Effects of Scaled-up Professional Development Courses About Inquiry-Based Learning on Teachers

Katja Maass[1], Katrin Engeln[2]

[1]International Centre for STEM Education, University of Education, Freiburg, Germany

[2]Leibniz Institute for Science and Mathematics Education (IPN), Kiel, Germany

Correspondence: Katja Maass, International Centre for STEM Education, University of Education, Freiburg, Germany.

Abstract

Although well researched in educational studies, inquiry-based learning, a student-centred way of teaching, is far away from being implemented in day-to-day science and mathematics teaching on a large scale. It is a challenge for teachers to adopt this new way of teaching in an often not supportive school context. Therefore it is important to provide high-quality professional development (PD) at a large scale. However, there is little empirical evidence about the effects of scaled-up professional development initiatives. Therefore, this paper presents an international research study for which long-term PD courses have been designed. These courses have been implemented across Europe in twelve different countries at scale by using the so-called "Cascade Model". Here, course leaders are educated, who in turn educate other teachers. The research study aimed at evaluating the overall impact the scaled-up PD had on teachers and at also identifying variables influencing this impact.

Keywords: inquiry-based learning, beliefs, professional development

1. Introduction

Research and experience reveals that innovative teaching approaches promoted by mathematics and science education researchers are far beyond the day-to-day practices of teachers in many countries (Akker, Gravemeijer, McKenney, & Nieveen, 2006; Boaler, 2008; Krainer, 2011). This is true in particular for inquiry-based learning (IBL) (Hudson, McMahon, & Overstreet, 2002; OECD, 2009a; Weiss, Pasley, Smith, Banilower, & Heck, 2003), a student-centered way of teaching in which students raise questions, explore situations, and develop their own ways towards solutions (Dorier & Maass, 2014).

As teachers are the core enablers of change in educational practices, they need support to do so and they need it on a large scale. However, here we face several challenges when implementing IBL:

First, it is simply a challenge for teachers to change their role from being just an instructor, who transmits knowledge, to being a facilitator of learning processes as required in relation to IBL (Swan, 2005, 2007). This is by no means easy and requires changes in teachers' knowledge, classroom practices and beliefs (Clarke & Hollingsworth, 2002; Zehetmeier & Krainer, 2011). Classroom management is often perceived to be more difficult when teaching inquiry-based (Walker, 2007).

Second, from teachers' perspective there are several impediments to the implementation of innovative teaching approaches such as IBL, which need to be overcome: Systemic factors such as the classroom and existing norms (e.g., students are not used to inquiry-based learning), teachers' community (e.g., limited collaboration with colleagues), school development and leadership (e.g., limited support by the head of school), the national educational and political system (e.g., non-supportive assessment) (Krainer & Wood, 2008; Skott, 2013), or a lack of resources (Engeln, Euler, & Maass, 2013) might hinder the implementation of IBL.

Consequently and, third, any professional development (PD) activity aiming to encourage change needs to consider these challenges and fulfill certain quality criteria (e.g., Guskey, 2000; Loucks-Horsley, Stiles, Mundry, Love, & Hewson, 2009) in order to actually encourage and enable teachers to implement IBL. However, designing PD courses with these features is no guarantee that teachers will actually implement what they have learnt during the PD course. The difficulties encountered when implementing new practices are often unexpected for teachers (Berman & McLaughlin, 1978; Fullan & Miles, 1992). There is also evidence that many teachers either do not or only slowly implement what they have learnt (e.g., Maass, 2011).

Fourth, the question of how to impact the instruction of a large number of teachers further complicates the issue (e.g., Maass & Artigue, 2013; Roesken-Winter, Hoyles, & Blömeke, 2015). So far, little is known on how to scale up professional development (Jackson et al., 2015; Roesken-Winter, Schüler, Stahnke, & Blömeke, 2015). The issue of how to scale-up PD courses is complex and closely related to the question of how programs for professional development can be implemented practically within different contextual settings (Adler & Jaworski, 2009). For example, what are the strengths and weaknesses of the so-called "Cascade Model" in which course leaders are educated, and then in turn educate other teachers (Maass & Artigue, 2013)? One of the major concerns with this model is the question of the forms of knowledge and practice that can actually be "handed down" the cascade (OECD, 1998). Thus, a major issue when talking about professional development is the question of scaling-up.

This paper presents an international research study for which long-term PD courses addressing the four challenges above have been designed and implemented across Europe in different cultural settings on a large scale by using the so-called "Cascade Model". It aimed at researching the overall impact of the scaled-up PD courses on teachers.

The research study was carried out within the project Primas, a research and implementation project within the Seventh Framework Program of the European Union. Between 2010 and 2013 fourteen universities from twelve countries worked together to promote the implementation of IBL in mathematics and science. Each of the partner countries was given the goal of educating at least 100 teachers within the years 2012 and 2013.

The quantitative evaluation of Primas focused on the overall impact of the implementation of IBL in day-to-day teaching with a pre-post study of 50-100 teachers and around 1500 students per partner country (Maass & Doorman, 2013). In this paper we focus on the perception of teachers. This seems meaningful as teachers can be considered as experts on various instructional approaches, methods and lesson features due to their education and teaching experience (Kunter & Baumert, 2006). This is in particular true for their own educational intentions, whilst they lack the possibility to compare their teaching to those of other teachers (Baumert et al. 2004).

