



The effect of efficient professional development on the teaching of code-focused skills in beginner readers: exploring the impact of professional development intensity and coaching on student outcomes

Verónica Martínez¹ · Luis Castejón¹ · Soledad González-Pumariega¹

Accepted: 23 December 2021 / Published online: 10 February 2022
© The Author(s) 2022

Abstract

Professional development (PD) of teachers working with students in the first years of learning to read is a privileged way of preventing initial reading difficulties and its effects in the long and short term. This research studies the effects of PD in student reading performance, although the results are not conclusive with regard to which PD format is more adequate. The objective of this study is to determine which modality (face-to-face with a coach vs. without a coach) and intensity (number of contact hours) of PD are more efficient in achieving fluency improvement in student performance in the code-focussed skills in the first years of learning to read. Both pre-schoolers and their teachers took part in the study with a quasi-experimental pre-test post-test design. The experimental group (n=71) was provided literacy instruction from teachers (n=8) who received 40 h of face-to-face training with a coach and the control group (n=29) was provided literacy instruction from teachers (n=8) who received only 8 h of initial training (without coaching). The results showed significant intra-group improvements with a large reduction in students at risk for reading difficulties. No significant differences were obtained between groups in reading performance. This suggests greater efficiency in a lower intensity format of PD without a coach in the development of code-focussed skills. The study considers the need to adequately assess reading ability in the light of attitude and motivation of teachers as variables which influence the efficiency of PD.

Keywords Professional development · Reading · Code-focussed skills · Pre-schooler · Coaching · Intensity

✉ Verónica Martínez
martinezveronica@uniovi.es

Extended author information available on the last page of the article

Introduction

Research into professional development (PD) shows that it is relevant as a means of promoting the teaching of reading based on research among teachers (Connor et al., 2014; Hudson et al., 2021). The impact of PD on student outcomes has also been established. The theory proposed by Desimone (2009) considers a sequence of actions, according to which the PD develops the teacher's knowledge and skills, thus leading to a change in attitudes and finally impacting on student outcomes. In fact, a fundamental criterion of quality of proposed PD is its effect on student performance (Biancarosa et al., 2010; Didion et al., 2020; Kennedy, 2016; McCutchen et al., 2002; Paige et al., 2019; Yoon et al., 2007). In accordance with this concept, the study of PD should be based on a dynamic model, like the Lattice Model (Connor, 2016) which includes the determining factor of the adaptation of teaching to the level of the student in specific implied processes (texts, linguistic socio-emotional and cognitive processes).

Since the 1980s, when Shulman (1986) identified professional knowledge as the “missing paradigm”, a growing body of research has addressed instruction and teacher knowledge. Teachers are currently considered to be the most important factor in shaping student achievement (Darling-Hammond et al., 2017). From this perspective, it is necessary to study PD in greater depth so as to offer training proposals to those leaders responsible for educational training. The up-dating of teachers may facilitate the adaptation of instruction to student level, especially with early reading difficulties (Brownell et al., 2017; Connor et al., 2014; Landry et al., 2011; Lonigan et al., 2011; Scanlon et al., 2008; Scarparolo & Hammond, 2018). Students cannot learn what their teacher does not know, and this is a major contributor to poor reading skill outcomes, especially in the first stages of learning (Brady et al., 2009; Pfof et al., 2014; Scanlon et al., 2008). For example, Lane et al. (2008) showed that a student from the 2nd year of primary school, whose teachers had acquired greater professional knowledge of fluency, obtained higher scores in reading precision and speed.

Improvement in PD for teachers in early primary education is, to a certain extent, urgent, since the persistence of initial reading difficulties may lead to negative consequences in the short, medium and long term (low academic performance, abandonment, lack of professional qualifications, delinquency, and so on) (Connor et al., 2014) which are difficult to make up for (Pfof et al., 2014). It has been shown that the reading level in third grade of primary education predicts both low performance in different knowledge areas as well as early abandonment (Lesnick et al., 2010).

Generally, educational research considers three categories of teachers' professional knowledge (Evens et al., 2018; McCutchen et al., 2002; Moats, 2009; Shulman, 1986). Content knowledge (CK) is the knowledge of the teacher about the specific subject which they teach. In the area of reading, CK was identified by The National Reading Panel (NPR, 2000) and five essential reading components (“Five Big Ideas”) were highlighted. The teacher should have knowledge of phonological awareness, phonics, fluency, vocabulary and comprehension

(Brady et al., 2009; Ehri & Flugman, 2018). Pedagogical knowledge (PK) refers to knowledge of methodologies for the teaching of reading, for example, the proposal for intervention in the classroom put forward by the Response to Intervention Model (RTI) (Scanlon et al., 2008). Pedagogical content knowledge (PCK) refers to an amalgamation which integrates the two previous categories (CK and PK) in practice. This is implemented through a repertoire of answers that the teacher activates in the classroom to construct specific teaching situations in a specific circumstance (Evens et al., 2018).

The PD modality which has been most used is the *traditional conference format* (Connor et al., 2014; Desimone & Garet, 2015; Garet et al., 2008; Kennedy, 2016; Landry et al., 2011). This consists of one-time conferences or workshops where a group of teachers from different schools listen to a 1–3 h lecture on a topic. Research has shown repeatedly that this procedure is not effective in promoting PD in the area of reading (Biancarosa et al., 2010; Garet et al., 2008; McMahan et al., 2019) because it does not deal with the series of actions needed to connect theory and practice (Desimone, 2009). Didion et al. (2020) highlight this: “it is unclear what happens between when teachers receive instruction through PD and the distal outcome of students reading performance” (p. 31).

New proposals of demonstrated effectiveness are needed, but currently, consensus has yet to be reached as to which is the best format of PD to develop high-quality, professional knowledge in the area of reading. Research into PD has identified five core features for implementing an efficient PD format (Desimone, 2009; Desimone & Garet, 2015; Pak et al., 2020). *Intensity*, which refers to the number and distribution of hours, highlights the effectiveness of PD when it is extended over the duration of the academic year and includes 20 or more hours of training. *Coherence* means acting in conjunction with the student, the teacher and context characteristics, for example, starting with that which the teacher knows (CK and PK), does and feels when teaching and learning (PCK) (Osman & Warner, 2020). *Collective participation* refers to *active learning* and *collaboration* among professionals as two characteristics which make PD effective so that PCK may be transformed (Darling-Hammond et al., 2017). Finally, *content focus* refers to how content can influence a PD proposal, for example, each area of the teaching of reading requires a specific PD proposal.

The combination of these five features gives rise to a combination of formats that Darling-Hammond et al. (2017) group together under the title ‘*collaborative learning process*’. Evidence supporting these PD formats comes from transversal and longitudinal studies with reference to different content (sciences, reading, mathematics), carried out frequently in large-scale programmes (Biancarosa et al., 2010; Brady et al., 2009; Landry et al., 2011; Paige et al., 2019). However, some studies and revisions have identified limitations in the effect of collaborative PD based on the five features previously mentioned (Didion et al., 2020; Garet et al., 2008, 2011; Kennedy, 2016; Lonigan et al., 2011), or, on factors linked to learning on the part of teachers, like attitude and motivation (Brady et al., 2009; Osman & Warner, 2020). Kennedy (2016) points out that “a conception that is based on more nuanced understanding of what teachers do, what motivates them, and how they learn and grow” is necessary (p. 30). These aspects related to the

teacher as learner have been frequently ignored in large-scale programmes that relied on intermediaries whose familiarity with teacher learning may have been limited. These may also have reduced the initiative and autonomy of teachers, and this would be minimised in short-term programmes such as in the case of learning communities (McCutchen et al., 2002).

