphere. The PDU members expressed the viewpoint that VC, in comparison to PO and TT, fostered a greater degree of authenticity. This authenticity was attributed to the absence of external observers during VC. This perspective resonates with Hockly's (2018) study, which highlights the unique quality of VC to provide genuine insights into classroom dynamics for teachers, offering valuable insights into their own teaching practices as well as those of their colleagues.

These findings answer the gap outlined in the Introduction: they deliver the first comparative effectiveness profile of four PDAs in a Turkish school of foreign languages, revealing clear instructor preferences (VC > TT ≈ PO, CS superior overall) and pinpointing resource-light modifications (e.g., video-supported CS, control mechanisms for PO). As such, the study provides the evidence base that Turkish PDUs and accreditation bodies have lacked when prioritizing PD formats.

4.4. Institutional Recommendations for the Professional Development Activities

Concerning CS, the English instructors emphasized the potential enhancement of CS through the integration of video technology, highlighting its capacity to improve facilitation and overall effectiveness. Similarly, the PDU members concurred that incorporating video in the observational phase of CS could augment its effectiveness. This aligns with Kaneko-Marques' (2015) argument regarding the undeniable contribution of video recordings to teachers' PD. The utilization of video in CS could have provided more flexibility in observing diverse classroom settings and learning environments, as suggested by Marrow et al. (2002).

The English instructors proposed updates to the forms utilized for RT to enhance its effectiveness within the research context. Their rationale stemmed from the perception that existing forms were outdated and overly complex. They believed that revisions needed to consider institutional dynamics and provide comprehensive details, supporting Quesada Pacheco's (2005) argument for tailored tools to amplify pedagogical efficacy through RT. Improved forms may result in more insightful feedback and consequently bolster the potency of RT, as posited by Prieto et al. (2020). In contrast, the PDU members suggested an alternative approach to enhance the effectiveness of RT, which involved recruiting younger English instructors. They proposed that these recruits, motivated by novelty and openness to new ideas, might exhibit a greater interest in PD. This perspective corresponds with Farrell's (2011) study regarding the potential reluctance of seasoned teachers to embrace reflective practices. However, it differs from Afshar and Farahani's (2015) research, which suggests an inverse correlation, where greater experience in English teaching correlates with increased participation in RT.

During the interviews with both the English instructors and the PDU members, a common theme emerged regarding suggestions for PO. Both groups mentioned the necessity of implementing a control mechanism for PO. Without such control, the effectiveness of PO suffered, with instances where instructors refrained from engaging in pre- and post-observation phases or mutual observation. This supports Dos Santos' (2016) findings, which suggest that teachers may deviate from meticulous execution of PO and thus require oversight. The presence of a control mechanism may hold the participants accountable for their engagement in the PO process, as noted by Hammersley-Fletcher and Orsmond (2004). Similarly, the PDU members shared this view and emphasized the need for oversight. Nonetheless, logistical and administrative limitations hindered full implementation. With only two members, the unit's limited personnel capacity exacerbated this issue, and the PDU members were unable to fully address this administrative requirement.

In order to enhance the effectiveness of TT, both the English instructors and the PDU members endorsed the strategic pairing of instructors teaching similar courses, especially within the same unit. This finding is in line with Carless' (2006) findings, asserting that pairing teachers with similar courses enhances the effectiveness of TT. Such pairings provide diverse techniques and novel concepts, aiding the assimilation of differentiated approaches within the same framework, a point echoed in Tim's (2018) study.

Regarding VC, a common theme emerging from the interviews was the need for technical equipment support from the institution, such as cameras and tripods. In the absence of such support, the instructors often relied on fellow instructors or students for VC. The PDU members also noted that these resources could enhance the effectiveness of VC and encourage its adoption. This is in accordance with Yücel et al.'s (2016) research, which advocate for technical assistance to optimize the effectiveness of VC as a RT form, emphasizing its instructional benefit.