We attempt to answer the following research questions:

1. Is there a significant change regarding the implementation of IBL in mathematics and science teaching from the perspective of the teacher after participation in our scaled-up PD using the Cascade model?

This main question relates to the overall effectiveness of our scaling – up approach: In how far did teachers participating in our PD using the Cascade model perceive a change in their teaching? The following two questions take into account two variables which probably have an impact on the teachers' development, at least from the perspective of teachers: First, the impediments teachers face should be taken into account. Second, considering the fact that it is indeed a challenge for teachers to adopt this new way of teaching their perceived prior teaching practice should be taken into account in order to explain later results. Consequently our research questions 2 and 3 are:

2. Are classroom management, systemic factors and the availability of resources, predictors for change?

3. Is the degree of implementation of IBL at the beginning of the course a predictor for the change?

The results reveal that our long term PD courses using the Cascade-Model indeed had a positive impact on teachers' perceived use of IBL. They also indicate that the degree of change depends on teachers' perceived pre-knowledge of IBL. Restricting factors, such as systemic restrictions, availability of materials and classroom management have a constricted impact on the change. Altogether, these results hint to the positive impact scaled up PD using the Cascade model can have on teachers.

2. Theoretical Background and Implementation Within Primas

In the following we will address the four challenges as listed above. The theoretical background as discussed here lays an important basis for the design of the study as such, of the PD courses and the evaluation instruments. First, we will give our definition of IBL as this will also outline why it is such a challenge to teach in this way.

2.1 Inquiry-Based Learning (IBL)

IBL is discussed both in mathematics education (e.g., Artigue & Blomhøj, 2013) and in science education (e.g., Minner, Levy, & Century, 2010). A definition of inquiry often quoted in science education is that presented by the National Research Council (NRC) (1996, p. 23): "Inquiry is a multifaceted activity that involves making observations; posing questions; examining books and other sources of information to see what is already known; planning investigations; reviewing what is already known in light of experimental evidence; using tools to gather, analyze, and interpret data; proposing answers, explanations and predictions; and communicating the results. Inquiry requires identification of assumptions, use of critical and logical thinking, and consideration of alternative explanations and scientific inquiry refers to the diverse ways in which scientists study the natural world and propose explanations based on the evidence derived from their work."

The existence of such definitions does not exclude the co-existence of different interpretations of IBL. Important differences regard, for example: the degree of autonomy given to the students in the selection of problems and in the inquiry process itself; the ratio between the development of inquiry competences and the development of scientific techniques aimed at by the curriculum; and the importance given to authentic connections to real life questions (Artigue & Blomhøj, 2013).

In addition, IBL is often used interchangeably with other terms that describe similar teaching approaches such as hands-on, problem-solving, constructivist learning, project-based, or student-centered approaches (Anderson, 2002).The failure to give a concrete definition has led to misunderstandings and is one reason for discussions about the effectiveness of IBL. For example, Kirschner, Sweller, and Clark (2006) characterize IBL as a minimal guidance approach and conclude that this approach does not work. In a direct response Hmelo-Silver, Duncan, and Chinn (2007) argue that inquiry learning is not a minimally guided approach but rather provides extensive scaffolding and guidance. Based on a science survey of 138 studies Minner et al. (2010) conclude that "the evidence of effects of inquiry-based instruction from this synthesis is not overwhelmingly positive, but there is a clear and consistent trend indicating that […] having students actively think about and participate in the investigation process increases their science conceptual learning" (p. 493). Concerning mathematics, there are also several studies, which point to positive effects of IBL of students' achievements and attitudes (for details, see Bruder and Prescott (2013)).

Due to the different understandings of IBL as discussed above and due to the different perspectives of the countries involved in Primas, a pragmatic definition of IBL was elaborated within the project focusing on four core aspects:

In IBL students are invited to: observe phenomena and create their own questions; select scientific tools; carry out experiments; create representations to clarify relationships; seek for evidence-based explanations; interpret and evaluate solutions; connect explanations to scientific knowledge (Dorier & Maass, 2014; National Research Council (NRC), 2000, p. 27). Summing up, they are supposed to *do inquiry* and conduct investigations.

Additionally, students should engage in actual *hands-on activities* (e.g., Prince & Felder, 2007; Savery, 2006). These hands-on activities can carried out either in student-led or teacher-led inquiry. As the first requires a particular level of competence, teachers therefore often provide students with a problem or questions and even give them the method required to solve the problem (Staver & Bay, 1987). Although of course hands-on activities are an important component of science education and lead to a deeper understanding of science and mathematics (Schoenfeld, 1992), science education researchers have critically discussed the practice where students follow recipes to reach particular, pre-determined outcomes (Dewey, 1938; Hodson, 1996; Hofstein & Kind, 2012).

On the teachers' part, orchestrating and facilitating learning processes, for example through modeling and coaching, is a subtle skill that is critical in making IBL function well (Colburn, 2006; Hmelo-Silver, 2004). Thus, the role of the teacher in such a setting is different to traditional teaching approaches: pedagogies make a shift away from a 'transmission' orientation, in which teacher explanations, illustrative examples and exercises dominate, towards a more collaborative orientation, in which students work together on 'interconnected', 'challenging' tasks. Here, the teacher's role includes making constructive use of students' prior knowledge, challenging students through effective questions; managing small group and whole class discussions and encouraging the discussion of alternative viewpoints (Swan, 2005, 2007). Summing up, the interaction between students and teachers in class changes and turns from a more teacher-centred way of teaching to a more *student-centred* way. This change of role is a major reason why it is so difficult for teachers to teach in this way.