Desimone and Garet (2015) point out that the content focus influences the design of PD. In the area of reading, Didion et al. (2020) have differentiated two interconnected areas in the teaching of reading which could require PD in different formats. The first refers to code-focussed skills (i.e., phonological awareness, phonics, fluency) which allow accuracy at the initiation of learning to read and the gradual acquisition of fluency through reading practice (Castejón et al., 2015). The second refers to meaning-focussed skills (i.e., vocabulary, comprehension) which allow the comprehension of texts and reading to learn. It must be taken into account that both areas are related and sequenced, given that, at advanced levels, comprehension and learning may be affected if code-focussed skills have not been adequately mastered (Connor, 2016).

Given that, in practice, the content focus determines differences in PD for the teaching of reading, the central question in this study is to advance in the knowledge of the most adequate characteristics for promoting instructional abilities implied in the teaching of code-focused skills (Didion et al., 2020; Landry et al., 2011).

In PD in code-focussed skills with teachers in early primary education, diverse results have been obtained with or without coaching (for a wider summary see Hudson et al., 2021). Biancarosa et al. (2010) studied the effects on reading performance of a coaching programme with a large sample. While in the first year, the baseline, the students obtained an improvement of 16%. In the second year, in which coaching was implemented, the improvement in performance was 28%. The results obtained by Landry et al. (2011) were along the same lines. They carried out a longitudinal study with pre-schoolers to prevent difficulties in reading. However, other studies have not found effects on the performance of students in relation to collaborative PD proposals with coaching (Garet et al., 2008; Lonigan et al., 2011). This question is relevant to the design of PD since monitoring with coaching may condition intensity (contact hours) and intensity depends largely on the cost of PD. Collaborative PD proposals of greater intensity and with monitoring (PD with a face-to-face coach) would be more costly than those collaborative proposals which are more discreet, one-off, and spread out over time (PD without coach).

The objective of this research is to study the adequacy of intensity (contact hours) and the type of follow-up (PD with a face-to-face coach vs. PD without a coach) in the area of code-focussed skills. Intensity is specified in the number of hours and whether this is accompanied by the participation and monitoring of a literacy coach. Efficiency is determined by studying the effect of the PD proposal on reading performance (Desimone, 2009; Kennedy, 2016). Within this framework the hypothesis is considered, just as Didion et al. (2020) suggested, that there will not be significant differences in student performance in code-focussed skills where teachers receive more intense PD with a face-to-face coach.

Method

Participants

The initial sample for this type of intentional, non-probabilistic study was made up of 100 students, all of whom attend the third year of Preschool (5 years of age) at three different state schools in a province in the north of Spain. Of the total number of participants in the sample, 71 formed part of the experimental group who was provided literacy instruction from teachers ($n=8$), named the High Intensity Group with a face-to-face coach (High Intensity Group HIG-40) and 29 formed the control group who was provided literacy instruction from teachers ($n=4$), named the Low Intensity Group without a coach (Low Intensity Group LIG-8). The average chronological age of the total of the sample was 65.14 months ($TD=3.545$). Table 1 shows the number of participants in each group according to gender, as well as the mean (M) and the standard deviation (SD) in chronological age. All parents or legal guardians gave their consent to participation in the study.

In each school for the group HIG-40, working groups were formed by five professionals. There were two coordinators (one for each centre) who were specialist teachers in Hearing and Language (known in Spanish as “Audición y Lenguaje-AL”), two teachers of Therapeutic Pedagogy (known in Spanish as “Pedagogía Terapéutica-PT”) (one in each centre), two class teachers (one for each class group and two for each centre) and two external assessors from the University who acted as coaches (one for each centre). In this proposal, the coordinator-AL was a language specialist and supported the teacher in the classroom. Furthermore, this person was responsible for managing group work and for following the four principles (development of teacher knowledge, design of collaborative activities, systematicity and autonomy). The PT was specialised in learning difficulties and supported the work of the teacher in the classroom. The teacher participated in the design and implementation of the activities and constructs specific teaching situations in the classroom in specific circumstances (Evens et al., 2018). The coach (external assessor from university) brought knowledge which comes from research (NRP, 2000); guided all the members of the group in the construction, change and deepening of knowledge; analysed and helped to interpret the results in order to design adequate intervention (help with lesson planning); supervised the meetings about the design of activities; as well as monitoring the effect of

Table 1 Number of participants in each group by gender and chronological age, Mean (M) and Standard Deviation (SD)

| | | Gender | | Chronological age | |
|-------------|----------|--------|-------|-------------------|-------|
| | | Boys | Girls | M | SD |
| HIG-40 (EG) | School A | 9 | 20 | 65.62 | 3.448 |
| | School B | 19 | 23 | 63.45 | 2.634 |
| LIG-8 (CG) | School C | 19 | 10 | 67.10 | 3.726 |

coaching in teaching practice and student outcomes. In the school with the LIG-8 group, the working group was set up in the same way (coordinator-AL, PT and two teachers), but without a coach.

The study was developed in accordance with the code of ethics of the World Medical Association (Declaration of Helsinki) for experiments involving human subjects in research and the Spanish Law for Personal Data Protection (15/1999 and 3/2018) principles.

Instrument

The university assessors elaborated ad hoc a log of six tasks based on Artiles and Jiménez (2011) and on González-Seijas and Cuetos (2008) related to the area of code-focussed skills (i.e., phonological awareness, phonics, RAN, fluency) which were presented in the initial training for the practical development of CK with teachers. The Cronbach alpha is 0.79 suggesting that the items have relatively high internal consistency.

The first task consisted of quickly naming a list of 36 letters written in capital letters (NCL), which included both vowels and consonants. In the second task of phonemic awareness (PMA), the participant was required to identify the initial sound of 10 words, of which two were monosyllabic and eight bi-syllabic (six had a CVCV structure and two had a CCVCV structure). The instruction for this second task was: *What is the first sound in these words?* In the third task, also dealing with phonological awareness (PPA), the participant was asked, on hearing the word, to recognise which word in a group of three began with a different phoneme. The task was made up of 10 groups of three bi-syllabic words, where the five first groups were of the structure CVCV and the last five of the structure CCVCV. In this case the instruction was: *Which of these words begins with a different sound?* The fourth task consisted of reading 20 short, high-frequency, bi-syllabic words (RWO), 17 with the structure CVCV and three bi-syllabic words with the structure VCCV. To select these words a process of discussion was carried out in the working group and curricular vocabulary was taken as a reference. The fifth task consisted of reading 10 pseudowords (RPS), three monosyllabic, six bi-syllabic with the structure CVCV and one with the structure VCV. The sixth task consisted of the rapid naming (RNT) of 20 stimuli (nine black and white images which are repeated) selected by teachers and known by the students (geometric figures, numbers and drawings). The instruction was: *Say the name of these drawings as fast as you can.*

Procedure

First contact: initial training in the traditional conference form

Following the strategic line of the Regional Government Orientation Service, first contact was established with the schools in a training course organised at the beginning of the academic year by the Centre for Teachers and Resources (CTR) in the district. Preschool and early primary teachers from different centres of the province

attended voluntarily and free of charge. The course was titled: *Difficulties in the teaching of reading in schools: towards an inclusive model* and it consisted of a one-time workshop which, for the group LIG-8 was the main training and for the group HIG-40 initial training. This is now described in detail due to its relevance to the present study.