The findings suggest three practical recommendations for implementation in tertiary language education institutions. First, the integration of VC should be prioritized through the provision of a limited number of portable recording devices (e.g., cameras and tripods) and the adoption of a concise peer-feedback template. This approach ensures the alignment of institutional resources with practices that instructors identify as most

beneficial. Second, CS processes may be optimized by focusing on novice instructors while transitioning experienced staff to biennial review cycles, a change that could alleviate reported stress and reduce procedural redundancy. Third, traditional PO practices characterized by extensive documentation and limited impact could be replaced with brief, goal-oriented "teaching huddles" (collaborative sessions wherein two instructors engage in mutual goal setting, classroom observation and debriefing within a one-week period). Oversight by the PDU is recommended to ensure meaningful engagement and mitigate performative compliance. These modifications can be trialed over a single academic term and assessed using the same survey and interview instruments employed in the current study.

5. Conclusion

The principal objective of this research study entailed a comprehensive exploration of the English instructors' perspectives regarding the effectiveness of CS and RT. Moreover, it endeavored to scrutinize the effectiveness of specific RT modalities, namely, PO, TT and VC. Finally, where deemed essential, this research aspired to proffer constructive recommendations for the enhancement of extant PDAs. To begin with, the findings of this study revealed that CS was perceived to possess a higher level of effectiveness in contrast to RT. Next, among the three forms of RT, VC was deemed to exhibit the highest level of perceived effectiveness by the participants, followed by TT and PO. Furthermore, the participants expressed the belief that CS needed to be modified, citing reasons such as the necessity for an overall change, repetitiveness, ineffectiveness, stressfulness, and considering it as a burden. They advocated for the removal of peer observation, and this was underscored by their characterization of PO as ineffective, useless, often being conducted superficially, unreliability, superficiality, and a sense of insecurity during the process. Similarly, they believed that TT needed to be removed due to reasons such as complexity, ineffectiveness, futility, and redundancy. In contrast, the participants advocated for maintaining the current practice of VC, emphasizing its

effectiveness, facilitation of self-observation and self-reflection, enhancement of self-awareness, beneficial attributes, conduciveness to a natural classroom atmosphere and the comfort it provided during the process. Lastly, the English instructors indicated that the implementation of CS via video technology could significantly enhance its facilitation and overall effectiveness. They recommended the modernization of the forms employed for RT to increase its effectiveness within the research context. Alternatively, the PDU members proposed an alternative strategy to enhance the effectiveness of RT, which involved the recruitment of younger English instructors. Both the English instructors and the PDU members emphasized the necessity of implementing a control mechanism to bolster the effectiveness of PO. In the context of TT, the participants advocated for the strategic pairing of instructors responsible for teaching similar courses as a means to enhance its effectiveness. Regarding VC, technical equipment support from the institution was suggested to optimize its effectiveness.

To translate these findings into practice, it is recommended that tertiary language-teaching units and relevant governing bodies (1) allocate a small recurring budget for portable video-coaching kits and a one-page peer-feedback template, (2) adopt a developmental CS cycle that concentrates on novice instructors while moving experienced staff to biennial reviews, (3) replace paperwork-heavy peer observation with brief "teaching huddles" that are completed and logged within one week and (4) embed evidence-based PD targets into accreditation criteria and institutional key performance indicators. By implementing these steps over a single academic year and evaluating them, institutions can align PD resources with activities that English instructors themselves deem most effective.

The study highlights the importance of enhancing PD for English teachers, starting with pre-service and continuing with in-service education (Evişen, 2021). An undergraduate course could introduce prospective English teachers to the value of PD. Besides, ongoing in-service training sessions might be implemented across educational levels, from kindergarten to tertiary education, to

maintain positive perceptions of PD among English teachers. An additional implication may be to conduct a survey allowing teachers to select their preferred RT activity each term. Categorizing teachers based on preferences and enabling partner/group selection from a published list could encourage greater active engagement and reduce feigned participation, enhancing the effectiveness of these activities (Diaz-Maggioli, 2004). Finally, institutions might enhance RT by offering a wider range of activities for teachers to choose from. This diversification may reduce the perception of repetitiveness and stimulate greater engagement, ultimately improving the overall effectiveness of these activities (Fakazlı, 2021).