Another important aspect of IBL, the project Primas focused on, is the relevance the subject of inquiry has to the students. This aspect goes back to the American educator John Dewey (1859-1952), to whom the importance of inquiry in education is generally attributed. Dewey's perspective on IBL implies a practice of teaching based on projects closely linked to students' life and interests (Artigue & Blomhøj, 2013). Therefore, inquiry-based learning should also be *authentic for students' lives*.

Summing up, our definition of IBL includes student-led investigation, hands-on activities, student-centeredness in the teaching approach and authentic connections to students' lives.

2.2 Impediments to the Implementation of IBL

Implementing innovative teaching approaches and changing ones' way of teaching are not a straight-forward and easy matter for teachers. Impediments might prevent them from doing so and thereby may also have an influence on the effects of a related PD course (Joubert & Southerland 2008).

First, the requirements of the school system which can hinder the uptake of IBL are specifications of the curriculum, especially on content, the time being available for instruction, the existing assessment practice, a lack of cooperation with colleagues, a lack of support by the school administration or students' and parents reactions (Dorier & García, 2013; Maass, 2011; Walker, 2007).

Second, quite often teachers also do not feel able to include IBL into their teaching because they experience a lack of resources, such as classroom materials, textbooks, computers, laboratories (Dorier & Garcín, 2013; Engeln, Euler, & Maass, 2013 ; Walker, 2007).

Third, as regards innovative teaching approaches such as IBL in particular, teachers do not feel confident with rethinking their role from initiator and controller to guide and facilitator. They worry about the possibility for discipline problems, classroom management of student-centred work, exorbitant preparation requirements and their lack of knowledge in a particular topic (Engeln et al., 2013; Maass, 2011; Walker, 2007).

Altogether, in our review we found three different categories expected to affect the adoption of IBL from the teachers' perspective and his beliefs: *system restrictions, resources and classroom management*.

2.3 Designing Effective Professional Development Courses

Teachers' PD relates to all different facets of professional competences (Roesken-Winter, Schüler et al., 2015). It includes teachers' content knowledge (knowledge about the subject, CK), pedagogical content knowledge (knowledge about how to teach the subject, PCK) and pedagogical knowledge (general pedagogical knowledge, PK) (Shulman, 1986), teachers' classroom practice (Clarke & Hollingsworth, 2002; Zehetmeier & Krainer, 2011) as well as teachers' beliefs, their relatively long-standing subject knowledge of certain objects and the attitudes linked to that knowledge (Pehkonen & Törner, 1996), teachers' motivation; and teachers' competence in self-reflection (Baumert & Kunter, 2013).

To date a considerable body of research exists as regards features of effective PD (Loucks-Horsley et al., 2009; Roesken-Winter, Schüler, et al., 2015).

In a meta-analysis of research in relation to PD, Lipowsky and Rzejak (2012) distinguish four levels of PD effectiveness: (1) teachers' opinions; (2) teachers' knowledge; (3) teaching; and (4) effects on students' performance and list features of PD which support effects on these levels.

In regard to aspect (1) teachers' opinions, PD initiatives are considered to be effective if they have clear relevance for day-to-day teaching. In addition, teachers appreciate the exchange of experiences with colleagues (Lipowsky, 2004).

Turning to aspect (2) professional knowledge and competences, research shows that during reflection, teachers should be requested to deal with their own beliefs about the nature of their subject and its teaching and learning as well as their teaching experiences (Joubert & Sutherland, 2008; Tirosh & Graeber, 2003). In this respect, PD courses should also address the challenges teachers face in relation to IBL and the needs they have (Guskey, 2000; Maass, 2011). Putnam and Borko (2000) recommend that teacher educators should treat teachers the way they expect teachers to treat students. This seems to be an extremely important factor when it comes to IBL. If teachers are to encourage students to participate in IBL, they must first have experienced inquiry themselves.

When it comes to aspect (3) teaching, PD interventions have proven to be effective if: the interventions are long-term; they combine learning-off-job in courses with learning-on-job in school; and teachers are given feedback about their teaching (Lipowsky, 2004).

Speaking to aspect (4) effects on students' performance, the Lipowsky meta-analysis (2004) shows that PD courses seem to be effective if they have a clear focus on a certain aspect of teaching.

Clearly, in PD the focus should not only be on individual teachers, but also on the context in which they are working (Joubert & Sutherland, 2008; Maass & Artigue, 2013).

Also, when designing a PD activity it is not only important to consider the effects it is supposed to have. One has to be aware that the "task of preparing teachers for inquiry teaching is much bigger than the technical matters...the matter must be addressed... at a level that includes central attention to beliefs and values" (Anderson, 2002). As changing from more teacher-centred ways to IBL is a challenge for teachers, there can be disaffirmation towards IBL on the level of the individual teachers' beliefs. Teachers' beliefs therefore need to be taken into account, when developing a PD course.

However, some studies suggest that beliefs have an influence on classroom practices (e.g., Schoenfeld, 1992),whilst others (e.g., Swan & Swain, 2010) argue that rather than having an influence on practice, beliefs are more the result of practice. Therefore, these linear approaches were modified by Clarke and Hollingsworth (2002) within their cyclic Interconnected Model.