Three university professors explained the initial difficulties in the teaching of reading, relating these to the need to promote professional knowledge (Connor, 2016; Didion et al., 2020; Paige et al., 2019). Evidence collected by the NRP (2000) was presented, highlighting the fact that, after 20 years, this had not been put into practice in the classroom. So, the idea was put forward of the risk to initial readers and the consequences, related to both the possible difficulties for the student and with the difficulties of teaching related to the professional knowledge of teachers (Didion et al., 2020). The difficulties experienced by teachers in practice and the effects on students were analysed on the basis of a number of studies (Castejón et al., 2019; Lane et al., 2008; Lesnick et al., 2010; Scanlon et al., 2008; Scarpapolo & Hammond, 2018). The solution to the problem under consideration was put forward in the PD, which also included an agreement by the teachers to take up the challenge of constructing new practice for teaching code-focussed skills in a collaborative, systematic and explicit way.

Thus, teachers were motivated and challenged to participate in a PD proposal which was based on a RTI and was considered as a preventative, educational intervention developed in the classroom (Tier 1) with the support of one of the speakers as a coach. The experience of Scanlon et al. (2008) in a model for intervention in Tier 1 was explained in detail, given that these authors highlight the value of quality of instruction, autonomy and leadership of teachers.

Following this, the practical proposal was specified by presenting, on both a theoretical and practical level (with results obtained from other previous experiences) (Castejón, et al., 2019), the six tasks which would be used for the initial and final assessments. The details were explained of how the teacher could establish criteria measurements in the classroom. In doing this, the teachers were trained to: (1) differentiate three levels of initial reader performance: at risk, emerging and established (Good et al., 2001) and (2) establish performance objectives for each group (Castejón et al., 2019). Furthermore, once levels were identified, the participating teachers selected a list of five activities for content (i.e., phonological awareness). The initial workshop concluded with the presentation of examples of previous experience, in which a comparison was made of different routes to the learning of code-focussed skills through an analysis of recordings (Castejón et al., 2015). The main objective of this initial training was to create CK, PK, interest, motivation, and commitment on the part of teachers. Lastly, those who attended were invited to participate voluntarily in a PD proposal.

Setting up of working groups: the HIG-40 and the LIG-8

The creation of working groups of participants from the three schools arose from a negotiation of the mode of PD among university professors, teachers, and management teams of the three centres. In each of the three cases the same collaborative

mode of PD for 36 h was offered with monitoring by university coaches. The 4 h of initial training must be added to this (40 h: HIG-40). A schedule of 1 h weekly meetings was established with the coach (a total of 36 h from October to June). The CTR of the district endorsed and financially supported the experience by covering the cost of teaching and materials and also asked each group to supply an initial project as well as a final report which described the results of the proposal, the resources used, the schedule and the minutes from the weekly meetings. These meetings would take place in the school of each working group guided by four principles: (1) Start with what the teacher knows, does and feels about the teaching of reading. (2) Define and propose collaborative solutions in a process of accompaniment based on open dialogue to design activities and promote changes in instructional strategies. (3) Systematicity, without changing the activity before having reached the objective and maintaining the motivation of teachers and students. (4) Autonomy, leaving a space for the adaptation of proposals and the creativity of teacher (McCutchen et al., 2002).

Two schools accepted the PD proposal, but the third, for collateral reasons related to the circumstances of the centre, proposed an adjustment by which the intensity of the training was considerably reduced. In this case it was limited to two initial hours and two final hours in the centre, to which must be added the 4 h of initial training. Thus, the low intensity proposal LIG-8 (PD without a coach) was set up, in which other characteristics of the initial proposal were maintained. The teachers of the centre agreed to meet weekly (without a coach) to revise and adjust the experience and they also agreed to carry out systematic practice, to be autonomous and to take on leadership in the classroom. The coordinator of the group would contact the external assessor when necessary by email (online monitoring) and they also agreed that, on certain dates, they would send the results of the initial and final assessments in an Excel format designed by the external assessor. This assessor would then analyse the results to establish criteria measurements and performance objectives and to determine the levels of each group (*Risk, Emerging and Established*). Also, the assessors would prepare an initial and a final report, which would be incorporated into the overall report. The proposal was accepted in these terms by the participants as a challenge and with a high level of motivation and, so, the LIG-8 constituted the control group to study the effect of intensity variables (40 h vs. 8 h) on the performance of the student in code-focussed skills. Table 2 shows the distribution of intensity in each of the two groups.

In the two modes of assessment (HIG-40 and LIG-8), with the objective of establishing a map of initial reader performance of the class group, three levels of

Table 2 Distribution of intensity (contact hours) in each group

| | Initial workshop (hours) | First contact (hours) | Working groups | Final contact (hours) |
|--------|--------------------------|-----------------------|--|-----------------------|
| LIG-8 | 4 | 2 | Light intensity without a coach | 2 |
| HIG-40 | 4 | 2 | High intensity a face-to-face coach (32 h) | 2 |

achievement were established following the proposal by Good et al. (2001): *Risk*, *Emerging and Established*. In order to do this, the Index of Lexical Fluency (ILF) was used since it is an ideal measure which allows the combination of precision and speed which are the two key elements for measuring the initial level of code-focussed skills in a transparent, orthographic system like Spanish (Castejón et al., 2015). This is the total number of words read correctly divided by the time spent in reading them, and all of this multiplied by 100. In this way the ILF was obtained for each participant in order to establish the criteria for measurement and performance objectives. The mean and the standard deviation of ILF were calculated and it was estimated that those participants whose scores were a standard deviation below the mean would also form part of the risk level. The participants whose scores were a standard deviation above the mean would form the level *Established*. All of those whose scores were found between the two previous values would form the level *Emerging*.

The ILF was calculated for both the measurements on the pre-test and those on the post-test with the objective of establishing whether the students had evolved from one level to another along the continuum. Between both measurements, the teachers designed and carried out systematic activities put forward by the working groups in accordance with the established principles. In the group HIG-40, this was with the support of a face-to-face coach, and, in the group LIG-8, this was autonomous.

Finally, in the last sessions with the experimental group (HIG-40) and the control group (LIG-8) the results of the final assessment were shared to show the effects of the work carried out. Then, the teachers were presented with the patterns of compensation of the initial differences in reading skill acquisition: the Mateo effect, stable performance and compensation pattern (Pfost et al., 2014). The experience of the two groups concluded with an evaluation by the participants with regard to what they had learned from the experience and to its continuation.