Future research can investigate the viewpoints of English instructors in different educational contexts, including various schools of foreign languages within foundation universities with similar PD programs. Additionally, conducting comparative analyses between English instructors in foundation universities' schools of foreign languages and those in state universities may provide valuable insights for subsequent studies. Prospective research could also delve into the disparities between English instructors' professed beliefs and their actual practices in PD. Initially, a survey might be used to gather instructors' beliefs about PDAs. Subsequently, interviews with simulated scenarios could be employed to prompt participants to discuss their anticipated actions and emotional responses. This approach would provide insights into the alignment between beliefs and actual practices in PDAs among English instructors.

References

Abbott, S., & Carter, R. M. (1985). Clinical supervision and the foreign language teacher. *Foreign Language Annals, 18*(1), 25-30. https://doi.org/10.1111/j.1944-9720.1985.tb01762.x

Afshar, H. S., & Farahani, M. (2015). Reflective thinking and reflective teaching among Iranian EFL teachers: Do gender and teaching experience make a difference? *Procedia-Social and Behavioral Sciences, 192*, 615-620. https://doi.org/10.1016/j.sbspro.2015.06.107

Ahmed, E., Nordin, Z. S., Shah, S. R., & Channa, M. A. (2018). Peer observation: A professional learning tool for English language teachers in an EFL institute. *World Journal of Education, 8*(2), 73-87. https://doi.org/10.5430/wje.v8n2p73

Amini, S., & Gholami, J. (2018). Professional development of EFL teachers through rotatory peer supervision. *Journal of Teacher Education for Sustainability, 20*(2), 101-117. https://doi.org/10.2478/jtes-2018-0018

Baharom, M. (2002). *Perceptions of teachers towards teaching leadership in computer literacy in Johor state schools* (Doctoral dissertation). Universiti Kebangsaan Malaysia.

Bailey, K. M. (1996). The best laid plans: Teachers' in-class decisions to depart from their lesson plans. In K.M. Bailey, & D. Nunan (Eds.), *Voices from the language classroom* (pp. 15-40). Cambridge: Cambridge University.

Bailey, N. M., & Van Harken, E. M. (2014). Visual images as tools of teacher inquiry. *Journal of Teacher Education, 65*(3), 241-260. https://doi.org/10.1177/0022487113519130

Baltaci-Goktalay, S., Kesner, J., Uzun, A., Bulunuz, N., Gürsoy, E., & Bulunuz, M. (2014). Redefining a teacher education program: Clinical supervision model and UludagKDM. *International Journal on New Trends in Education and Their Implications, 5*(2), 1-11.

Barahona, M. (2017). Exploring models of team teaching in initial foreign/second language teacher education: A study in situated collaboration. *Australian Journal of Teacher Education, 42*(12), 9.

Bartlett, L. (1990). Teacher development through reflective teaching. *Second Language Teacher Education*, 202-214.

Batlle, J., & Miller, P. (2017). Video enhanced observation and teacher development: Teachers' beliefs as technology users. *Proceedings of the 9th International Conference on Education and New Learning Technologies,* 2352-2361. https://doi.org/10.21125/edulearn.2017.1487

Beach, D. M., & Reinhartz, J. (2000). *Supervisory leadership: Focus on instruction*. Boston: Allyn & Bacon.

Bell, M. (2005). *Peer observation partnerships in higher education.* Milperra: Higher Education Research and Development Society of Australasia.

Borer, V. L., Muller, A., & Flandin, S. (2018). Referentiality in secondary teachers' video observation of others' teaching. *Contemporary Issues in Technology and Teacher Education, 18*(1), 149-167.

Borg, S. (2011). Language teacher education. In Simpson (Ed.), *The Routledge handbook of applied linguistics* (pp. 215–228). London: Routledge.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher, 33*(8), 3-15. https://doi.org/10.3102/0013189X033008003

Brockbank, A., & McGill, I. (2007). *Facilitating reflective learning in higher education* (2nd ed.). Open University Press.

