



**A Corpus-driven Scoping Systematic Review of  
Four Decades of Teacher Professional Development Research:  
Exploring Research Foci, Content Areas, Designs Methods and Trends**

**Farzaneh Arjmand**

*Department of English Language & Literature, Imam Khomeini International University. Qazvin, Iran.*

[Farzaneh826@yahoo.com](mailto:Farzaneh826@yahoo.com)

**Mohammad Bagher Shabani** (Corresponding Author)

*Department of English Language & Literature, Imam Khomeini International University. Qazvin, Iran.*

[mshabani@hum.ikiu.ac.ir](mailto:mshabani@hum.ikiu.ac.ir)

**Reza Khani**

*Department of English Language & Literature, Ilam University. Ilam, Iran.*

[r.khany@ilam.ac.ir](mailto:r.khany@ilam.ac.ir)

**Abbas Ali Zarei**

*Department of English Language & Literature, Imam Khomeini International University. Qazvin, Iran.*

[a.zarei@hum.ikiu.ac.ir](mailto:a.zarei@hum.ikiu.ac.ir)

**ARTICLE INFO:**

Document Type: **Research Paper**

Received date:

**2021.09.02**

Accepted date:

**2021.11.24**

Print ISSN: **2251-7995**

Online ISSN: **2676-6876**

**Keywords:**

Teacher professional development, systematic review, content areas; research methods, data collection procedures, research findings.

**Abstract**

This study represents the findings of a systematic review (SR) of literature in the teacher professional development (TPD) domain to outline the research patterns through content examination of 199 research articles (RAs) in the area of TPD over the previous 40 years (1982 -2021). RAs were investigated and their research content areas, utilized research methods, data collection procedures, and findings were analyzed and coded. The broad investigation of the RAs showed a wide variety of themes that corresponded to 22 research areas. ‘TPD program effects’, ‘TPD & technology’, and ‘TPD & Sociolinguistics’ were the most searched content areas. It was also found that the qualitative method with 52.26% of occurrences appeared to be the dominant research method used in RAs. Exploring data collection procedures, it was uncovered that interview, questionnaire and observation were the main data collection strategies utilized within the TPD RAs. Analyzing the findings, ‘changes in teacher practices, attitudes and knowledge’, ‘learner achievements’, and ‘determining priorities for TPD programs’ were the most reported findings in TPD RAs. This corpus-driven SR underpins the notion that TPD makes a difference in altering teachers’ practices and attitudes and improves learner abilities if specific characteristics are taken into account in the planning and administration of TPD programs.

DOI: 10.22034/ELT.2021.47750.2435

Citation: Arjmand, F., Shabani, M. B., Khani, R., Zarei, A. A., (2021). A Corpus-driven Scoping Systematic Review of Four Decades of Teacher Professional Development Research: Exploring Research Foci, Content Areas, Designs Methods and Trends. *Journal of English Language Teaching and Learning*, 13(28), 39-59. Doi: 10.22034/ELT.2021.47750.2435

## 1. Introduction

Participating in a determined course in order to accomplish a lifetime profession does not suffice anymore. This applies to most careers, especially teaching. Many teachers have agreed upon this reality and many of them in developed countries attempt to take part in various types of professional development (PD) programs (OECD, 2016), with organizations providing large supplies on scheduling and implementing TPD courses (e.g., Swedish Ministry of Education, 2012; U.S. Department of Education, 2014). As a consequence of emerging developments of technology and innovations that are accessible for present-day learners, their expectations, demands, and requirements transform as well. To fulfill these altered demands of pupils, teachers need to bring up to date their professional information, attitudes, and practices of their disciplines that are conceivable by progressive professional education and development. Professionalism has been spoken of with various phraseology including 'teacher development', 'professional development', 'professional education', 'professional learning groups', 'training', and 'in-service education'. TPD can be achieved through different ways, from unofficial, transitory, and distinctive practices like inspecting special articles to involvement in wide-ranging and detailed teacher development courses arranged widely. As TPD can be considered as an instrument to develop theoretical knowledge and experiences of teachers, Teacher professional development is an eminent facet in teacher professionalism and teaching experiences (Lee, 2010).

Glattorn (1995) considers professional development as "the professional growth a teacher attains as a result of achieving increased experience and inspecting his or her teaching systematically" (p. 46). From another point of view, INSET applies to the professional and educational actions in which teachers participate, and are mostly proposed by a supervising originator like the ministry of education (August & Calderón, 2006). INSET and PD have been often differentiated in the related articles. Training concentrates on instant and transient aims while development focuses on the teachers' long-term growth (Gardner, 1995). In spite of the fact that these two terms have been differentiated in the literature, they have a lot of overlap. With the purpose of professionally developing themselves, it's necessary for teachers to participate in programs that contribute to the accomplishment of short as well as long-term purposes.

TPD attempts are based on the belief that they assist teachers to develop their competence and enhance teaching abilities. Although this is indicated that TPD can promote positive transformations (Jacob & McGovern, 2015), this is not incontrovertibly beneficial in all situations (Hattie, 2009). Rather, the achievement of TPD programs can vary intrinsically between domains, depending on diverse factors, including the attributes of TPD interventions and the characteristics of individual schools and teachers (OECD, 2016). However, school leaders and policymakers are required to certify that effective TPD opportunities are available for teachers (Lipowsky, 2014). The focus of this SR will be EFL teachers' professional development (TPD). In order to investigate the current position of EFL teachers' PD, the present SR of literature is very well-timed. The purpose of this SR is:

- To discover the previous research articles investigating TPD of EFL teachers.
- To survey and review the effects and suggestions of previous research.

- To provide suggestions for subsequent research in EFL teachers' TPD domain.

To accomplish the purposes of the present study, three research questions were regarded:

1. Which research Trends and developments emerge from SR of RAs on TPD?
2. Which content areas have gone under investigation in TPD RAs over 40 years (from 1982 to 2021)?
3. Which trends are uncovered in research methods, data collection procedures, and findings of TPD RAs?

## 2. Literature Review

Teachers are encountered by many challenges in the contemporary educational environment. Technological developments and expanding worldwide network necessitate that nowadays' instructors be sympathetic issue resolvers competent of valuing and recognizing different viewpoints (Adamson, Astrand, & Darling-Hammond, 2016; Darling-Hammond, 2015; Reardon, 2013). Besides, significant gaps in academic achievement persist among socio-economic communities (Boix Mansilla & Jackson, 2013), and the global force towards privatization could lead to fewer students with access to quality education (Darling-Hammond, 2015; Reardon, 2013). It is therefore essential that TPD expand the abilities, standards and orientations essential to meet the needs of educational situations in the 21st century.

What EFL teachers are, what they need to know, and what they should be able to do are the factors needed to be embedded well in EFL teachers' professional development programs. We perceive that the expression "professional development" has been employed to be used in a vast area of complicated, connected, official, and unofficial opportunities of teacher learning (Firestone & Mangin, 2014). For instance, Kriewaldt (2008) applies expressions like professional development and professional learning which are closely and widely connected, while others (e.g. Lieberman & Miller, 2014) discriminate them. In the present study, we consider Mayer and Lloyd (2011) to define professional development in a way that involves the designed practices which teachers participate in to enrich their sources of knowledge and practice, whereas professional learning results in their classroom practice alterations. The alterations are likely a consequence of professional development programs, but they may also be the result of the peripheral learning which occurs in daily classroom situations. Accordingly, to go beyond the constraints of single studies, there is a need to integrate the research literature on TPD.