Data analysis

The design of this research is quasi-experimental pre-test post-test with two groups that received two modes of assessment: the experimental group (EG) named the High Intensity Group (HIG-40 which is characterised by a higher intensity of weekly support by experts (PD with a face-to-face coach) and the control group (CG) named the Low Intensity Group (LIG-8) with lower intensity and without weekly support by experts (PD without a coach). Statistical analyses have been carried out with the statistical programme SPSS version 25.0 for Mac and the programme G*Power 3.1. With the former, the corresponding descriptive statistics were obtained. Furthermore, a factor analysis and the post-hoc Scheffé test were carried out in order to know if the ILF was useful for establishing a continuum and whether there were differences between the reading levels. Similarly, the Student *t* test was used for independent samples to determine whether there were statistically significant intragroup differences before and after the intervention. The post hoc calculation of the observed power and the effect sizes were obtained with the programme G*Power 3.1 in order to assess the effectiveness of the proposal on each of the variables. Cohen's *d* index (Cohen, 1988) was used

to measure effect size, this being an index pertaining to the family of indexes which measure the size of the treatment effect. This index is also one of the most commonly used measurements for the calculation of statistical power (Ledesma et al., 2008). This establishes small effects when situated around 0.20, medium effects if d is close to 0.50 and large effects when it is situated around 0.80 (Cohen, 1988).

Results

The first section presents the results obtained by the HIG-40 and the LIG-8 in the reading variables before the application of the PD proposal, as well as the diversity of the participants, taking ILF as a reference. Following this, the results are presented after having implemented the proposal in the different variables under study, considering the intra-group and inter-group differences. Finally, a continuum of the distribution of levels of reading in the variables after the PD proposal is presented in order to assess whether the number of readers with possible difficulties have been reduced.

Diversity of the participants before the proposal

As can be observed in Fig. 1, the distribution of the participants, with time (LPT) and the number of correct answers in the reading of words (LPA) as variables, is very heterogeneous in both groups (each point represents a participant). The mean of the words read correctly before implementing the PD proposal is of 15 words in both groups (vertical line), while the mean for time of reading the 20 words, correctly or incorrectly (horizontal line), is 98.18 s in the HIG-40 and 75.73 in the LIG-8. In Fig. 1, the diversity that the teachers had identified as a difficulty in teaching can be clearly observed, given that there was a group of participants who read the 20 words correctly in a short time. These would be on one extreme of the continuum and the other group who read fewer words, and took a longer time to do so, would be on the other extreme of the continuum.

Table 3 shows the results obtained by the HIG-40 and the LIG-8 in all the studied variables in the initial assessment (PRE), the value of the t statistic and p -values. There were no statistically significant differences in any of the studied

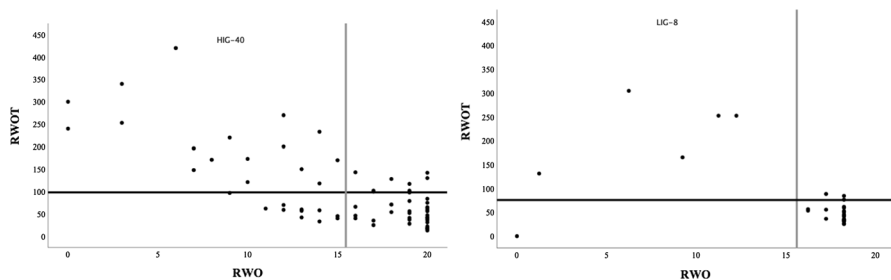


Fig. 1 Distribution of participants of HIG-40 and LIG-8 according to time and correct reading of words in PRE

Table 3 Descriptive statistics, difference in means and *p*-values for HIG-40 and LIG-8 in PRE

| Variable | HIG-40 (n=71) | | LIG-8 (n=29) | | t | <i>p</i> |
|----------|---------------|--------|--------------|--------|--------|----------|
| | M | SD | M | SD | | |
| NCL | 35.12 | 3.157 | 34.55 | 5.378 | 0.824 | 0.412 |
| NCLT | 48.23 | 18.161 | 45.32 | 23.608 | 0.665 | 0.508 |
| PMA | 8.87 | 1.529 | 8.69 | 1.430 | 0.549 | 0.584 |
| PPA | 5.78 | 2.518 | 6.51 | 2.029 | -1.387 | 0.169 |
| RWO | 15.46 | 5.231 | 15.61 | 5.148 | -0.127 | 0.899 |
| RWOT | 98.18 | 83.545 | 75.73 | 74.900 | 1.255 | 0.212 |
| RPS | 7.46 | 3.097 | 8.24 | 2.344 | -1.362 | 0.178 |
| RPST | 37.12 | 27.646 | 40.09 | 36.635 | -0.442 | 0.659 |
| RNT | 27.84 | 9.293 | 28.20 | 8.794 | -0.327 | 0.745 |
| ILF | 33.60 | 32.801 | 34.58 | 21.119 | -0.150 | 0.881 |

NCL Naming capital letters, *NCLT* Naming capital letters time, *PMA* Phonemic awareness, *PPA* Phonological awareness, *RWO* Reading words, *RWOT* Reading words time, *RPS* Reading pseudowords, *RPST* Reading pseudowords time, *RNT* Rapid naming time, *ILF* Index of Lexical Fluency

variables, which leads to the conclusion that the groups were homogenous at the beginning of the proposal.

Before establishing the critical or cut-off values, the extreme values in the ILF were eliminated from the total. Those participants whose ILF was below the percentile 20 (5.294) would directly form part of the Risk level and would not be used to calculate the cut-off point. Once eliminated, the measurement criteria was established from the calculation of the mean (39.58) and the standard deviation (28.807) of the ILF to identify the levels in each class group.

It was estimated that those children whose scores were a standard deviation below the mean ($M-1SD$), would also form part of the Risk level (< 10.78). Those whose scores were a standard deviation above the mean ($M+1SD$), would make up the Established level (> 68.38).

The factor analysis of variance (ANOVA) of a factor between groups, taking ILF as an independent variable in the HIG-40 showed statistically significant differences in the three reading levels $F(2, 68) = 165.561$; $p = 0.000$. The post hoc Scheffé test established significant differences between the Established and Emerging levels ($p < 0.001$), between the Established and Risk levels ($p < 0.001$), and between the Emerging and Risk levels ($p < 0.001$). As can be observed in Fig. 2, the Risk level is the lowest ILF sample and the diversity within this level is lower in both groups, with the distribution of the percentages of participants being 23.94% in the HIG-40 and 20.68% in the LIG-8. The Emerging and Established levels are seen to have greater diversity. The percentage of participants which are to be found in the Emerging level in the HIG-40 is 66.19% and in the LIG-8 it is 72.41%, while the distribution in the Established level in the HIG-40 is 9.85% and in the LIG-8 it is 6.89%.