Byrne, J., Brown, H., & Challen, D. (2010). Peer development as an alternative to peer observation: A tool to enhance professional development. *International Journal for Academic Development, 15*(3), 215-228. https://doi.org/10.1080/1360144X.2010.497685

Carless, D. R. (2006). Good practices in team teaching in Japan, South Korea and Hong Kong. *System, 34*(3), 341-351. https://doi.org/10.1016/j.system.2006.02.001

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

Centra, J. A. (1993). R*eflective faculty evaluation: Enhancing teaching and determining faculty effectiveness.* San Francisco: Jossey Bass.

Chien, C. W. (2019). Integration of technical vocabulary into peer observation of teaching for Taiwanese elementary school English teachers' professional learning. *Education 3-13, 47*(2), 176-190. https://doi.org/10.1080/03004279.2017.1420672

Copland, F. (2010). Causes of tension in post-observation feedback in pre-service teacher training: An alternative view. *Teaching and Teacher Education, 26*(3), 466-472. https://doi.org/10.1016/j.tate.2009.06.001

Creswell, J. W. (2013). *Qualitative inquiry and research design: Choosing among five approaches*. California: SAGE.

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed.). SAGE Publications.

Cruickshank, D. R., & Applegate, J. H. (1981). Reflective teaching as a strategy for teacher growth. *Educational Leadership*, *38*(7), 553-54.

Darling-Hammond, L., & Richardson, N. (2009). Teacher learning: what matters? *Educational Leadership, 66*(5), 46–53.

Desimone, L. M. (2011). A primer on effective professional development. *Phi Delta Kappan*, *92*(6), 68-71. https://doi.org/10.1177/003172171109200616

Diaz-Maggioli, G. (2004). *Teacher-centered professional development*. Association for Supervision and Curriculum Development.

Dos Santos, L. M. (2016). Foreign language teachers' professional development through peer observation programme. *English Language Teaching*, *9*(10), 39-46. https://doi.org/10.5539/elt.v9n10p39

Dos Santos, L. M. (2017). How do teachers make sense of peer observation professional development in an urban school. *International Education Studies*, *10*(1), 255-265. https://doi.org/10.5539/ies.v10n1p255

DuFour, R. (2004). What is a "professional learning community"? *Educational Leadership*, *61*(8), 6-11.

Evişen, N. (2021). Turkish EFL pre-service and in-service teachers' views on professional development and related activities. *Focus on ELT Journal*, *3*(1), 43-64. https://doi.org/10.14744/felt.2021.00048

Fakazlı, Ö. (2021). Different ways of promoting reflective teaching. *Journal of Social and Humanities Sciences Research, 8*(69), 1255-1267. https://doi.org/10.26450/jshsr.2429

Farrell, T. S. (2011). Exploring the professional role identities of experienced ESL teachers through reflective practice. *System*, *39*(1), 54-62. https://doi.org/10.1016/j.system.2011.01.012

Fiszer, E. P. (2004). *How teachers learn best: An ongoing professional development model*. Lanham: Scarecrow Education.

Flint, A. S., Zisook, K., & Fisher, T. R. (2011). Not a one-shot deal: Generative professional development among experienced teachers. *Teaching and Teacher Education, 27*(8), 1163-1169. https://doi.org/10.1016/j.tate.2011.05.009

Florez, M. A. C. (2001). *Reflective teaching practice in adult ESL settings. ERIC Digest*. Washington: National Center for ESL Literacy Education.

Gaudin, C., & Chaliès, S. (2015). Video viewing in teacher education and professional development: A literature review. *Educational Research Review, 16*, 41-67. https://doi.org/10.1016/j.edurev.2015.06.001

Ghapanchi, Z., & Baradaran, M. (2012). English language teachers' Perceptions of educational supervision in relation to their professional development: A case study of Iran. In *The 1st. Conference on Language Learning & Teaching: An Interdisciplinary Approach (LLT-IA)* (pp. 1-21).