TPD is an extensive expression and is used variably in the literature. In the present study, TPD appoints face-to-face, determined, organized, and formalized training and/or learning occasions for teachers. As is broadly acknowledged - for instance, by August & Calderón (2006) and Lipowsky & Rzejak (2015), attributed to Goldschmidt & Phelps, (2007) and Kirkpatrick (1979) - TPD adequacy could be determined by the following elements: Teachers' acknowledgment and delight with TPD programs, teacher education (advances in information, inspiration, and convictions) and teachers' practices in the classroom as well as student's learning (Guskey, 2000).

Lipowsky & Rzejak (2015) and Wade (1984) suggested that there is an uninterrupted relation among successful TPD, promoted instructor's information, upgraded instructing, and

developed learner accomplishment. The association of the four mentioned elements is, though, neither linear nor straight (Yoon, Duncan, Lee, Scarloss, & Shapley, 2007). For example, changes in teachers' thoughts may be the reason or the result of enhanced learner performance (Hattie, 2009), and teacher's gratification in TPD projects does not inevitably results in improvements in teacher's knowledge (Lipowsky & Rzejak, 2015). Numerous researchers are capable of evaluating TPD program effects in different aspects on students and teachers. Guskey (2000), for example, mentioned an effect size of  $d = 0.71$  in a reported meta-analysis.

Within the domain of language teaching, it has been acknowledged that intensive TPD is an important common element in effective programs (Birman, Desimone, Porter, & Garet, 2000). It is of paramount importance to declare that several elements concussively impress teacher training results. In addition to the addressed program characteristics, these components may involve the attributes of the teachers who are participating in the program, coordinators, institution context, and the interaction between these aspects (Lipowsky, 2014). A majority of researchers inferred a variety of frequent characteristics of beneficial teacher training programs by evaluating TPD in various disciplines. They could be classified into characteristics that pertain to the program's frame and structure (structural aspects), the program's content (content-based aspects), and the method in which the contents are transformed (didactic aspects). One of the relevant structural factors is the TPD program's time period. In spite of the fact that there's no straight connection between a TPD program's duration and its accomplishments, longer training periods appear to be essential to transform and expand teachers' attitudes, professional knowledge, and determined classroom exercises (Lipowsky, 2014). Consistency with other learning schemes and opportunities (Müller & Papenkort, 2013; Zehetmeier, 2010); connectivity with participants' own needs, interests, and experiences (Biancarosa & Snow, 2006; Cheung & Slavin, 2012) and the experts' engagement were also important structural characteristics of TPD interventions (Lipowsky & Rzejak, 2015).

Regarding content-related characteristics, efficient TPD can be identified by focusing on students' learning practices and the effectual teaching of subject matter (Ingvarson, Meiers, & Beavis, 2005). Opportunities for active learning process (Müller & Papenkort, 2013) and collaboration among teachers (Birman et al., 2000; Lipowsky & Rzejak, 2015), as enhanced by, for instance, cooperative participation are significant didactic features for efficient TPD. Furthermore, successful TPD is intently associated with the individual teacher's practices; this also contains reflection and connects input phases, the pursuance of recent information in classroom setting, and times of reflecting on creative activities (Lipowsky & Rzejak, 2015).

In some countries, there seems to be limited research analyzing TPD projects with the aim of the information, attitudes, and capabilities which teachers require to be able of helping learners in achieving language competence in their subject matters. The United States can be an exception, where many researchers have evaluated the same interventions. Several researchers collected features of the related literature and, therefore, determined a lot of constituents of efficient TPD in this domain and in various disciplines (August & Calderón, 2006; Knight & Wiseman, 2006). After all, to date, no SR which investigates TPD within the scope of language over the educational curriculum is known.

In consideration of the necessity to assist the professionalization of teachers, a majority of studies have offered essential characteristics of high-quality teacher TPD (e.g., Russell, Kleiman, Carey, & Douglas, 2009). A few years ago, many researchers in the field reached a consensus upon five critical characteristics of TPD programs (e.g., Borko, Jacob, & Koelner, 2010; Darling-Hammond & McLaghlin, 1995). These five characteristics stressed the issue that teachers collaborate with colleagues (*collective participation*), that the TPD encompass several sessions lasting over a long time period (*duration*), that it offers teachers chances to involve in their classroom-related tasks actively (*active learning*), that the TPD topics concentrate on a mix of subject-specific pedagogy and contents (*content focus*), and be associated to teachers' attitudes, knowledge and policy standards (*coherence*).

However, this agreement has been conflicted recently. Firstly, researchers (Jacob, Hill, & Cory, 2017; Kennedy, 2016) have discussed that the mentioned characteristics' evidence mainly comes from local studies and does not focus on the sustainability questions. Secondly, current research has demonstrated that TPD projects including the essential characteristics addressed by the research literature explain varying impacts on student's success (Dede, Jass Ketelhut, Whitehouse, Breit, & McCloskey, 2008; Hill, Beisiegel, & Jacob, 2013). Thirdly, even though similar terminology has been utilized in the literature for TPD, they are defined and applied in various ways (Kennedy, 2016). Indeed, in spite of the fact that the studies on TPD have considerably promoted our information in the domain, there have been comments which declared that the research community has not supplied the type of apparent structures which are necessary to direct TPD contributions (Garet, Heppen, Walters, Smith, & Yang, 2016; Guskey, 2014b; Scher & O'Reilly, 2009; Sztajn, Borko, & Smith, 2017).

Therefore, more research is needed, and research focusing on characteristics of TPD which result in encouraging outcomes are specifically called for (Desimone & Garet, 2015; Hill et al., 2013; Lindvall, Helenus, & Wiberg, 2018). The priority outstandingly addresses the key characteristics of content focus (Firestone, Mangin, Cecilia Martinez, & Polovsky, 2005) and coherence (Kennedy, 2016). In recent years, meta-analyses and reviews have been published that have analyzed the impact of TPD programs related to their content focus (e.g., Cobb & Jackson, 2011; Salinas, 2010; Sztajn, Campbell, & Yoon, 2009).

### 3. Methodology

It is apparent that, like primary research, SRs differ in their breadth, research questions, overarching approach, and data expanding those in the purpose of compiling findings and "pile-up" the evidence from homogenous researches to access greater particularity in their empirical reports, to those intending to find various kinds of studies related to each other to draw a bigger picture (Sandelowski, Voils, Leeman, & Crandell, 2012). Regarding the claim that there is not a general concept map in TPD literature, we discuss that, considering TPD, gathering piles of studies on "what has been done" in the TPD area with qualitative explanations of how and leading to what consequences is a necessity. Accordingly, our primary purpose is to find what aspects of TPD have gone under research during 40 years; what methods and data collection procedures have been utilized and what are the research findings accessed.

Furthermore, the review procedure was carried out to discover the manifestation of TPD in the research domain of teacher education, and also to draw a picture of the aspects and foci of

TPD gone under research during 40 years. This has much correspondences with meta-narrative approach (Greenhalgh, Robert, Macfarlane, Bate, Kyrikidou, & Peacock, 2005) or thematic synthesis (Thomas, Harden, & Newman, 2012), which involves categorizing and collecting main notions and/or results from the researches composed as a systematic chart (Gough, 2007). The systematic map produced in the present study both supplies a description of the various notions of TPD which were (or were not) frequent in the literature on teacher education (as a product) as well as synthesizing the paper results considering TPD (as a tool). In the following parts, a review methodology is outlined which includes: literature search, search terms, search strategies, study corpus selection, inclusion and exclusion criteria, quality and relevance appraisal, codification, and data analysis procedures.