This model explains teacher professional growth in terms of domains that are connected through mediating processes. The domains are the following ones:

- The personal domain: teacher knowledge, beliefs and attitudes
- The domain of practice: professional experimentation
- The domain of consequence: salient outcomes

- The external domain: sources of information, stimulus or support

Within this model professional growth can start at any point of the cycle and would ideally involve all four aspects. It combines teachers' knowledge, beliefs and attitudes, the external stimulus, the implementation within the classroom and the salient outcomes as factors all impacting on professional growth.

The model of Clarke and Hollingsworth (2002) also emphasizes the role of teachers' priors knowledge. Indeed, literature suggests that teachers' prior knowledge and experiences influence what impact a professional development course has on teachers. For example, Schoenfeld (2010) considers teaching as a function of knowledge, goals and beliefs. Cohen (1990) conducted a case study of a teacher and identified her teaching after the professional development as a melange of something old and something new. In the same vain, Pehkonen and Törner (1999) point to the fact that teachers can adapt a new curriculum by interpreting their teaching in a new way. Broad and Evans (2006) highlight the different needs of teachers in their different phases of professional development. However, so far we do not know of any studies examining the impact of teachers' prior experience in IBL on the effects of a PD course about IBL.

2.4 The Primas Model for PD

The Primas model for PD in IBL takes into account all the domains of the Interconnected Model and draws on all the principles for PD as described above.

Considering the time needed for an effective change of day-to-day teaching and taking into account the importance of both formal and informal learning opportunities, we offered long-term PD courses. Our PD activities involved cycles of analysis, implementation and reflection. In the analysis phase (in the PD course), teachers worked on collaborative classroom activities that illustrated the pedagogical challenges of IBL and watched other teachers using these same activities using video. These provided a 'challenge' to existing knowledge, practices and beliefs, which were discussed and provided a clear connection to their day-to-day teaching. In the implementation phase at school, teachers were encouraged to adapt and use the activities in their own classrooms. We encouraged teachers to visit and observe each other's lessons in order to give further opportunities for feedback and reflection. Finally, in the reflection phase (in the PD course), teachers shared their classroom experiences, discussed the pedagogical implications and reflected on the growth of new practices and beliefs. This cyclic process was repeated at each professional development event, as new pedagogical issues were addressed.

In total, the PD activities were designed as seven units. The content of these is described in Table 1. Teachers were encouraged to sequentially focus on major characteristics of IBL and relate these to their day-to-day teaching. Methods used in the modules allow teachers to work in groups, to investigate problems, to reflect on issues and to discuss open questions. In other words, teachers participating in the course experience IBL themselves. Further discussion of the design and trialing of the PD materials may be found in Swan, Pead, Doorman, and Mooldijk (2013).

The concept of the Primas PD courses was designed for an international use. The seven individual modules were developed by the British partner who collected feedback from all other countries so as to make them usable in all other 11 partner countries. Table 1 shows the topics addressed in the Primas PD course.

Table 1. Topics addressed in the Primas PD course

Name	Content
Student-led Inquiry	In this unit, teachers were presented phenomena and invited to pose and pursue their own questions.
Tackling Unstructured Problems	Teachers considered the decisions that are made for students when we present them with structured problems and compared them with unstructured versions of problems.
Learning Concepts Through IBL.	We considered how the processes of IBL may be integrated into the teaching of content.
Asking Questions that Promote Reasoning	Teachers were supposed to reflect on: characteristics of questioning that encourage students to reflect and reason.
Students Working Collaboratively	We offered teachers opportunities to reflect on the characteristics of student-student discussion that benefit learning.
Building on What Students Already Know.	We considered the different ways teachers may use formative assessment to make effective use of students' prior knowledge.
Self and Peer Assessment	Teachers reflected on how students can be encouraged to assess and improve each other's work?

Table 1 gives an overview of the topics addresses. We will now have a look at one of these modules in detail. For example in Module 2 (see table 1) teachers were given the table tennis task as shown in Fig. 1. They were asked to work through the structured problem carefully, list all the decisions that are being made for the students, and revise the problem so that some of these decisions are handed back to students. In the next step, they were asked to discuss pedagogical issues that may arise when using the open version in class and how to overcome them. Next teachers watched a video on the implementation of the problem at school and discussed emerging issues. Afterwards they

planned their own lesson with the tennis tournament problem (see Figure 1). Back at school they implemented the task. In the next seminar lesson, they presented their lesson, showed student work, and highlighted challenges which were discussed in the group.

We implemented the seven modules in all 12 partner countries. For this, the overall international concept of the Primas PD courses was adapted nationally. However, the adaptations made did not conflict with the overall principles as described above and mainly concerned the organizational framing. In order to ensure quality across countries, we discussed the overall PD principles and their implementation into the PD course at the biannual project meetings. In each country, project partners selected and briefed further math education experts who then ran the courses. In 2012 and 2013, about 100 teachers in each country took part in the PD courses. As we had to take into account the national contexts in each country for a successful implementation, the actual timeframe varied across countries (from several weeks to a few months).

Organizing a table tennis tournament

You have the job of organizing a table tennis tournament.

- 7 players will take part
- All matches are singles.
- Every player has to play each of the other players once.