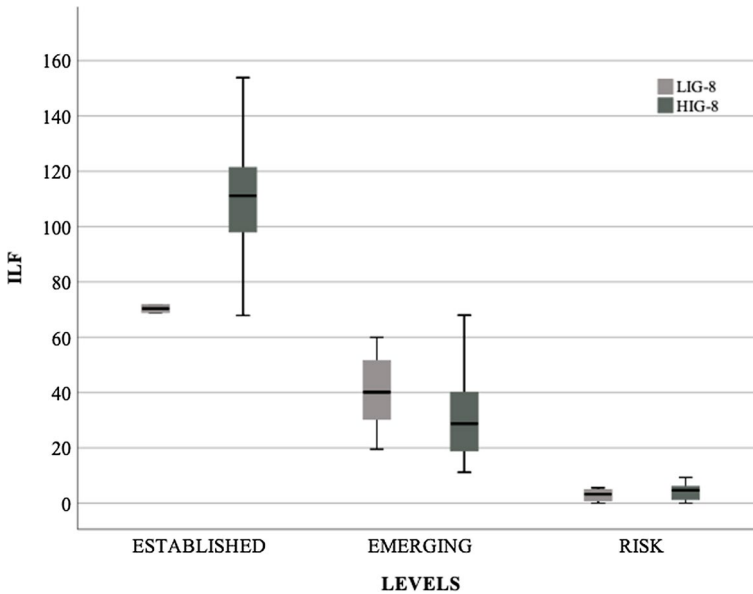


Fig. 2 Box-plot of the three levels within the performance continuum in the PRE

Diversity among participants after the PD proposal

Statistically significant differences can be observed between the pre-test (PRE) and the post-test (POST) intra-group in all the variables (Table 4). In the case of the HIG-40, the level of significance is $p < 0.001$ in all the variables, while in the case

Table 4 Difference in means, observed power and effect sizes (d) intra-groups between PRE and POST

| Variable | HIG-40 (n=71) | | | | LIG-8 (n=29) | | | |
|----------|---------------|------|----------------|-------|--------------|------|----------------|-------|
| | t | p | Observed power | d | t | p | Observed power | d |
| NCL | -4.043 | .000 | 0.999 | 0.479 | -2.199 | .036 | 1.000 | 0.408 |
| NCLT | 8.040 | .000 | 1.000 | 0.954 | 12.412 | .000 | 1.000 | 2.304 |
| PMA | -5.165 | .000 | 0.999 | 0.612 | -5.551 | .000 | 0.999 | 1.048 |
| PPA | -8.850 | .000 | 1.000 | 1.050 | -10.744 | .000 | 1.000 | 1.995 |
| RWO | -6.484 | .000 | 0.999 | 0.769 | -6.089 | .000 | 0.999 | 1.130 |
| RWOT | 6.241 | .000 | 0.999 | 0.740 | 4.022 | .000 | 0.988 | 0.746 |
| RPS | -6.008 | .000 | 0.999 | 0.712 | -4.692 | .000 | 0.998 | 0.871 |
| RPST | 5.286 | .000 | 0.999 | 0.627 | 3.744 | .001 | 0.977 | 0.695 |
| RNT | 5.934 | .000 | 0.999 | 0.704 | 6.415 | .000 | 0.999 | 1.191 |
| ILF | -11.555 | .000 | 1.000 | 1.371 | -7.140 | .000 | 0.999 | 1.325 |

NCL Naming capital letters, *NCLT* Naming capital letters time, *PMA* Phonemic awareness, *PPA* Phonological awareness, *RWO* Reading words, *RWOT* Reading words time, *RPS* Reading pseudowords, *RPST* Reading pseudowords time, *RNT* Rapid naming time, *ILF* Index of Lexical Fluency

of the LIG-8, it is also $p < 0.001$ in all the variables except in NCL ($p = 0.036$) and in LPST ($p = 0.001$). Furthermore, the observed power in all the variables is close to the value 1.000 and the effect size is large. Both groups, in the variable NCL, obtained a smaller effect size in comparison with other variables while in the ILF this is large.

To confirm the existence of statistically significant differences between the HIG-40 and the LIG-8 after implementing the PD proposal, a t test was used on independent samples. The most outstanding result was that, despite differences in intensity and ways of monitoring, in both groups there was a reduction in the initial Risk group, and no significant intergroup differences were found. Table 5 shows the results obtained by the HIG-40 and the LIG-8 in all the variables studied after the application of the PD proposal (POST) and its comparison, the value of the t statistic, its significance, the observed power and the effect size. Significant statistical differences are only observed in PPA ($t = 2.362$; $p = 0.020$), the effect size being medium ($d = 0.554$). Only in the variables NCL and ILF effect size is small ($d = 0.232$, and $d = 0.200$, respectively) and close to reaching this in LPT ($d = 0.186$) and LPST ($d = 0.198$).

As can be observed in Fig. 3, the distribution of the participants, after having implemented the PD proposal and taking time (LPT) and the number of correct answers in the reading of words (LPA) as variables, is less diverse (each point represents a participant) in both groups. In this case, the mean for words that the participants in both groups read correctly increased to 19 words (vertical line) and the mean for time of reading has been reduced in the HIG-40 (44.56 s) and in the LIG-8 (53.15 s) (horizontal line). In Fig. 3 it can be seen that, although there are still participants of the HIG-40 who present difficulties in reading the words correctly in a shorter time, the whole group appears to be more homogenous.

Table 5 Descriptive statistics, difference in means, observed power and effect sizes (d) for EG and CG in POST

| Variable | HIG-40 (n=71) | | LIG-8 (n=29) | | t | p | Observed power | d |
|----------|---------------|--------|--------------|--------|--------|--------|----------------|-------|
| | M | SD | M | SD | | | | |
| NCL | 35.87 | 0.412 | 35.31 | 3.526 | 0.857 | 0.399 | 0.261 | 0.223 |
| NCLT | 37.21 | 12.767 | 36.74 | 24.420 | 0.123 | 0.903 | 0.062 | 0.024 |
| PMA | 9.85 | 0.669 | 9.79 | 0.568 | 0.414 | 0.680 | 0.113 | 0.096 |
| PPA | 8.07 | 1.922 | 9.00 | 1.389 | -2.362 | 0.020* | 0.803 | 0.554 |
| RWO | 19.18 | 1.667 | 19.07 | 2.764 | 0.254 | 0.800 | 0.076 | 0.048 |
| RWOT | 44.56 | 25.031 | 53.15 | 60.292 | -0.742 | 0.464 | 0.210 | 0.186 |
| RPS | 9.62 | 0.947 | 9.62 | 1.347 | 0.004 | 0.997 | 0.054 | 0.008 |
| RPST | 22.30 | 13.563 | 26.75 | 28.741 | -0.799 | 0.430 | 0.0225 | 0.198 |
| RNT | 23.31 | 7.448 | 22.71 | 7.955 | 0.355 | 0.724 | 0.097 | 0.007 |
| ILF | 57.49 | 34.590 | 65.08 | 40.725 | -0.945 | 0.347 | 0.229 | 0.200 |

NCL Naming capital letters, NCLT Naming capital letters time, PMA Phonemic awareness, PPA Phonological awareness, RWO Reading words, RWOT Reading words time, RPS Reading pseudowords, RPST Reading pseudowords time, RNT Rapid naming time, ILF Index of Lexical Fluency

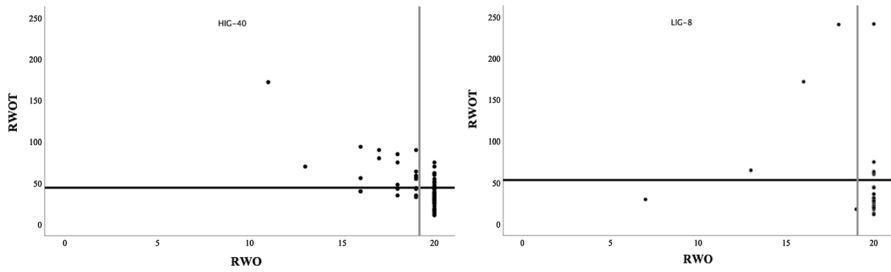


Fig. 3 Distribution of the participants in the HIG-40 and LIG-8 according to time and correct reading of words in the POST

The results of the ANOVA of the variables analysed by level (Fig. 4) within the continuum after the PD proposal showed that there were statistically significant differences in LPA ($F=9.093$; $p<0.001$), in LPT ($F=61.930$; $p<0.001$), in LPS ($F=11.434$; $p<0.001$) and in RNT ($F=5.289$; $p=0.007$). In the variable LPA there were statistically significant differences between the Established and Risk levels ($p=0.005$) and between the Emerging and Risk levels ($p=0.001$). In the LPT there were significant differences between the Established level, the Emerging level ($p=0.001$) and the Risk level ($p<0.001$) and between the Emerging and Risk levels ($p=0.001$). In RPS there were differences between the Established and Risk levels ($p=0.004$) and between the Emerging and Risk levels ($p<0.001$). In the LPST there were significant differences between the Established and Risk levels ($p<0.001$).