### 3.1. Literature Search

The study begins with a literature search in which we referred to eight databases totally (Elsevier, Sage, Taylor & Francis, Oxford Publications, De Gruyter, Springer, Wiley, and Google Scholar) which have collected worldwide educational studies, searching for RAs in English peer-reviewed journals related to our domain of interest. The time criterion considered for inclusion was from 1982 until 2021. Compiling copies of the combined results of the eight databases, an initial pool including 3268 RAs was retrieved. Review articles, book chapters, theses, comments as well as RAs written in languages other than English were kept out of the review. Table 1 illustrates the number of RAs retrieved from each database in the initial search separately.

**Table 1.** Number of articles according to their place of publication

Databases	Number of Articles				
	1982-1991	1992-2001	2002-2011	2012-2021	total
Elsevier	5	38	96	229	368
Sage	17	93	168	199	477
Taylor & Francis	108	321	575	811	1485
Oxford Publications	0	6	2	3	11
De Gruyter	0	0	7	6	13
Springer	3	27	57	156	242
Wiley	21	36	111	174	342
Google Scholar	26	51	97	183	357
total	180	572	1113	1761	3268

#### 3.1.1. Search Terms

In order to determine the search terms for the study, a broad academic search was done in TPD literature especially in Language Teacher Professional Development (LTPD) and English Language Teacher Professional Development (LTPD) domains. Definitions, concepts, and synonyms related to the aforementioned concepts have been accumulated and the following string has been selected as the most important terms related to TPD. It is crystal clear that there are plenty of terms in the TPD area, however, they are excluded for the reason of the width of the domain, for instance, the terms related to education, in general, have not been included and the study has been limited to TPD in the English Language Teaching domain. The search strings used to retrieve the related RAs in this study included: “Professional Development”, “Teacher Professional Development”, “Language Teacher Professional Development”,

“English Language Teacher Professional Development”, “EFL Teacher Professional Development”, “ELT Professional Development”, “In-service Teacher Professional Development Program”, and “Language Teacher Career Development”.

### 3.1.2. Search Strategies

In order to prepare an answerable pool of data, two main points have to be considered, first determining databases and resources from which the articles have to be retrieved and second, the search strategies utilized. Eight databases including Elsevier, Sage, Taylor & Francis, Oxford Publications, De Gruyter, Springer, Wiley, and Google Scholar were selected that have collected worldwide educational studies in the “Language Teaching” domain. In the advanced search part of the databases, the filters including searching just in the resources’ titles, in the time period of 1982-2021 have been activated. Besides, in the advanced search part, just RAs have been selected determining the type of publication in the present study.

### 3.2. Study Corpus Selection

In the second phase, *Study Corpus selection*, the 3268 RAs were subjected to a selection process. RA’s titles, abstracts, and keywords have been examined in order to find those RAs that were related to the purpose of the present study. The inclusion criterion in the following step was that the articles have to clearly be concerned with TPD for English teachers. RAs concentrating on, for instance, development and educational programs for pre-service teachers, or expanded systemic interventions (with no special focus on TPD), were not included in the sample. Applying the second part of the inclusion criteria provided 234 research articles. RAs which in some way researched TPD (or synonym terms) related to English teachers entered into the next level, data extraction. TPD concept had to be somewhat detailed and central in the paper.

In the *data extraction* phase, the RAs’ texts published in the mentioned journals were accessed and stored in various files. Each file was specified to a year and included that year’s various issues. Articles were saved with a number showing the publication year, volume, issue, as well as the name of the journal in which the article was published. For example, a RA coded as “2011.5.3.1.System” was published in 2011, while it is the first RA in the fifth volume and the third issue of System journal. Identifying the cases of analyses’ incompatibility was the merit of the used coding strategy. It needs to be mentioned that the analysis team included four analysts, two authors of the present study, and two applied linguistics’ Ph.D. candidates, analyzed the RAs’ texts, and classified research variables. The team spent seven two hour meetings in order to explain the aims of the study, discussing the procedure of identifying content areas, research methods, and designs, data collection procedures and findings elaborating on the applicable strategies for classifying the data, and finally conducting a pilot analysis to assure that the analyst team attempted a compatible methodology for the goals of the study.

In the process of analyses, the researchers had to determine the content areas, research methods, and designs, data collection procedures, and findings of every RA on a word page including the paper’s code. In the last meeting, the researchers had to deliver the classified data to the head researchers. RAs were stored in four decades groups: 1982-1991, 1992-2001, 2001-2011, and 2012-2021. After analyzing the RAs of each decade set, the researchers delivered

the results to the head researchers. Head researchers reconsidered the delivered data and conducted a problem-solving meeting with the analysts to remove the problem. The discussion sessions continued till they reached an agreed decision.

### 3.2.1. Inclusion and Exclusion Criteria

The following selection criteria were developed for study inclusion and exclusion suggested by the relevant literature: The RAs have to (a) focus on TPD in English language education; Studies with a focus on TPD in other fields of study were excluded from the corpus; (b) have been published from 1982 to 2021; (c) and be written in English; Further, d) we limited our review to only include scholarly RAs and exclude dissertations, reports, editorial reviews, book chapters and book reviews.; Finally, e) we included RAs from eight aforementioned databases including Elsevier, Sage, Taylor & Francis, Oxford Publications, De Gruyter, Springer, Wiley and Google Scholar as the main sources of most international educational research in the “Language Teaching” domain. RAs out of these eight databases were not selected to be included in the present study. Table 2 summarizes the Exclusion and Inclusion Criteria for the selection of RAs’ included in this study:

**Table 2.** Exclusion and Inclusion Criteria

Criteria	Inclusion	Exclusion
The Study focus	TPD in English language education	TPD in other fields of study
Publication Period	from 1982 to 2021	Prior to 1982 and after 2021
Publication Type	RAs	dissertations, reports, editorial reviews, books, book chapters and book reviews
Databases	Elsevier, Sage, Taylor & Francis, Oxford Publications, De Gruyter, Springer, Wiley and Google Scholar	Databases other than Elsevier, Sage, Taylor & Francis, Oxford Publications, De Gruyter, Springer, Wiley and Google Scholar
Language	written in English	written in languages other than English

### 3.2.2. Quality and Relevance Appraisal

After appraising each study against inclusion and exclusion criteria, the remaining 234 RAs went into the evaluation process based on Gough’s (2007) framework called the weight of evidence (WoE). Applying Gough’s framework, each study was evaluated based on three basic areas: methodological quality, methodological relevance, and topic relevance. In order to expedite the evaluation process, a rating of adequate and inadequate was established. Methodological quality evaluation attempted to measure the extent to which studies included sufficient breadth and clarity on the following areas: the purpose of studies, sampling, data analysis procedure, data collection instruments, and so on. Methodological relevance included evaluating the relevance and quality of chosen and employed designs in determining the generic structure of RAs. Moreover, concerning the topic relevance, researchers determined the degree to which the studies included adequate findings on the generic structure of RAs. Appraising the studies against the WoE framework resulted in 199 studies. Thirteen studies did not satisfy the quality criteria, while 22 studies were considered inadequate in methodological relevance evaluation. There is an illustration of the steps of the literature search as well as the number of identified, assessed, excluded, and retrieved articles in Figure 1 (conform to Moher, Liberati, Tetzlaff, Altman, & the PRISMA Group, 2009, p. 3)