1. Call the players A, B, C, D, E, F, G.

   Complete the list below to show all the matches that need to be played.

   A v B   B v C   ....

   A v C   B v D   ....

2. There are four tables at the club and each game takes half an hour. The first match will start at 1:00 p.m. Copy and complete the poster below to show the order of play,   so that the tournament takes the shortest possible time. Remember that a player cannot be in two places at once! You may not need to use every row and column in the table!

Start Time	Table on which the game will be played			
	1	2	3	4
1:00 p.m.	A v B			
1:30 p.m.				
2:00 p.m.				
2:30 p.m.				
3:00 p.m.				
3:30 p.m.				
4:00 p.m.				
4:30 p.m.				

Figure 1. The worksheet for the PD session about the tennis table tournament

3. Design of the Study

3.1 Hypotheses

The aim of this research is to analyze the impact of our scaled up PD interventions applied within the 12 Primas countries regarding the implementation of IBL. In particular we wanted to test the following hypotheses:

   (1) After the Primas PD course teachers perceive their teaching as being

- more student-centred;
- more authentic;
- more hands-on;
- more investigative (see part 2.1).

   (2) An unsupportive school system, a lack of teaching resources and concerns regarding classroom management

negatively influence the perceived change in classroom teaching regarding the four core elements of IBL.

(3) The prior teaching experience regarding our four core elements of IBL as perceived by teachers predicts respectively their changes.

3.2 Implementation of the Scaled up PD and Data Collection

In all 12 Primas countries the scaled up PD courses on IBL were implemented. The PD courses were all run in the years 2012 – 2013, all involving several cycles of the spiral model.

All countries worked with course leaders to scale up PD in their countries. These course leaders were either engaged teachers from school, pre-service educators from higher education or persons from institutions with a main responsibility in running PD. Depending on their experience with PD and IBL their preparation varied from country to country.

Naturally, in order for Primas to succeed, the 12 different national contexts involved had to be taken into account. Therefore, we first carried out an analysis of the context in each country in which Primas was to be implemented. This included analysing curricula, examining methods used to assess student performance, and looking at the existing conditions for teachers to participate in professional development (Dorier & Garcá, 2013). Second, a baseline study was conducted which informed the project about teachers' beliefs in regard to IBL and their teaching practice (Engeln et al., 2013).

Based on the analysis of the context and the results of the baseline study, the national teams (consisting of members from institutions such as universities, schools, school authorities, teacher education institutes) developed concepts for the adaption of the international model and implementation in the given setting. Subsequently, the national implementations varied to some extent (mainly as regards the organisational framework), whilst subordinating to the major pedagogical principles and content (using the seven modules) as described above (for details, Maass & Doorman, 2013).

In each of the 12 Primas countries a pre-post study was conducted to gain insight into the effects of the Primas intervention. Here, we focus on teacher questionnaires.

3.3 Design of the Teacher Questionnaire

Pre- and post-questionnaires were partly identical and contained three main sections: personal data, attitude and believes on IBL (including impediments to IBL) and teaching practice with respect to IBL (Engeln, 2013).

We assessed the IBL teaching practice following our theoretical concept of IBL and thus using the four core elements of IBL (see 2.1): The scales student-centeredness (STC), authentic connections to students' life (AUT), hands-on activities (HON) and investigative teaching (INV) each comprises three items (Table 2). They items were adapted from the OECD Programme for International Student Assessment (PISA), more concretely from the scales interaction, focus on models or applications, hands-on activities and student investigations (OECD, 2009b, pp. 333-336). These four scales capture our four core elements of IBL and therefore give insight into the implementation of IBL from the perspective of the teacher. Cronbach alpha for these scales is .61, .65, .74 and .80 respectively and therefore shows that the adapted items are applicable for teachers. For each item respondents were expected to rate the frequencies of occurrence with respect to their own classroom practice on a Likert-Scale ("never or hardly ever", "in some lessons", "in most lessons", "in almost all lessons"). Table 2 gives an overview of the items used.

Table 2. Overview of the items used

Name	Item	Pisa 2006
inv1	In my lessons the students design their own experiments/investigations.	(ST34Q08)
inv2	In my lessons the students do experiments/investigations to test out their own ideas.	(ST34Q16)
inv3	In my lessons the students have the chance to choose their own experiments/investigations.	(ST34Q11)
hon1	In my lessons the students spend time doing practical experiments/ investigations.	(ST34Q02)
hon2	In my lessons the students do experiments/investigations by following my instructions.	(ST34Q14)
hon3	In my lessons the students draw conclusions from experiments/investigations they have conducted.	(ST34Q06)
aut1	In my lessons I use this subject to help the students understand the world outside school.	(ST34Q12)
aut2	In my lessons I show the students how this subject is relevant to society.	(ST34Q17)
aut3	In my lessons I explain the relevance of this subject to our daily lives.	(ST34Q15)
stc1	In my lessons the students are given opportunities to explain their ideas.	(ST34Q01)
stc2	In my lessons the students have discussions about the topics.	(ST34Q13)
stc3	The students are involved in class debate or discussion.	(ST34Q09)

We assessed teachers' beliefs on impediments in relation to the implementation of IBL by developing items based on our theoretical knowledge (see 2.2). As suggested by theory, we developed three scales: Missing resources (RES), systemic restrictions (SYS) and problems with classroom management (CLA) (see 2.2). The scales have been confirmed by utilizing the pre-data confirmatory factor analysis with the MPLUS 5.21 program (Muthén & Muthén,

1998-2010) (CFI = .96, RMSEA = .050, SRMR = .03). The resources scale comprises two items while the system scale and the classroom management each comprise four items (Table 3). The response format for these three scales was a four-point Likert scale to reflect teachers' agreement ("strongly disagree", "disagree", "agree", "strongly agree"). Table 3 gives an overview of the items on hindrances to the implementation of IBL.