Glatthorn, A. (1995). Teacher development. In L. Anderson (Ed.), *International Encyclopedia of Teaching and Teachers Education* (pp. 41-57). London: Pergamon.

Glickman, C. D., & Tamashiro, R. T. (1980). Clarifying teachers' beliefs about discipline. *Educational Leadership, 37*(6), 459-464.

Glickman, C. D., Gordon, S. P., & Gordon, J. M. (1995). *Supervision of instruction: A developmental approach (3rd ed.)*. Boston: Allyn & Bacon.

Glickman, C. D., Gordon, S. P., & Ross-Gordon, J. M. (2004). *Supervision and instructional leadership: A development approach* (6th ed.). Boston: Allyn and Bacon.

Goldhammer, R. (1969). *Clinical supervision: Special methods for the supervision of teachers*. New York: Holt McDougal.

Gonen, S. I. K. (2016). A study on reflective reciprocal peer coaching for pre-service teachers: Change in reflectivity. *Journal of Education and Training Studies*, *4*(7), 221-235. https://doi.org/10.11114/jets.v4i7.1452

Gosling, D. (2002) Models of peer observation of teaching. Keynote address at the *LTSNGC Peer Observation of Teaching Conference*, Birmingham, 29 May.

Hamilton, E. R. (2013). His ideas are in my head: Peer-to-peer teacher observations as professional development. *Professional Development in Education*, *39*(1), 42-64. https://doi.org/10.1080/19415257.2012.726202

Hammersley-Fletcher, L., & Orsmond, P. (2004). Evaluating our peers: is peer observation a meaningful process? *Studies in Higher Education*, *29*(4), 489-503. https://doi.org/10.1080/0307507042000236380

Hammersley-Fletcher, L., & Orsmond, P. (2005). Reflecting on reflective practices within peer observation. *Studies in higher education*, *30*(2), 213-224. https://doi.org/10.1080/03075070500043358

Hendry, G. D., Bell, A., & Thomson, K. (2014). Learning by observing a peer's teaching situation. *International Journal for Academic Development*, *19*(4), 318-329. https://doi.org/10.1080/1360144X.2013.848806

Hişmanoğlu, M., & Hişmanoğlu, S. (2010). English language teachers' perceptions of educational supervision in relation to their professional development: A case study of Northern Cyprus. *Novitas-ROYAL (Research on Youth and Language), 4*(1), 16- 34.

Hockly, N. (2018). Video-based observation in language teacher education. *ELT Journal*, *72*(3), 329-335. https://doi.org/10.1093/elt/ccy022

Igawa, K. (2009). EFL teachers' views on team-teaching: In the case of Japanese secondary school teachers. *The International Journal of Language Society and Culture, 47*, 145-172.

Inyamah, C. K. (2011). *Educational the 7–point agenda for national development some critical issues*. Paper presented at the 12th annual national conference of National Association for Advancement of Knowledge.

Jo, J., & Woo, S. (2010). Teachers' beliefs on English team teaching in Korean elementary schools. 교육논총, *47*(2), 163-189.

Kaneko-Marques, S. (2015). Reflective teacher supervision through videos of classroom teaching. *Profile Issues in Teachers' Professional Development*. *17*(2). 63-79.

Kayaoğlu, N. (2012). Dictating or facilitating: The supervisory process for language teachers. *Australian Journal of Teacher Education (Online)*, *37*(10), 103-117.

Keig, L., & Waggoner, M. D. (1994). *Collaborative peer review: The role of faculty in improving college teaching. Ashe-Eric higher education report no. 2*. Washington: ASHE-ERIC Higher Education Reports.

Ken-Maduako, I., & Oyatogun, A. T. (2015). Application of team teaching in the English language class. *Journal of Education and Practice*, *6*(29), 178-181.

Kim, Y., & Silver, R. E. (2016). Provoking reflective thinking in post observation conversations. *Journal of Teacher Education*, *67*(3), 203-219. https://doi.org/10.1177/0022487116637120

Lasagabaster, D., & Sierra, J. M. (2011). Classroom observation: desirable conditions established by teachers. *European Journal of Teacher Education*, *34*(4), 449-463. https://doi.org/10.1080/02619768.2011.587113

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge University Press.