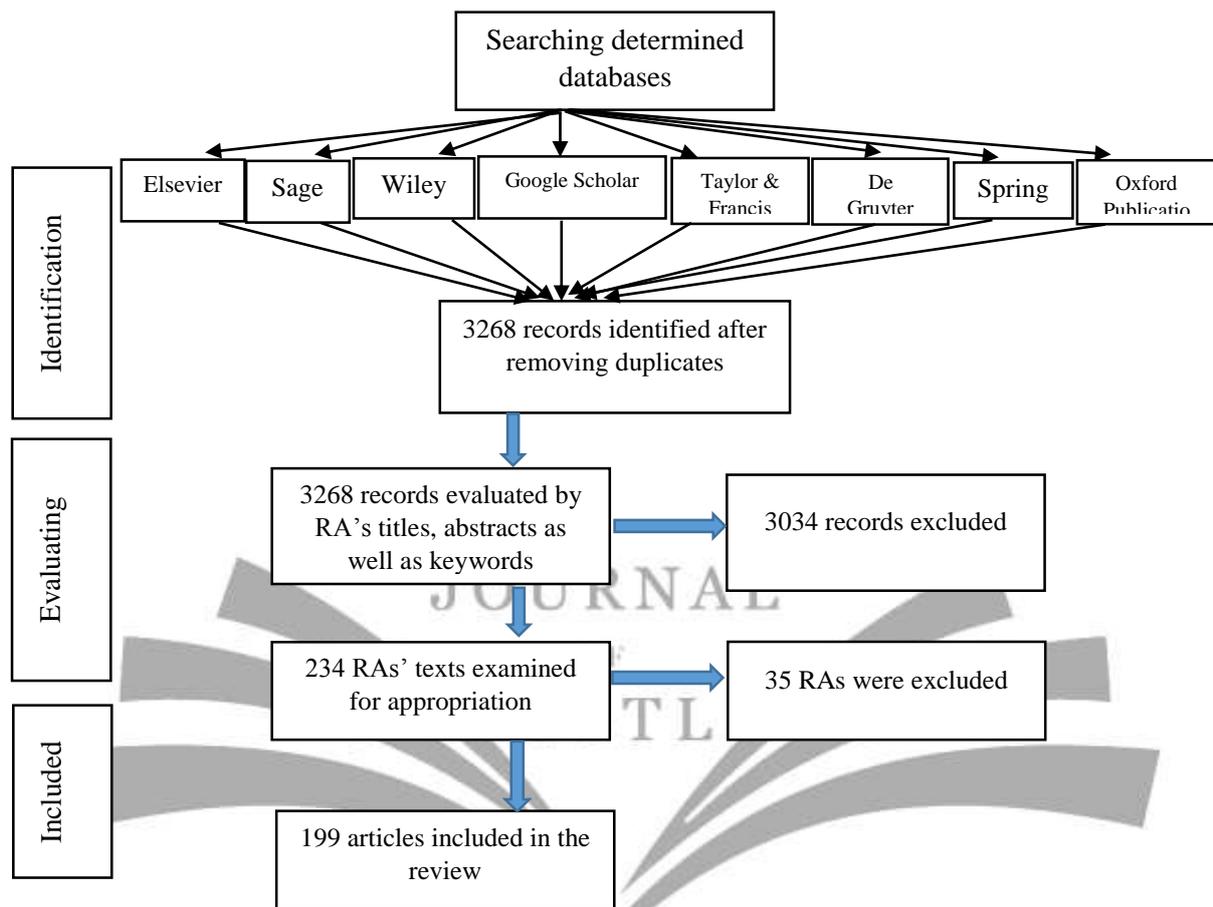


Figure 1. Steps of the literature search as well as the number of identified, assessed, excluded, and retrieved articles

### 3.3. Codification and Data Analysis Procedures

In the codifying phase, the codes which RAs were analyzed according to were determined as content areas, methods of data collection, data collection procedures as well as findings. In the next phase, information was extracted from the RAs considering their content areas, research methods, data collection procedures, and their findings related to the study's general picture. To find the RAs' content areas, the researchers analyzed the papers' topics, abstracts, and keywords. Through reviewing these sections, the appropriate content areas related to the papers were determined. If decision-making was difficult based on analyzing these sections or there was more than one area of study discerned in these sections, other parts of the RAs were analyzed too. The variables' definition parts and the studies' implications helped in finding the content areas a lot. In order to determine the RAs' methodology in the TPD area, Gao, Li, and Lu's (2001) framework was applied in which the research categories and subcategories were divided under two main titles: Empirical (quantitative and qualitative) and non-empirical (Table 3).

**Table 3.** Research types by Gao et al. (2001)

Research types		
Empirical		Non-empirical
<b>Quantitative</b>	<b>Qualitative</b>	
Experimental vs. non-experimental	Text analysis	Theory and implication
Inferential vs. descriptive statistics	Ethnography	Educational operation
Multivariate vs. Univariate	Narrative accounts	Personal attitudes and experiences
	Verbal reports	
	Classroom communication analysis	

Gao et al. (2001) classified qualitative and quantitative studies as empirical because they both involve systematic data gathering and analysis. Consequently, non-empirical studies include studies that do not apply first-hand data. In line with this categorization, Henning (1986) defines quantitative research as the process of “the tallying, manipulation, or systematic aggregation of quantities of data” (p. 702). Accordingly, inferential as well as descriptive statistics can be used in quantitative research or it can test hypotheses in experimental studies. Henning (1986) also believes that qualitative or non-quantitative research are data-driven articles that present research results qualitatively. Qualitative research, as stated by Gao et al. (2001, p. 7), involves exploring formal characteristics of discourse (text analysis); meticulous analysis of interaction in a domain (ethnography); the collection of life experiences or everyday events (narrative reports); content-focused narratives, discussions or interviews (verbal descriptions) and analyzing and coding classroom communications (classroom communication investigations).

Non-empirical research, however, includes reporting theoretical topics as well as their instructional significance (theory and implication); representing meticulous descriptions of instructional issues like curriculum, syllabus, test, software and etc. (instructional operation); collecting personal perspectives and incidents in a study (personal experience and attitudes). Finally, mixed-methods studies are research projects including features of quantitative and qualitative research collectively. According to this categorization reported from Gao et al. (2001) and the propositions suggested by Henning (1986), methodological features of the RAs were determined as Quantitative, Qualitative, Mixed-methods, and Non-empirical research studies. In the next stage of the study, the data collection strategies applied in the RAs were reviewed, coded, and listed in detail.

In some papers, more than one data collection procedure may be used. The overall technique frequencies were tallied and reported from the highest to the lowest. In order to get access to the last part of the data entailed for the purposes of this study, abstract, results, and conclusion parts of the RAs were scrutinized. The main findings of each article were identified and categorized. Just like the first depiction, the process was an inductive one, projecting the configurative analysis method. Numerous coding cycles were applied to the data and the findings were thematize under some predetermined categories including (a) student consequences (b) teachers' pedagogical activities (c) teachers' attitudes and (d) teachers' attendance in high-quality TPD or under the appearing categories from the data. In the results

section of the final report, we will elaborate on the classifications with instances from the papers.

## 4. Results and Discussion

### 4.1 The Number of RAs and Years of publication

Analyzing the published RAs in the mentioned databases showed changes in the number of RAs published during different decades. Therefore, the number of RAs during 40 years of TPD articles was assessed (Table 4).

**Table 4.** Number of RAs according to their years of publication

	Years of publication				total
	1982-1991	1992-2001	2002-2011	2012-2021	
Number of articles	2	3	36	158	199

As presented in Table 4, the number of RAs published in the area of TPD has been increasing over the decades. The number of RAs increased from 2 in 1982-1991 to 3, 36, and 158 over the next three decades. As presented in Table 4, the number of RAs is almost the same in the first two decades, however, it experiences a sharply rising trend from 3 during 1992-2001 to 36 during 2002-2011 and finally, it reaches the pick in the fourth decade from 36 during 2002-2011 to 158 in 2012-2021. It is implied that TPD has received a central focus in recent years. The availability of technology can be a notable reason.