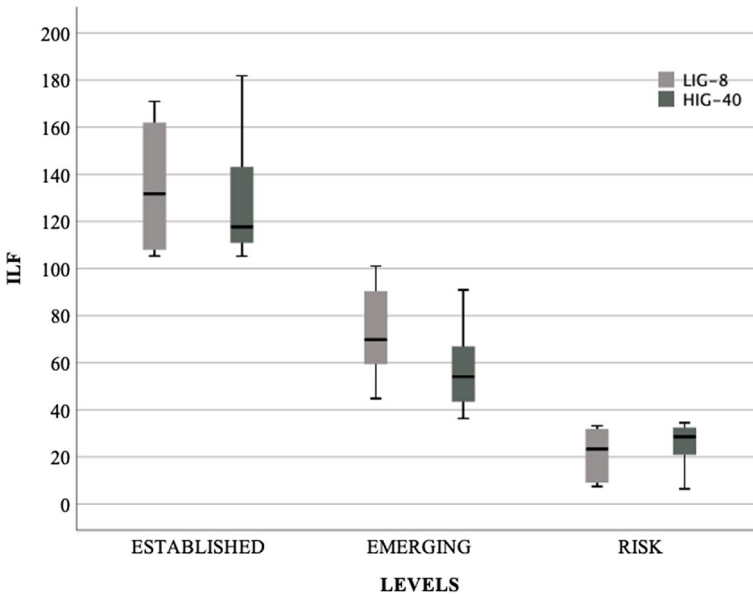


Fig. 4 Box-plot of the three levels within the performance continuum in the POST

Lastly, in RNT there were only significant differences between the Established and Risk levels ($p=0.010$).

The ANOVA of a factor between levels in the variable ILF in the HIG-40 after implementing the PD showed statistically significant differences in the three levels [$F(2, 68)=157.630$; $p=0.000$]. The post-hoc Scheffé test showed significant differences between the Established and Emerging levels ($p<0.001$), between the Established and Risk levels ($p<0.001$), and between the Emerging and Risk levels ($p<0.001$). Also, the ANOVA in the LIG-8 showed statistically significant differences in the three levels of reading [$F(2, 28)=54.664$; $p=0.000$]. In this group, the post-hoc Scheffé test also showed statistically significant inter-levels with a level of significance of $p<0.001$ (Established and Emerging; Established and Risk; Emerging and Risk). As can be observed in Fig. 4, it is the Risk level where above the mean is lower in both groups, the distribution being 26.78% of the participants in the HIG-40 and 31.03% in the LIG-8. In the Emerging and Established levels greater homogeneousness is observed within the levels as is observed in the PRE. Similarly, the percentage of participants who are in the Emerging level in the HIG-40 is 60.56% and in the LIG-8 this is 55.17%, while the distribution in the Established level in the HIG-40 is 12.67% and 13.7% in the LIG-8.

Discussion

This article has studied the effect of two proposals for PD on the teaching of code-focussed skills (Didion et al., 2020). The implemented PD proposals shared initial training with two objectives. In the first place, this was to disseminate CK (“Big Five Ideas”—NRP, 2000) and PK (proposed RTI Model) in a coherent and relevant way for teachers. Secondly, the objective was to promote motivation and a positive attitude towards PD among teachers (Brady et al., 2009; Kennedy, 2016; Osman & Warner, 2020; Scanlon et al., 2008).

From the initial training, an experimental group was formed (EG: HIG-40) and a control group (CG: LIG-8) which presented no significant differences in the pre-test and which made possible the study of the effect on student performance of the two PD formats. These shared some of the five core features which, according to research, are part of high-quality PD (Darling-Hamond et al., 2017; Desimone, 2009; Desimone & Garet, 2015; Pak et al., 2020), but they differed in intensity and the use of a coach. In the two PD groups there was a positive effect on students considered to have possible reading difficulties, but no significant differences were obtained between groups in performance on the tasks carried out. So, it may be affirmed that, for code-focussed skills in the teaching of reading, the LIG-8 format was more efficient since it required less time and resources to obtain similar results.

The results from the experimental group (HIG-40) support those of previous studies which, with large-scale programmes, obtained a significant effect in reducing initial reading difficulty through coaching and high levels of intensity (Erhi & Flugman, 2018; Hudson et al., 2021). Biancarosa et al. (2010), with a hierarchical, crossed-level, value-added-effects model, compared student literacy learning over 3 years, obtained that PD coaching may have significant effects on students’ reading.

Paige et al. (2019) carried out a longitudinal study (PD lasting from 90 to 180 h) and found that with third-grade teachers, there was statistically significant growth in reading knowledge among teachers with generally large effect sizes and also large effects on all student outcomes with gains in year one exceeded in years two and three.

At the same time, considering the comparison between the two groups, the results indicate that to master code-focussed skills the PD proposal of lower intensity and without a coach (LIG-8) obtained similar results and so, is more efficient. This result does partially support descriptions by Garet et al. (2011) and Lonigan et al. (2011) in the sense that they did not find that the incorporation of coaching showed more benefits to student performance. The main result of this study suggests greater efficiency of lower intensity PD without a coach. This indicates that 32 h of difference in intensity between the two proposals means a high cost in time and results. This information could be very useful when offering proposals to those leaders responsible for educational training. However, this result does not align well with other research (Hudson et al., 2021) and deserves some consideration.

Interpretation of the results is not only related to the increase in intensity, since it leads to the consideration of other aspects of the experience which could be relevant for PD. Firstly, the two proposals have been placed in a category named 'the collaborative learning process' and so, the group LIG-8 should not be identified with the traditional conference format (Connor et al., 2014; Desimone & Garet, 2015; Kennedy, 2016), but rather with a different form of collaboration to that used in the group HIG-40. This was conditioned by the difficulties of the center to adopt the proposal with greater initial intensity. Secondly, this reduction entailed the absence of an external university assessor in the group HIG-8. Such an assessor acted as a coach in the group HIG-40. In this sense, the most notable difference between the two groups is the presence or absence of an external assessor who acted as a coach in the group HIG-40. Thirdly, this lack of an external coach in LIG-8 did not mean that these teachers acted alone. They had internal support from the AL group coordinator who supported them in their practice and met them weekly so that a natural form of collaboration was generated which could be identified with a professional learning community (PLC) (McCutchen, 2002).