Manning, R. C. (1986). Evaluation strategies can be improved with peer observers. *School Administrator*, *43*(1), 14.

Marrow, C. E., Hollyoake, K., Hamer, D., & Kenrick, C. (2002). Clinical supervision using video-conferencing technology: A reflective account. *Journal of Nursing Management*, *10*(5), 275-282. https://doi.org/10.1046/j.1365-2834.2002.00313.x

Martin, G. A., & Double, J. M. (1998). Developing higher education teaching skills through peer observation and collaborative reflection. *Innovations in Education and Training International*, *35*(2), 161-170. https://doi.org/10.1080/1355800980350210

Marwati, B., Wldan, W., & Baehaqi, B. (2019). Implementation of supervisor clinical supervision in increasing the teachers' pedagogic competence of Madrasah Tsanawiyah (MTs) in Masbagik District. *International Journal of Multicultural and Multireligious Understanding*, *6*(4), 374-388.

Masats, D., & Dooly, M. (2011). Rethinking the use of video in teacher education: A holistic approach. *Teaching and Teacher Education*, *27*(7), 1151-1162. https://doi.org/10.1016/j.tate.2011.04.004

Mathew, P., Mathew, P., & Peechattu, P. J. (2017). Reflective practices: A means to teacher development. *Asia Pacific Journal of Contemporary Education and Communication Technology*, *3*(1), 126-131.

McLaughlin, M. W., & Talbert, J. E. (2006). *Building school-based teacher learning communities: Professional strategies to improve student achievement* (Vol. 45). New York: Teachers College.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.

Moradi, K., Sepehrifar, S., & Khadiv, T. P. (2014). Exploring Iranian EFL teachers' perceptions on supervision. *Procedia-Social and Behavioral Sciences, 98*, 1214- 1223. https://doi.org/10.1016/j.sbspro.2014.03.536

Motallebzadeh, K., Hosseinnia, M., & Domskey, J. G. H. (2017). Peer observation: A key factor to improve Iranian EFL teachers' professional development. *Cogent Education, 4*(1), Article 1277456. https://doi.org/10.1080/2331186X.2016.1277456

Ng, M. L. (2015). Difficulties with team teaching in Hong Kong kindergartens. *ELT Journal*, *69*(2), 188-197. https://doi.org/10.1093/elt/ccu057

Opfer, V. D., & Pedder, D. (2011). Conceptualizing teacher professional learning. *Review of Educational Research*, *81*(3), 376-407. https://doi.org/10.3102/0034654311413609

Paba, L. A. (2017). Clinical supervision, a proposal for ensuring the effectiveness of English language teaching at public universities in Colombia. *English Language Teaching, 10*(9), 171-180. https://doi.org/10.5539/elt.v10n9p171

Quesada Pacheco, A. (2005). Reflective teaching and its impact on foreign language teaching. *Revista Electrónica Actualidades Investigativas en Educación, 5*, 1–19.

Peyton, J. K. (1997). *Professional development of foreign language teachers.* Washington: ERIC Clearinghouse.

Phelps, D. L. (2013). *Supervisee experiences of corrective feedback in clinical supervision: A consensual qualitative research study* (Doctoral dissertation). ProQuest Dissertations and Theses Database. (UMI No. 3493809)

Prieto, L. P., Magnuson, P., Dillenbourg, P., & Saar, M. (2020). Reflection for action: Designing tools to support teacher reflection on everyday evidence. *Technology, Pedagogy and Education, 29*(3), 279-295. https://doi.org/10.1080/1475939X.2020.1762721

Race, P., & Fellows, L. B. T. (2009). *Using peer observation to enhance teaching*. Leeds: Met.