### 4.2 Content Areas in RAs

Findings of this review revealed that 22 content areas have gone under investigation over 40 years of TPD journals' publications. Frequencies of these content areas in different decades are presented in Table 5. It needs to be declared that the percentages of occurrences are also presented in Table 5 according to the number of articles analyzed and coded over every decade. For example, 19.44% for 'TPD program effects' is computed from 36 articles over the third decade. The total percentages and frequencies of content areas in 199 RAs are displayed in the last column. In order to further describe the data in Table 5, the content areas in the center of consideration in each decade are displayed first, and the overall frequency of the content areas in all the RAs are demonstrated subsequently.

**Table 5.** Percentages and frequency of content areas in RAs

Content areas	Publication Years				
	1982-1991	1992-2001	2002-2011	2012-2021	total
TPD program effects	-	-	7(19.44)	23(14.55)	30(15.07)
TPD & Technology	-	-	2(5.55)	27(17.08)	29(14.57)
TPD & Sociolinguistics	-	1(33.33)	7(19.44)	19(12.02)	27(13.56)
TPD programs analysis	-	-	2(5.55)	16(10.12)	18(9.04)
TPD & Need analysis	1(50)	1(33.33)	5(13.88)	9(5.69)	16(8.04)
TPD & Teachers' Perceptions	-	-	1(2.77)	13(8.22)	14(7.07)
TPD & curriculum development	-	1(33.33)	4(11.11)	4(2.53)	9(4.52)
TPD models	-	-	2(5.55)	6(3.79)	8(4.02)
TPD & reflective teaching	1(50)	-	-	5(3.16)	6(3.01)
TPD & Skills Teaching	-	-	1(2.77)	5(3.16)	6(3.01)

TPD & teaching practices	-	-	1(2.77)	5(3.16)	6(3.01)
TPD & Teacher Identity	-	-	-	6(3.79)	6(3.01)
TPD & Teacher research	-	-	3(8.33)	2(1.26)	5(2.51)
TPD & Teacher's self- efficacy	-	-	-	5(3.16)	5(2.51)
Factors affecting TPD	-	-	-	4(2.53)	4(2.01)
TPD & Psycholinguistics	-	-	1(2.77)	2(1.26)	3(1.50)
TPD & Critical Pedagogy	-	-	-	2(1.26)	2(1.00)
TPD & discourse analysis	-	-	-	1(0.63)	1(0.50)
TPD & Assessment	-	-	-	1(0.63)	1(0.50)
TPD & Teacher Motivation	-	-	-	1(0.63)	1(0.50)
TPD & lesson study	-	-	-	1(0.63)	1(0.50)
TPD & pragmatics instruction	-	-	-	1(0.63)	1(0.50)
total	2(100)	3(100)	36(100)	158(100)	199(100)

Table 5 depicts the topic variability investigated in the domain of TPD over various decades. It is revealed that in 2 RAs of 1982-1991 in the field of TPD, 2 content areas went under investigation each of which covered (50%) of the researched topics in the mentioned decade. These two content areas included 'TPD & Need analysis' (50%) and 'TPD & reflective teaching' (50%). In the second decade (1992-2001), the number of published articles was not much more than the published articles in the previous decade. However, the content areas were different including 'TPD & Sociolinguistics' (33.33%), 'TPD & Need analysis' (33.33%) and 'TPD & curriculum development' (33.33%). 'TPD & Need analysis' was the same topic published in the two decades.

In the third decade (2001-2011), the observed variability in the content areas (12 content areas) was more than two previous decades. Research on 'TPD program effects' (19.44%) and 'TPD & Sociolinguistics' (19.44%) were the most investigated content areas which accounted for 38.88% of articles in this decade. 'TPD & Need analysis' (13.88%), 'TPD & curriculum development' (11.11%), and 'TPD & Teacher research' (8.33%) were also dominant in this decade's in the publications. As it is prevalent in the data, 'TPD & Need analysis' and 'TPD & curriculum development' are the content areas going under research from the previous decades. Between 2012 and 2021, 22 content areas were investigated of which 'TPD & Technology' (17.08%), 'TPD program effects' (14.55%), 'TPD & Sociolinguistics' (12.02%), 'TPD programs analysis' (10.12%), 'TPD & Teachers' Perceptions' (8.22%), and 'TPD & Need analysis' (5.69%) were the most investigated ones. The diverse and broad scope of content areas reported in the journals in the last decade indicates the increasing interest of researchers in TPD areas. According to Pica (2003) concurrent with the developments of research on language, teacher, and education, the TPD area expanded significantly. Consequently, researchers tried to investigate new theories related to the TPD projects and achieved new consequences in the area. Therefore, the investigated topics diverged noticeably and explored multiple issues from diverse areas of study like psycholinguistics, sociolinguistics, linguistics, and so on.

‘TPD program effects’ (15.07%) was the most frequently investigated content area among the total content areas. ‘TPD & Technology’ (14.57%), ‘TPD & Sociolinguistics’ (13.56%), ‘TPD programs analysis’ (9.04%), and ‘TPD & Need analysis’ (8.04%) were other dominant content areas studied. These content areas reflect the determined domains of inquiry that include a bunch of subareas for research. Generally, ‘the TPD program effects’ were the dominant area of investigation, and ‘TPD & discourse analysis’, ‘TPD & Assessment’, ‘TPD & Teacher Motivation’, ‘TPD & lesson study’, and ‘TPD & pragmatics instruction’ are the least frequent content areas of study. It was also found that areas including ‘factors affecting TPD’, ‘TPD & Psycholinguistics’, and ‘TPD & Critical Pedagogy’ were not frequently searched contents areas in RAs. An inquiry into the content areas signifies that the focus of concentration in the field in the recent decades is on investigating the effects of TPD programs and the role of technology in TPD programs which is a correspondence of the new era of technology and the interconnectedness of sociology factors in TPD programs, encouraging collaborative and cooperative aspects into TPD programs.

### 4.3 Research Methods in the RAs

Over the 40 years of publications on TPD, quantitative, qualitative, mixed-methods, and non-empirical research methods have been used in the published RAs. The frequencies and percentages of using research methods in every decade of publication are presented in Table 6.

**Table 6.** Frequencies and percentages of research methods used over 40 years of publications on TPD

Research Methods	Years of publication				
	1982-1991	1992-2001	2002-2011	2012-2021	total
<b>Non-empirical</b>	1(50)	1(33.33)	1(2.77)	20(12.65)	23(11.55)
<b>Quantitative</b>	-	-	7(19.44)	24(15.18)	31(15.57)
<b>Qualitative</b>	1(50)	2(66.66)	21(58.33)	80(50.63)	104(52.26)
<b>Mixed-methods</b>	-	-	7(19.44)	34(21.51)	41(20.6)
<b>total</b>	2(100)	3(100)	36(100)	158(100)	199(100)

As shown in Table 6, in the first decade of our interest (1982-1991), non-empirical (50%) and qualitative (50%) methods have been the deployed methods in the articles. There were no quantitative and quantitative and mixed-method studies done on TPD in this decade. The studies applying non-empirical (33.33%) and qualitative (66.66%) methods were continuing to be done in the TPD area in the second decade with different percentages. The expanding range of qualitative studies (58.33%) seemed noteworthy in this decade. This tendency was perceptible in the third decade (2002-2011). 19.44% of the 36 articles published over this decade applied quantitative and another 19.44% employed mixed-methods.