Therefore, the experience of the control group suggests two approaches already highlighted in the literature and which are of interest for future research. The first is the value which professional groups acquire when they are given initiative, autonomy and leadership (McCutchen, 2002; Scanlon et al., 2008). Here, the role assumed by the AL coordinator is fundamental. The fact that they belong to the center, and their familiarity with teacher learning, gives them a different role to that of the mediators of large-scale programmes (Biancarosa et al., 2010). The second approach is related to the first, since the autonomy which characterized the group favored, as usually occurs in short-scale programmes, motivation and learning on the part of teachers (Didion et al., 2020; McCutchen, 2002; Osman & Warner, 2020). This fact promotes consideration of the teacher as a learner, which is a key factor in PD (Kennedy, 2016).

Intensity appears to be modulated, not only by advice given, but also by other factors, like coherence, attitude, and motivation of teachers (Brady et al., 2009;

Brownell et al., 2017; Kennedy, 2016; McCutchen et al., 2002; Osman & Warner, 2020). Even though these variables have not been measured in this study, they have been taken into account, especially in the initial training.

Another fundamental aspect has been the implication of responsible leaders and trainers involved in the PD. The attitude of leaders and trainers may form a barrier in PD (Didion et al., 2020; Kennedy, 2016; Pak et al., 2020). District and school leaders play a key role in PD, where a positive, encouraging attitude towards teachers has repercussions in the motivation and attitude of trainers and teachers (Brady et al., 2009; Desimone & Garet, 2015; Pak et al., 2020). Management conditions the educational organisation of the project on a school level by providing time for teachers to participate in PD sessions and in the activities proposed in PD. In practice, the flexibility of managers has helped the three centres to find space and time to meet and, what is more important, to organise a support system which helps in the process of evaluation and intervention in the classroom. In this sense, McCutchen et al. (2002) highlight the relevance of teacher communities.

To conclude, eight actions which have been key to the project, and which have influenced the result may be highlighted: (1) To openly discuss the difficulties associated with PD and teacher learning, for example, open reflections were made on the fact that after 20 years the “Big Five Ideas” (NRP, 2000) had not been put into practice. (2) To discuss the project as a challenge to creating interest and commitment among teachers when compensating for the initial differences among students (Pfost et al., 2014). Also, to predict the consequences of risk in the short, medium and long term (Connor et al., 2014). (3) To identify reading levels in the classroom through assessment tests given by teachers and to establish measurement criteria from these (Good et al., 2001) and to define an objective related to student performance based on the criteria established with the ILF. (4) To compare the initial and final criteria to promote the competence of teachers. (5) To carry out systematic action during the intervention. (6) To design and adapt classroom activities to each level of reading performance (Established, Emerging and with Possible Difficulty) while allowing teacher autonomy. To give instructions but without following a fixed programme of sequenced stages. (7) To reflect as a group on the difficulties, solutions, assessment, tasks and results. (8) To present a reference model which uses RTI in tier 1, like that in the research by Scanlon et al. (2008) which is presented in initial training and is maintained as a reference.

With regard to limitations, the study may have been improved by the application of specific tools to evaluate some implied factors, such as motivation, perception of competence and attitudes (Brady et al., 2009; Osman & Warner, 2020). Similarly, CK, PK and PCK could have been measured which would have allowed greater objectivity in the differences between the pre-test and the post-test (Brady et al., 2009; Scarparolo & Hammond, 2018). The study could have been followed up with classroom observations so as to facilitate the monitoring of teachers (Piastra et al., 2009). A further limitation is the size of the sample, which is much smaller than those of other studies which have been cited. Studies with large samples are indispensable for obtaining evidence on general aspects of PD, for instance, the five core features of high-quality PD (Biancarosa et al., 2010; Connor et al., 2014; Pak et al., 2020). However, research is needed with smaller groups of voluntary participants

to study the value of variables related to the process of learning on the part of the teacher (McCutchen et al., 2002). Desimone and Garet (2015) pointed out “we need more information about specific aspects of the five features that are important in different contexts, in order to form a better understanding of why some PD works and some doesn’t” (p. 260). This study, which looks at a small sample, shows the relevance of motivation and commitment of teachers. The idea of expanding the sample refers to the development of a bottom to top network in centres, taking into consideration what Kennedy (2016) points out: “mandated PD creates a problem: attendance is mandatory, but learning is not” (p. 29).

Acknowledgements This article has been written within the project PAPI-20-EMERG-21 of the University of Oviedo.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Artiles, C., & Jiménez, J. E. (2011). *Normativización de instrumentos para la detección e identificación de las necesidades educativas del alumnado con trastornos por déficit de atención con o sin hiperactividad (TDAH) o alumnado con dificultades específicas de aprendizaje (DEA)*. Universidades, Cultura y Deportes.
- Biancarosa, G., Bryk, A. S., & Dexter, E. R. (2010). Assessing the value-added effects of literacy collaborative professional development on student learning. *The Elementary School Journal*, 111(1), 7–34. <https://doi.org/10.1086/653468>
- Brady, S., Gillis, M., Smith, T., Lavalette, M., Liss-Bronstein, L., Lowe, E., North, W., Russo, E., & Wilder, T. D. (2009). First grade teachers’ knowledge of phonological awareness and code concepts: Examining gains from an intensive form of professional development and corresponding teacher attitudes. *Reading and Writing*, 22(4), 425–455. <https://doi.org/10.1007/s11145-009-9166-x>
- Brownell, M., Kiely, M. T., Haager, D., Boardman, A., Corbett, N., Algina, J., Dingle, P., & Urbach, J. (2017). Literacy learning cohorts: Content-focused approach to improving special education teachers’ reading instruction. *Exceptional Children*, 83(2), 143–164. <https://doi.org/10.1177/0014402916671517>
- Castejón, L. A., González-Pumariega, S., & Cuetos, A. (2019). Enseñar fluidez lectora en el aula: de la investigación a la práctica (Teaching reading fluency in the classroom: From research to practice). *Ocnos: Revista de Estudios sobre Lectura*, 18(2), 75–84. <https://doi.org/10.18239/ocnos.2019.18.2.1961>
- Castejón, L. A., González-Pumariega, S., & Cuetos, F. (2015). Development of word reading fluency along primary education: A six-year follow-up. *Infancia y Aprendizaje*, 38(4), 842–871. <https://doi.org/10.1080/02103702.2015.1076266>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd). Academic Press.
- Connor, C. M. (2016). A Lattice model of the development of reading comprehension. *Child Development Perspectives*, 10(4), 269–274. <https://doi.org/10.1111/cdep.12200>