Rao, Z., & Chen, H. (2020). Teachers' perceptions of difficulties in team teaching between local-and native-English-speaking teachers in EFL teaching. *Journal of Multilingual and Multicultural Development*, 1-15. https://doi.org/10.1080/01434632.2019.1620753

Richards, J. C., & Farrell, T. S. C. (2005). *Professional development for language teachers: Strategies for teacher learning*. Ernst Klett Sprachen.

Richards, J. C., & Lockhart, C. (1994). R*eflective teaching in second language classrooms*. Cambridge: Cambridge University.

Sari, F. D., Kusmayanti, I. N., Iskandar, R. F., & Utami, A. R. I. (2013). Team teaching challenge: Professionalism or personality issue? In *Proceedings of the International Conference on Psychology in Health, Educational, Social, and Organizational Settings*.

Schon, D. (1983). *The reflective practitioner: How professionals think in action.* New York: Basic Books.

Setyaningrahayu, I., Widhi, P. R., & Murtisari, E. T. (2019). The use of video-based reflection to facilitate pre-service English teachers' self-reflection. *Indonesian JELT: Indonesian Journal of English Language Teaching, 14*(2), 147-165. https://doi.org/10.25170/ijelt.v14i2.1443

Shortland, S. (2004). Peer observation: A tool for staff development or compliance? *Journal of further and higher education, 28*(2), 219-228. https://doi.org/10.1080/0309877042000206778

Simons, M., Baeten, M., & Vanhees, C. (2020). Team teaching during field experiences in teacher education: Investigating student teachers' experiences with parallel and sequential teaching. *Journal of Teacher Education, 71*(1), 24-40. https://doi.org/10.1177/0022487118789064

Solis, M., Vaughn, S., Swanson, E., & Mcculley, L. (2012). Collaborative models of instruction: The empirical foundations of inclusion and co-teaching. *Psychology in the Schools, 49*(5), 498-510. https://doi.org/10.1002/pits.21606

Suhrheinrich, J., & Chan, J. (2017). Exploring the effect of immediate video feedback on coaching. *Journal of Special Education Technology, 32*(1), 47-53. https://doi.org/10.1177/0162643416681163

Tezcan-Unal, B. (2018). Action research on a collegial model of peer observations. *Educational Action Research, 26*(4), 641-654. https://doi.org/10.1080/09650792.2017.1358199

Thomas, T. (2008). Fixing teacher evaluation. *Journal of Educational Leadership, 66*(2), 32-37.

Thompson, M., & Schademan, A. (2019). Gaining fluency: Five practices that mediate effective co-teaching between pre-service and mentor teachers. *Teaching and Teacher Education, 86*, 102903. https://doi.org/10.1016/j.tate.2019.102903

Tim, S. (2018). Team teaching collaborations: Contact, conflict, and empowerment. *Jacet Journal, 62*, 29-47. https://doi.org/10.32234/jacetjournal.62.0_29

Tosriadi, T., Asib, A., Marmanto, S., & Azizah, U. A. (2018). Peer observation as a means to develop teachers' professionalism. *International Journal of Multicultural and Multireligious Understanding, 5*(3), 151-158.

Tsoulou, G. (2016). Attitudes towards conducting peer observation for teacher development purposes within the Greek state school teaching context. *Research Papers in Language Teaching and Learning, 7*(1), 7.

Ur, P. (1996). *A Course for language teaching*. Cambridge: CUP.

Villegas-Reimers, E. (2003). *Teacher professional development: an international review of the literature*. Paris: International Institute for Educational Planning.

Wong, M. W., & Pow, J. W. (2012). Initiating small class teaching in Hong Kong: Video reflective narratives and the professional developmental learning model. *Teacher Development, 16*(4), 507-522. https://doi.org/10.1080/13664530.2012.730713

Yin, R. K. (2014). *Case study research: Design and methods (applied social research methods)*. New Castle: Sage.

Yücel, B., Arman, D., & Yapar, D. P. (2016). Through the looking glass: Video coaching. *ELT Research Journal, 5*(2).

Yürekli, A. (2013). The six-category intervention analysis: A classroom observation reference. *ELT Journal*, *67*(3), 302-312. https://doi.org/10.1093/elt/ccs102