Applying the non-empirical method in the third decade was the same as the previous two decades in number but considering the total number of articles published in this decade, the percentage of non-empirical articles decreased. The expanding application of quantitative and qualitative methods in this decade is indicating a methodological consciousness in researchers’ style (Gao et al., 2001). Over the last decade (2012-2021), the increase in applying mixed-methods was prominent. About 21.51% of the published RAs in this decade applied mixed-

methods. However, the most published papers reported in this data have used the qualitative method (50.63%). The quantitative method accounted for 15.18% of the papers and the non-empirical method covered 12.65% of the published articles in this decade.

#### 4.4 Data Collection Procedures in the RAs

Table 7 depicts the percentages and frequencies of data collection procedures in the RAs during 40 years. The reviewing process of 199 articles displayed that a total of 32 types of data collection procedures have been applied in the articles with the total number of 366 times. It is worth mentioning that in some articles two or more types of data collection procedures have been implemented.

**Table 7.** Data collection procedures used in RAs of TPD

Data Collection Procedures	Years of publication				
	1982-1991	1992-2001	2002-2011	2012-2021	total
interview	-	1(50)	16(25)	74(24.74)	91(24.86)
questionnaire	-	-	16(25)	65(21.73)	81(22.13)
observation	-	-	10(15.62)	47(15.71)	57(15.57)
discussion	-	1(50)	5(7.81)	21(7.02)	27(7.37)
reflection	-	-	6(9.37)	18(6.02)	24(6.55)
journals	-	-	2(3.12)	15(5.01)	17(4.64)
test	-	-	3(4.68)	10(3.34)	13(3.55)
field notes	-	-	-	10(3.34)	10(2.73)
narratives	1(100)	-	3(4.68)	4(1.33)	8(2.18)
lesson plan	-	-	1(15.62)	6(2.00)	7(1.91)
self-evaluation	-	-	-	3(1.00)	3(0.81)
teaching portfolio	-	-	-	2(0.66)	2(0.54)
Teacher report	-	-	1(15.62)	1(0.33)	2(0.54)
stimulated recall	-	-	-	2(0.66)	2(0.54)
discussion	-	-	-	-	-
diaries	-	-	-	2(0.66)	2(0.54)
comments	-	-	-	2(0.66)	2(0.54)
blog	-	-	-	2(0.66)	2(0.54)
written assignments	-	-	-	2(0.66)	2(0.54)
document analysis	-	-	-	2(0.66)	2(0.54)
learning artifact	-	-	-	1(0.33)	1(0.27)
teachers' written account	-	-	-	1(0.33)	1(0.27)
Instructional Aids	-	-	-	1(0.33)	1(0.27)
website statistics	-	-	-	1(0.33)	1(0.27)
handouts	-	-	-	1(0.33)	1(0.27)
oral and written feedback	-	-	1(15.62)	-	1(0.27)
presentations	-	-	-	1(0.33)	1(0.27)
scoring rubric	-	-	-	1(0.33)	1(0.27)
chat interactions	-	-	-	1(0.33)	1(0.27)
analytical memos	-	-	-	1(0.33)	1(0.27)
extensive annotations	-	-	-	1(0.33)	1(0.27)
email correspondence	-	-	-	1(0.33)	1(0.27)
Q-sorting	-	-	-	1(0.33)	1(0.27)
total	1(100)	2(100)	64(100)	299(100)	366(100)

Of the total incidences of data collection procedures, interview with 91 (24.86%) occurrences was by a great amount the most frequently applied data collection procedure in articles. Questionnaire was used in 81(22.13%) articles and observation was also used in 57 (15.57%) articles. These three procedures accounted for 62.56% of the overall incidences of the data collection procedures in the articles of TPD during 40 years. Discussion 27(7.37%), reflection 24(6.55%), and journals 17(4.64%) were the other most frequently applied data collection procedures used in RAs.

To further address the occurrences of data collection procedures applied in the RAs, it should be noticed that interview, discussion and narratives have been the procedures used in three decades and questionnaire, observation, reflection, journals, tests, lesson plan, and Teacher reports have been used in two recent decades. One noteworthy point observed in the data displayed in Table 7 is an expanding range in the data collection procedures' number in the years of publication (1 procedure in the first decade, 2 in the second decade, 11 in the third decade, and finally 31 in the last decade). In the last decade (2012-2021), the variety of data collection procedures employed in RAs in the TPD area has reached its peak. The variety increase in the data collection procedures in the RAs may be related to the increase in the availability of data sources and participants. Additionally, it indicates that attempts have been applied to supply validity and reliability rates for the techniques and strategies of data collection.

#### 4.5 Findings Reported in the RAs

Findings of our analysis revealed that 2\ finding categories with the total occurrences of 240 times resulted from 199 RAs in the 40 years of publications in TPD field. Incidences of these findings categories over diverse decades of publication are displayed in Table 7. It is worth mentioning that in some articles two or more types of finding categories have been reported.

**Table 7.** Findings reported in RAs of TPD

Findings	Years of publication				
	1982-1991	1992-2001	2002-2011	2012-2021	total
changes in teacher practices	1(33.33)	-	8(17.02)	34(18.18)	43(17.91)
changes in teacher attitudes	1(33.33)	-	7(14.89)	18(9.62)	26(10.83)
strengthen teachers' knowledge	-	-	3(6.38)	18(9.62)	19(7.91)
learner achievements	-	-	3(6.38)	16(8.55)	19(7.91)
determining priorities for TPD programs	-	-	3(6.38)	16(8.55)	19(7.91)
enhancing professional growth	-	-	3(6.38)	14(7.48)	17(7.08)
encouraging collaborative TPD	-	1(33.33)	7(14.89)	7(3.74)	15(6.25)
increasing teachers' ability to use technology	-	-	1(2.12)	9(4.81)	10(4.16)
enhancing teacher skills	-	-	3(6.38)	6(3.20)	9(3.75)
improving teacher self-efficacy	-	-	1(2.12)	7(3.74)	8(3.33)
TPD program not promising	-	-	1(2.12)	7(3.74)	8(3.33)

enhancing autonomous learning	-	-	-	7(3.74)	7(2.91)
promoting reflective competence	1(33.33)	-	1(2.12)	4(2.13)	6(2.5)
improving TPD programs	-	1(33.33)	-	5(2.67)	6(2.5)
enhancing curriculum development skills	-	1(33.33)	1(2.12)	3(1.60)	5(2.08)
increasing self-awareness	-	-	-	5(2.67)	5(2.08)
developing useful strategies	-	-	2(4.25)	3(1.60)	5(2.08)
presenting model for TPD	-	-	1(2.12)	4(2.13)	5(2.08)
helping teachers play a more active role	-	-	1(2.12)	2(1.06)	3(1.25)
challenges of TPD programs	-	-	-	2(1.06)	2(0.83)
emphasizing importance of TPD	-	-	1(2.12)	-	1(0.41)
total	3(100%)	3(100%)	47(100%)	187(100%)	240(100%)

Among the total occurrences of finding categories, ‘changes in teacher practices’ with 43(17.91%) occurrences was by far the most frequent finding categories reported in 199 RAs in all four decades. ‘Changes in teacher attitudes’ has been reported in 26(10.83%) articles and ‘strengthen teachers’ knowledge’, ‘learner achievements’, and ‘determining priorities for TPD programs’ with the same frequency and percentage were also reported in 19(7.91%) articles. These three finding categories accounted for 23.73% of the overall incidences of the data collection procedures in the articles of TPD during 40 years. ‘Enhancing professional growth’ 17(7.08%), ‘encouraging collaborative TPD’ 15(6.25%), and ‘increasing teachers’ ability to use technology’ 10(4.16%) were the other most frequently finding categories reported in the RAs.