Table 3. Overview of the items on hindrances to the implementation of IBL

Name	Item
res	I have difficulties in implementing IBL, because I have a lack of adequate teaching materials.
res	I have difficulties in implementing IBL, because IBL is not included in textbooks I use.
sys	I have difficulties in implementing IBL, because there is not enough time in the curriculum.
sys	I have difficulties in implementing IBL, because my students have to take assessments that don't reward IBL.
sys	I have difficulties in implementing IBL, because the curriculum does not encourage IBL.
sys	I have difficulties in implementing IBL, because the school system does not encourage changes.
cla	I have difficulties in implementing IBL, because I worry about students' discipline being more difficult in IBL lessons.
cla	I have difficulties in implementing IBL, because I don't feel confident with IBL.
cla	I have difficulties in implementing IBL, because I worry about my students getting lost and frustrated in their learning.
cla	I have difficulties in implementing IBL, because I think that group work is difficult to manage.

3.4 Sample

The teacher pre-questionnaire was distributed in all 12 participating countries. Altogether it was filled out by 1219 teachers. The number of teachers varies between 46 (4 percent, Malta) and 192 (16 percent, United Kingdom) within the countries. In all, 37 percent of the sample is male. The proportion of male teachers varies significantly between the countries. In Hungary and Slovakia, for example, only 11 percent of the sample is male while in Denmark, the Netherlands and Switzerland more than half of the teachers are male. In all, 63 percent of the sample refers to maths and 37 percent to science (physics, chemistry, biology, general science) when answering the questions about teaching practice (Table 4). Furthermore, the teachers in the sample refer to different age groups of students. Only 8 percent of the sample refers to the age group ten years and younger, while 47 percent chose the age group ten to 14 years and 45 percent the age group 14 years and older.

Table 4. Overview of the sample of the sample of the teacher pre-study

Country	Sample size (proportion)		Male proportion
Cyprus	117	(.10)	.33
Denmark	70	(.06)	.56
Germany	70	(.06)	.37
Hungary	103	(.08)	.11
Malta	46	(.04)	.37
Netherlands	83	(.07)	.52
Norway	68	(.06)	.48
Romania	106	(.09)	.33
Slovakia	119	(.10)	.11
Spain	156	(.13)	.36
Switzerland	89	(.07)	.60
UK	192	(.16)	.48
All	1219	(1.00)	.33

The post sample shows a similar structure. The differences between the countries are only naturally when considering the range of the study and naturally do not impact the individual pre-post comparison of the teachers.

After data collection, we selected those teachers for data evaluation from whom we had matching pre- and post-questionnaires. For several reasons the number of matches is lower than the number of pre- and post-questionnaires in total: Some teachers drop out of the course, others were absent, when one of the questionnaires was completed, some teachers did not fill in the code correctly, which was necessary for matching. Altogether, we were able to match 583 teacher pre- and post-questionnaires.

3.5 Data Evaluation Methods

We used a latent change model to answer our research questions and test hypotheses. Latent change models are structural equation models and like growth models they allow to model changes over time. They allow estimating the change that each individual has over time specifying latent differences. The changes of a variable over time are defined to be a latent difference variable (DIFF). Thus, the latent change model allows to test whether a significant difference exist between two measurements (pre- and post-measurement). In this respect the latent change model leads to similar results as a t-Test. However, latent change modelling offers a greater of flexibility in testing a variety of hypotheses

concerning the change. For example, in addition to a t-Test, it also gives us the opportunity to include predictors of these intra-individual changes and therefore to explain differences in inter-individual change (between the teachers).

We first specified for each of the four core elements of IBL a basic latent change model with auto-correlated errors. Correlated error variables of the same indicators account for the higher correlations of identical indicators over time (Sörbom, 1975) (Figure 2). This basic model does not include predictors for the interpersonal differences in change. PRE1, PRE2 and PRE3 symbolize the scores of three items at the pre measurement while POST1, POST2 and POST3 symbolize the scores of the same three items at the post measurement (measured variables). PRE and POST are the corresponding latent variables. Within this model the variable POST is perfectly determined through the variables PRE and DIFF. Through this the DIFF variable can be interpreted as the change of the PRE variable (Geiser, 2011; Steyer, Eid, & Schwenkmezger, 1997). This basic model has been specified to look for a change in relation to our four core elements of IBL (student-led investigation, hands-on activities, authentic connections to students' lives and student-centeredness).
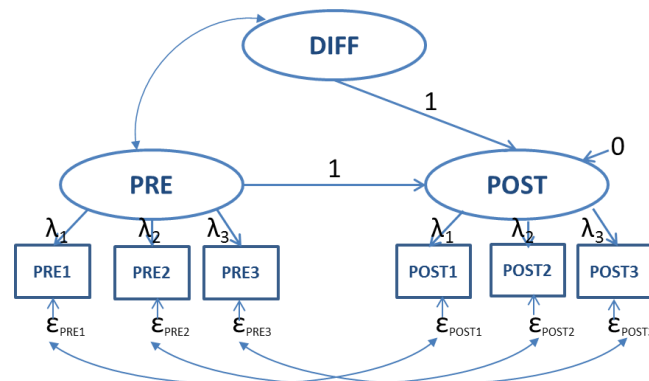