- Connor, C. M., Alberto, P. A., Compton, D. L., & O'Connor, R. E. (2014). *Improving reading outcomes for students with or at risk for reading disabilities: A synthesis of the contributions from the Institute of Education Sciences Research Centers (NCSEER 2014–3000)*. National Center for Special Education Research, Institute of Education Sciences, U.S. Department of Education.
- Darling-Hammond, L., Hyster, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute.
- Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, 38(3), 181–199. <https://doi.org/10.3102/0013189X08331140>
- Desimone, L. M., & Garet, M. S. (2015). Best practices in teachers' professional development in the United States. *Psychology, Society and Education*, 7(3), 252–263. <https://doi.org/10.25115/psyse.v7i3.515>
- Didion, L., Toste, J. R., & Filderman, M. J. (2020). Teacher professional development and student reading achievement: A meta-analytic review of the effects. *Journal of Research on Educational Effectiveness*, 13(1), 29–66. <https://doi.org/10.1080/19345747.2019.1670884>
- Ehri, L., & Flugman, B. (2018). Mentoring teachers in systematic phonics instruction: Effectiveness of an intensive year-long program for kindergarten through 3rd grade teachers and their students. *Reading and Writing*, 31, 425–456. <https://doi.org/10.1007/s11145-017-9792-7>
- Evens, M., Elen, J., Larmuseau, C., & Depaeppe, F. (2018). Promoting the development of teacher professional knowledge: Integrating content and pedagogy in teacher education. *Teaching and Teacher Education*, 75, 244–258. <https://doi.org/10.1016/j.tate.2018.07.001>
- Garet, M. S., Cronen, S., Eaton, M., Kurki, A., Ludwig, M., Jones, W., Uekawa, K., Falk, A., Bloom, H., Doolittle, F., Zhu, P., & Szejnberg, L. (2008). *The impact of two professional development interventions on early reading instruction and achievement (NCEE 2008–4030)*. National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences.
- Garet, M. S., Wayne, A. J., Stancavage, F., Taylor, J., Eaton, M., Walters, K., & Sepanik, S. (2011). *Middle school mathematics professional development impact study: Findings after the second year of implementation. NCEE 2011–4024*. National Center for Education Evaluation and Regional Assistance.
- González-Seijas, R.M., & Cuertos, F. (2008). *ALEI. Actividades para el aprendizaje de la lectura y la escritura. Cuaderno de evaluación inicial (ALEI. Activities for learning to read and write. Initial assessment notebook)*. CEPE.
- Good, R. H., Simmons, D., & Kame'enui, E. (2001). The importance and decision making utility of a continuum, of fluency-based indicators of foundational reading skills for third-grade high-stakes outcomes. *Scientific Studies of Reading*, 5, 257–288. https://doi.org/10.1207/S1532799XSR0503_4
- Hudson, A. K., Moore, K. A., Han, B., Wee Koh, P., Binks-Cantrell, E., & Malatesha, R. (2021). Elementary teachers' knowledge of foundational literacy skills: A critical piece of the puzzle in the science of reading. *Reading Research Quarterly*, 56(S1), S287–S315. <https://doi.org/10.1002/rrq.408>
- Kennedy, M. (2016). How does professional development improve teaching? *Review of Educational Research*, 86(4), 945–980. <https://doi.org/10.3102/0034654315626800>
- Landry, S. H., Swank, P. R., Anthony, J. L., & Assel, M. A. (2011). An experimental study evaluating professional development activities within a state funded pre-kindergarten program. *Reading and Writing*, 24(8), 971–1010. <https://doi.org/10.1007/s11145-010-9243-1>
- Lane, H. B., Hudson, R. F., Leite, W. L., Kosanovich, M. L., Strout, M. T., Fenty, N. S., & Wright, T. L. (2008). Teacher knowledge about reading fluency and indicators of students' fluency growth in reading first schools. *Reading and Writing Quarterly*, 25(1), 57–86. <https://doi.org/10.1080/10573560802491232>
- Ledesma, R., Macbeth, G., & Cortada, N. (2008). Tamaño del efecto: Revisión teórica y aplicaciones con el sistema estadístico ViSta. *Revista Latinoamericana De Psicología*, 40(3), 425–439.
- Lesnick, J., Goerge, R., Smithgall, C., & Gwynne, J. (2010). *Reading on grade level in third grade: How is it related to high school performance and college enrolment?* Chapin Hall at the University of Chicago.
- Lonigan, C. J., Farver, J. M., Phillips, B. M., & Clancy-Menchetti, J. (2011). Promoting the development of preschool children's emergent literacy skills: A randomized evaluation of a literacy-focused curriculum and two professional development models. *Reading and Writing*, 24(3), 305–337. <https://doi.org/10.1007/s11145-009-9214-6>

- McCutchen, D., Abbott, R. D., Green, L. B., Beretvas, S. N., Cox, S., Potter, N. S., Quiroga, T., & Gray, A. L. (2002). Beginning literacy: Links among teacher knowledge, teacher practice, and student learning. *Journal of Learning Disabilities*, 35(1), 69–86. <https://doi.org/10.1177/002221940203500106>
- McMahan, K. M., Oslund, E. L., & Odegard, T. N. (2019). Characterizing the knowledge of educators receiving training in systematic literacy instruction. *Annals of Dyslexia*, 69, 21–33. <https://doi.org/10.1007/s11881-018-00174-2>
- Moats, L. (2009). Knowledge foundations for teaching reading and spelling. *Reading and Writing*, 22(4), 379–399. <https://doi.org/10.1007/s11145-009-9162-1>
- National Reading Panel (US), National Institute of Child Health, & Human Development (US). (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction: Reports of the subgroups*. National Institute of Child Health and Human Development, National Institutes of Health.
- Osman, D. J., & Warner, J. R. (2020). Measuring teacher motivation: The missing link between professional development and practice. *Teaching and Teacher Education*, 92, 103064. <https://doi.org/10.1016/j.tate.2020.103064>
- Paige, D. D., Smith, G. S., & Magpuri-Lavell, T. (2019). Learning to improve: Report of a three-year capacity-building project leveraging professional development + coaching to improve third-grade reading outcomes. *Advances in Literary Study*, 7(4), 193–223. <https://doi.org/10.4236/als.2019.74013>
- Pak, K., Desimone, L. M., & Parsons, A. (2020). An integrative approach to professional development to support college-and career-readiness standards. *Education Policy Analysis Archives*, 28(111), 1–23. <https://doi.org/10.14507/epaa.28.4970>
- Pfost, M., Hattie, J., Dörfler, T., & Artelt, C. (2014). Individual differences in reading development: A review of 25 years of empirical research on Matthew effects in reading. *Review of Educational Research*, 44, 402–420. <https://doi.org/10.1177/0022219411417568>
- Piasta, S. B., Connor, C. M., Fishman, B. J., & Morrison, F. J. (2009). Teachers' knowledge of literacy concepts, classroom practices, and student reading growth. *Scientific Studies of Reading*, 13(3), 224–248. <https://doi.org/10.1080/10888430902851364>
- Scanlon, D. M., Gelzheiser, L. M., Vellutino, F. R., Schatschneider, C., & Sweeney, J. M. (2008). Reducing the incidence of early reading difficulties: Professional development for classroom teachers versus direct interventions for children. *Learning and Individual Differences*, 18(3), 346–359. <https://doi.org/10.1016/j.lindif.2008.05.002>
- Scarparolo, G. E., & Hammond, L. S. (2018). The effect of a professional development model on early childhood educators' direct teaching of beginning reading. *Professional Development in Education*, 44(4), 492–506. <https://doi.org/10.1080/19415257.2017.1372303>
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14. <https://doi.org/10.3102/0013189X015002004>
- 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. REL 2007-No. 033. Regional Educational Laboratory Southwest (NJ1).

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Verónica Martínez¹  · Luis Castejón¹  · Soledad González-Pumariega¹ 

Luis Castejón
luiscf@uniovi.es

Soledad González-Pumariega
ssoledad@uniovi.es

¹ Department of Psychology, University of Oviedo, Plaza Feijóo s/n, 33003 Oviedo, Asturias, Spain