To further address the occurrences of finding categories reported in the RAs, it should be noticed that ‘changes in teacher practices’, ‘changes in teacher attitudes’, ‘encouraging collaborative TPD’, ‘promoting reflective competence’, and ‘enhancing curriculum development skills’ have been the finding categories reported in three decades and ‘strengthen teachers’ knowledge’, ‘learner achievements’, ‘determining priorities for TPD programs’, ‘enhancing professional growth’, ‘encouraging collaborative TPD’, ‘increasing teachers’ ability to use technology’, ‘improving TPD programs’, ‘enhancing teacher skills’, ‘developing useful strategies’, ‘presenting a model for TPD’, ‘helping teachers play a more active role’, ‘improving teacher self-efficacy’, and ‘TPD program not promising’ have been reported in two decades. One indicative point observed in the data displayed in Table 7 can be the fact that the variety of finding categories in the last two decades is almost the same. As it was declared in the content area analysis part, the focal attention in the field in the recent decades has been on investigating the effects of TPD programs, therefore the number of reported finding categories indicating changes in teachers’ practices, attitudes, knowledge, skills, role, reflective competence, self-efficacy, autonomous learning, and learner achievements is explainable.

## 5. Conclusions

As displayed, the purpose of the present study has been determining the content areas, research methods, data collection procedures, and research findings in the TPD research area and their

frequencies manifested in the TPD journals. In spite of the fact that space restrictions did not allow an in-depth investigation of these factors, an endeavor was made to identify the state of the art of the area. Generally, the findings of this study displayed that the TPD research area has encountered noticeable innovations in the specified variables.

We found that different content areas have been studied in TPD RAs. The content areas' overabundance validates the interdisciplinary characteristic of the domain and the broad span of the publications in journals. However, it should be mentioned that the present study has indicated that a significant quantity of articles are associated with the 'TPD program effects', 'technology-related TPD programs', and 'TPD & sociolinguistics section of TPD'. 'TPD program effects' has been the dominant investigated content area in the articles. The study also has confirmed a vast exploration of content areas, and subsequently, variability in research methods, data collection procedures, and findings.

According to the results of the study, the domain is passing through innovative directions from which the ultimate goal of language learning will be achieved. In other words, the variability of research trends and themes is presumably indicating that the developing areas of research have caused complicated issues in the way in which the ultimate goal is obscured. In these incongruous circumstances, it is proposed that a regulatory section evidently describes the research needs and precisely concentrates on the heterogeneous and homogenous results.

Based on the frequency reported in research methods applied in TPD, the articles have transited from non-empirical methods into qualitative methods. In any case, no definitive and certain tendency could be characterized for the research methodologies used in the domain. The high frequency of qualitative studies is not powerful in the way that we can suggest a qualitative foundation for the research studies of the domain. Probably, the most important consequence is that the frequency of research methods during four decades has increased at the same pace for all the categories.

Regarding the data collection procedures used in the RAs, it was demonstrated that interview, questionnaire, and observation were the most commonly used procedures. As it is manifested, direct data collection procedures were at the top. Another characteristic of the studies was applying collective data collection procedures in order to access more reliable data. The use of mixed data collection procedures supplied a deep insight into the topic under investigation (Kirazlar, 2007). Applying questionnaires, interviews, surveys, diaries, and reflective journals in combination permitted researchers to access a deep awareness of participants' experiences in real life in TPD and provided an understanding of instructors considering their professions (Alan, 2003; Daloglu, 2004; Kirazlar, 2007; Ozcalli, 2007; Unal, 2010).

Regarding the TPD studies' findings, it was illustrated that 'changes in teacher classroom practices, attitudes, and knowledge' as well as 'learner achievements' have been the most commonly reported results of RAs of TPD area. Another interesting finding mostly reported was determined priorities for TPD programs from teachers' or other stakeholders' perspectives. These findings are consistent with the results of the first section of the present study reporting the 'TPD program effects' as the most common researched content area in the TPD RAs.

Analyzing and considering the findings of TPD RAs during 40 years can be a hint in organizing efficacious TPD programs for EFL teachers.

On the whole, the present study can be considered as a primary exploration introducing the state of the art of the domain. TPD specialists, readers, and researchers of journals may apply the findings of the present analysis to understand the direction of publishing in the mentioned journals. The finding of our analysis can benefit editorial boards of the journals too. Important data presenting an overall picture of the domain's history has been provided for them. They now know the position of the field and the steps that need to be taken to compensate for underrepresented research points.



## References

- Adamson, F., Astrand, B., & Darling-Hammond, L. (2016). *Global education reform: How privatization and public investment influence education outcomes*. New York: Routledge/Taylor & Francis Group.
- Alan, B. (2003). *Novice Teachers Perceptions of an Inservice Teacher Training Course at Anadolu University*. (Unpublished Thesis). Bilkent University, Ankara, Turkey.
- August, D., & Calderón, M. (2006). Teacher beliefs and professional development. *Review of Educational Research, 73*(2), 555–563.
- Biancarosa, G., & Snow, C. E. (2006). *Reading next—A vision for action and research in middle and high school literacy: A report to Carnegie Corporation of New York*. Washington, DC: Alliance for Excellent Education.
- Birman, B. F., Desimone, L. M., Porter, A. C., & Garet, M. S. (2000). Designing professional development that works. *Educational Leadership, 57*(8), 28–33.
- Boix Mansilla, V., & Jackson, A. (2013). *Educating for global competence: Learning redefined for an interconnected world*. New York: Solution Tree.
- Borko, H., Jacobs, J., & Koellner, K. (2010). Contemporary approaches to teacher professional development. *International encyclopedia of education, 7*(2), 548–556.
- Cheung, A. C. K., & Slavin, R. E. (2012). Effective reading programs for Spanish-dominant English language learners (ELLs) in the elementary grades: A synthesis of research. *Review of Educational Research, 82*(4), 351–395. doi:10.3102/0034654312465472
- Cobb, P., & Jackson, K. (2011). Towards an empirically grounded theory of action for improving the quality of mathematics teaching at scale. *Mathematics Teacher Education and Development, 13*(1), 6–33.
- Daloglu, A (2004). A professional development program for primary school English language teachers in Turkey: Designing a materials bank. *International Journal of Educational Development, 26*(2), 677- 690.
- Darling-Hammond, L., & McLaughlin, M. W. (1995). Policies that support professional development in an era of reform. *Phi Delta Kappan, 76*(8), 597–604.
- Darling-Hammond, L. (2015). *The flat world and education*. New York: Teachers College Press.
- Dede, C., Jass Ketelhut, D., Whitehouse, P., Breit, L., & McCloskey, E. M. (2008). A research agenda for online teacher professional development. *Journal of Teacher Education, 60*(1), 8–19.
- Desimone, L. M., & Garet, M. S. (2015). Best practices in teachers' professional development in the United States. *Psychology, Society, & Education, 7*(3), 252–263.
- Firestone, W. A., Mangin, M. M., Cecilia Martinez, M., & Polovsky, T. (2005). Leading coherent professional development: A comparison of three districts. *Educational Administration Quarterly, 41*(3), 413–448.
- Firestone, W. A., & Mangin, M. M. (2014). Leading professional learning in districts with a student learning culture. In L. E. Martin, S. Kragler, D. J. Quatroche, & K. L. Bauserman (Eds.). *Handbook of professional development in education: Successful models and practices, preK-12* (pp. 319–338). New York, NY: The Guilford Press.
- Gao, Y., Li, L., & Lu, J. (2001). Trends in research methods in applied linguistics: China and the West. *English for Specific Purposes, 20*(1), 1-14.
- Gardner, R. (1995). On-service teacher education. In: L. Anderson (Ed.): *International Encyclopedia of Teaching and Teacher Education*. 2nd Edition. London: Pergamon Press.
- Garet, M. S., Heppen, J., Walters, K., Smith, T., & Yang, R. (2016). *Does content-focused teacher professional development work? Findings from three Institute of Education Sciences studies*. Boston: Houghton Mifflin.
- Glattom, A. (1995). Teacher development. In: L. Anderson (Ed.): *International Encyclopedia of Teaching and Teacher Education*. (pp. 368–92). London: Pergamon Press.