Figure 2. Basic Latent change model with auto correlated errors

In a second step the model was expanded. We used a structural equation approach to examine inter-individual differences in change. The latent variable PRE as well as the variables "resources" (RES), "systemic restrictions" (SYS) and "classroom management" (CLA) are included as predictor of the latent change variable (DIFF). Here we use the sum scores as a predictor. Additionally, the variables CLA, RES and SYS are also taken as predictor of the latent variable PRE (see Fig. 3).
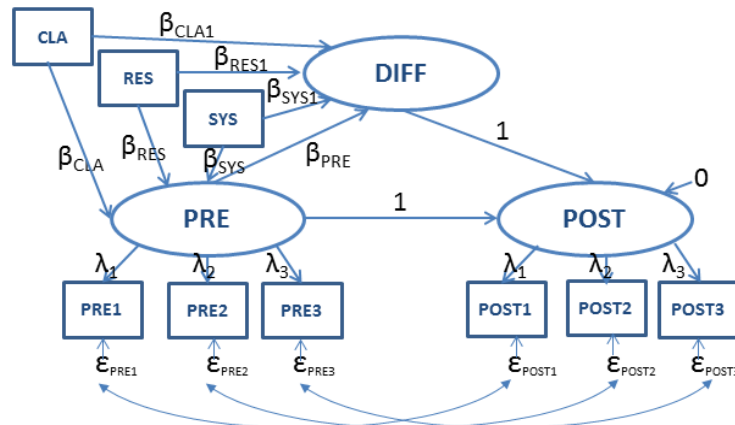

Figure 3. Expanded latent change model: Structural model with pre- level and impediments as predictor for change

We dropped missing data list wise. Within the data of teachers having a matching pre- and post- questionnaire model we suppose data to be missing at random. For all the models we specified less than 3 percent of the data was missing (Schafer & Graham, 2002).

4. Results

We estimated the basic latent change model based on data from 549 teachers for each of the four elements of IBL with the MPLUS 5.21 program (Muthén & Muthén, 1998-2010).

To illustrate our results we discuss the results of "student-centredness" (STC) in more detail. A correlation table is shown in Table 5.

Table 5. Correlation "Student-Centeredness"

Observe variable	pre1 (stc1)	pre2 (stc2)	pre3 (stc3)	post1 (stc1)	post2 (stc2)	post3 (stc3)
pre1 (stc1)	1.000					
pre2 (stc2)	.288	1.000				
pre3 (stc3)	.272	.412	1.000			
post1 (stc1)	.629	.256	.200	1.000		
post2 (stc2)	.221	.611	.326	.338	1.000	
post3 (stc3)	.195	.340	.544	.255	.467	1.000

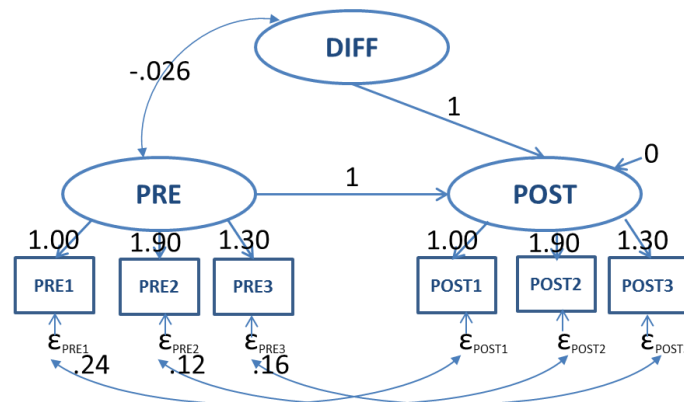The estimated parameters from the structural equation modelling are shown in Figure 4.

Figure 4. Results of the basic latent change model: IBL element STC

The hypothesized model appears to be a good fit to the data. The ratio of χ2 and df is 2.220, the comparative fit index (CFI) is .99; and the root-mean-square error (RMSEA) is .05 (Schermelleh-Engel, Moosbrugger, & Müller, 2003). The estimated p-values of the mean of the difference variable DIFF is less than .001. Therefore, a significant change of "student-centredness" (STC)" exists. As the mean of STC-DIFF is positive, it indicates a significant increase of STC between pre- and post-testing. Furthermore the estimate variances of PRE is .104, p<.001 and the estimate variance of DIFF is .049, p<.001. These variances indicate inter-individual differences of "student-centeredness" at the pre-measurement and also inter-individual differences in the change (see table 6). Table 6 also displays the estimated values for the scales AUT, HON and INV.