- Goldschmidt, P., & Phelps, G. (2007). Does teacher professional development affect content and pedagogical knowledge: How much and for how long?. *Economics of Education Review*, 29(3), 432-439.
- Gough, D. (2007). Weight of evidence: A framework for the appraisal of the quality and relevance of evidence. *Research Papers in Education*, 22(2), 213-228.
- Greenhalgh, T., Robert, G., Macfarlane, F., Bate, P., Kyriakidou, O., & Peacock, R. (2005). Storylines of research in diffusion of innovation: A meta-narrative approach to systematic review. *Social Science & Medicine*, 61(2), 417-430.
- Guskey, T. R. (2000). *Evaluating professional development*. Thousand Oaks, CA: Corwin Press.
- Guskey, T. R. (2014b). Planning professional learning. *Educational Leadership*, 71(8), 10-16.
- Hattie, J. A. C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London, UK: Routledge.
- Henning, G. (1986). Quantitative methods in language acquisition research. *TESOL Quarterly*, 20(1), 701-708.
- Hill, H. C., Beisiegel, M., & Jacob, R. (2013). Professional development research: Consensus, crossroads, and challenges. *Educational Researcher*, 42(9), 476-487.
- Ingvarson, L., Meiers, M., & Beavis, A. (2005). Factors affecting the impact of professional development programs on teacher's knowledge, practice, student outcomes & efficacy. *Education Policy Analysis Archives*, 13(10), 1-28.
- Jacob, R., Hill, H., & Corey, D. (2017). The impact of a professional development program on teachers' mathematical knowledge for teaching, instruction, and student achievement. *Journal of Research on Educational Effectiveness*, 10(2), 379-407.
- Jacob, A., & McGovern, K. (2015). *The Mirage. Confronting the hard truth about our quest for teacher development*. Brooklyn: TNT.
- Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research*, 86(4), 945-980.
- Kirazlar, C. (2007). *An Investigation into the Professional Development of Primary School English Language Teachers: The Effects of Keeping Diaries on Teacher Reflection and Attitudes towards Teaching*. (Unpublished Master Thesis). Canakkale Onsekiz Mart University, Canakkale.
- Kirkpatrick, D. L. (1979). Techniques for evaluating training programs. *Training and Development Journal*, 33(6), 78-92.
- Knight, S. L., & Wiseman, D. L. (2006). Lessons learned from a research synthesis on the effects of teachers' professional development on culturally diverse students. *Preparing quality educators for English language learners*, 14(3), 81-108.
- Kriewaldt, J. (2008). Research into relationships between teacher professional learning and teaching standards: Reviewing the literature. *Language Teaching Research*, 11(1), 43-62.
- Lee, J. F. (2010). Perceptions of ELT among English language teachers in China. *Education Journal*, 37(1), 137-154.
- Lieberman, A., & Miller, L. (2014). Teachers as professionals: Evolving definitions of staff development. *Handbook of professional development in education*, 3(1), 67-71.
- Lindvall, J., Helenius, O., & Wiberg, M. (2018). Critical features of professional development programs: Comparing content focus and impact of two large-scale programs. *Teaching and Teacher Education*, 70(2), 121-131.
- Lipowsky, F. (2014). Theoretical perspectives and empirical evidence on the effectiveness of teachers' professional development and training. *Education Journal*, 42(2), 511-541.

- Lipowsky, F., & Rzejak, D. (2015). Key features of effective professional development programs for teachers. *Ricercazione*, 7(2), 27–51.
- Mayer, D., & Lloyd, M. (2011). *Professional learning: An introduction to the research literature*. Melbourne: Australian Institute for Teaching and School Leadership.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & the PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, 6(7), 1–6.
- Müller, U., & Papenkort, U. (2013). Didactics for adults and the didactics of training. *Report*, 36(4), 22–31.
- OECD (2016). PISA 2015 results. In Excellence and equity in education (Vol. I) Paris: OECD publishing.
- Ozcalli, S. (2007). *Possible effects of in-service education on EFL teachers' professional development in terms of teacher efficacy and reflective thinking*. MA Thesis. Bogazici University, Istanbul.
- Pica, S. (2003). Second language acquisition research and applied linguistics. *Working Papers in Educational Linguistics*, 18(2), 1-26.
- Reardon, S. F. (2013). The widening income achievement gap. *Educational Leadership*, 70(8), 10-16.
- Russell, M., Kleiman, G., Carey, R., & Douglas, J. (2009). Comparing self-paced and cohort-based online courses for teachers. *Journal of Research on Technology in Education*, 41(4), 443–466.
- Salinas, A. (2010). *Investing in our teachers: What focus of professional development leads to the highest student gains in mathematics achievement?* Coral Gables, FL: University of Miami.
- Sandelowski, M., Voils, C. I., Leeman, J., & Crandell, J. L. (2012). Mapping the mixed methods–mixed research synthesis terrain. *Journal of Mixed Methods Research*, 6(4), 317–331.
- Scher, L., & O'Reilly, F. (2009). Professional development for k–12 math and science teachers: What do we really know?. *Journal of Research on Educational Effectiveness*, 2(3), 209–249.
- Swedish Ministry of Education (2012). *Cabinet decision I: 44*. Stockholm, Sweden: Swedish Ministry of Education.
- Sztajn, P., Campbell, M., & Yoon, K. S. (2009). Conceptualizing professional development in mathematics: Elements of a model. *Journal of Research on Educational Effectiveness*, 2(4), 311-324.
- Sztajn, P., Borko, H., & Smith, T. M. (2017). Research on mathematics professional development. In J. Cai. (Ed.). *Compendium for research in mathematics education* (pp. 793–823). Reston, VA: National Council of Teachers of Mathematics.
- Thomas, J., Harden, A., & Newman, M. (2012). *Synthesis: combining results systematically and appropriately*. California: Sage Publications.
- Unal, D. (2010). *Designing an In-service Teacher Training Program for English Language Instructors and Identifying the Effectiveness of the Program: An Implication at Afyon Kocatepe University*. (Unpublished Thesis). Gazi University, Ankara, Turkey.
- U.S. Department of Education (2014). *Fiscal year 2014 budget summary and background information*. Washington, DC: U.S. Department of Education.
- Wade, R. K. (1984). What makes a difference in in-service teacher education? A meta-analysis of research. *Educational Leadership*, 42(4), 48–54.
- Yoon, K. S., Duncan, T., Lee, S. W. Y., Scarloss, B., & Shapley, K. L. (2007). Reviewing the evidence on how teacher professional development affects student achievement: issues & answers. *Regional Educational Laboratory Southwest*, 21(3), 112-125.
- Zehetmeier, S. (2010). Action research in teacher education. *Working Papers in Educational Linguistics*, 13(2), 197-211.