Table 6. Result of the specified basic latent change model

	STC	AUT	HON	INV
Mean	2.949	2.744	1.974	1.662
z-Value	101.125	79.228	68.344	58.733
p	.000	.000	.000	.000
DIFF	.080	.063	.108	.160
z-Value	4.985	3.296	5.686	7.723
p	.000	.001	.000	.000
Variance	.104	.403	0.279	.249
Z	5.053	10.750	8.970	9.871
p	.000	.000	.000	.000
Variance DIFF	.049	.127	.099	.143
Z	4.661	8.844	7.161	8.364
p	.000	.000	.000	.000
Model fit				
χ^2	19.784	14.583	22.562	9.561
df	9	9	9	9
CFI	.989	.998	.990	1.00
RMSEA	.047	.034	.052	.011

The estimated results are similar. The estimated models have an acceptable fit. The mean values of DIFF are positive and the corresponding estimated p-values are less than .001. Furthermore, inter-individual variance exists in the pre-level and also in the change. Consequently, the basic latent change model estimates a significant increase for STC, AUT, HON and INV. These results support our hypothesis 1.

As has been described above the model was expanded to get more inside and to explain inter-individual differences. The three variables CLA (classroom), RES (resources) and SYS (system restriction) were included both as a predictor for PRE and for DIFF. Further, the PRE variable was taken into account as a predictor for DIFF.

Again, we start discussing the result for STC (student-centred) in more detail. The ratio of χ2 and df, the CFI and the RMSEA indicate a satisfactory model fit (Table 7) (Schermelleh-Engel et al., 2003).

As shown in table 7, inter-individual differences in intra-individual change in STC (STC-DIFF) are significantly predicted by STC-PRE and SYS. Teachers having less STC teaching practice at the beginning of the study report a higher change. Furthermore, SYS has a positive direct effect on STC-DIFF. Teachers reporting more systemic impediments at the beginning report a higher change of student-centeredness (STC-DIFF). Figure 5 and Table 7 show the estimated parameters.

Table 7. Result of the specified basic latent change model

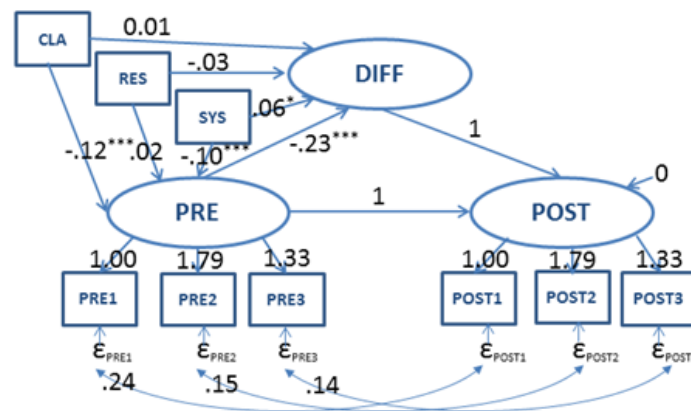|	STC	AUT	HON	INV
Mean	2.949	2.744	1.974	1.662
z-Value	101.125	79.228	68.344	58.733
p	.000	.000	.000	.000
DIFF	.080	.063	.108	.160
z-Value	4.985	3.296	5.686	7.723
p	.000	.001	.000	.000
Variance	.104	.403	0.279	.249
Z	5.053	10.750	8.970	9.871
p	.000	.000	.000	.000
Variance DIFF	.049	.127	.099	.143
Z	4.661	8.844	7.161	8.364
p	.000	.000	.000	.000
Model fit				
χ^2	19.784	14.583	22.562	9.561
df	9	9	9	9
CFI	.989	.998	.990	1.00
RMSEA	.047	.034	.052	.011

Figure 5. Results of the expanded latent change model: IBL element STC (* p<.05, *** p<.001)

The estimated model explains 15 percent of STC-DIFF. So, the included predictors explain a small part of the variance and therefore a small part of the interpersonal differences. Other reasons for these interpersonal differences (see also 3.4) are beyond the focus of this paper.

After this detailed look at STC, we have a look at all variables (STC, AUT, HON, INV). The estimated models have also a satisfactory model fit.

Value of the pre-variable as a predictor for change: PRE is in all four cases a predictor for the change. The regression coefficients show that more experience at the beginning of the PD course predicts a smaller change. The size of the increase (DIFF) can be predicted by the actual status of the core element of IBL in question. From a higher level of the pre-value of a certain variable it is more difficult to increase. Thus, less experience at the beginning (PRE) leads to a greater increase (DIFF). These results support our hypothesis 3 and specify the connection between both variables.

Impediments as a predictor of DIFF: As we have seen SYS predicts STC-DIFF in a way that teachers reporting more systemic impediments at the beginning report a higher change of student-centeredness (STC-DIFF). In addition, RES predicts HON-DIFF and INV-DIFF. Apparently, missing resources cause a smaller impact on teaching practice related to hands-on (HON) and to investigation (INV). These more differentiated results contradict our hypothesis 2, which assumes a general impact of the impediments on the change in relation to all four core elements.

Impediments as a predictor for the pre-variable: In addition, the model predicts the value of the PRE in relation to classroom management (CLA), resources (RES) and system (SYS). CLA is a significant predictor for all pre-values of the four core elements, SYS for all except HON, whilst RES only predict HON (Table 7). Since POST is perfectly determinated through the variables PRE and DIFF it is also predicted by the impediments taken into account. Table 7 shows the results of the specified basic latent change model.

5. Discussion

5.1 Summary

In our project we developed a PD course to encourage and enable teachers to implement IBL. The course followed relevant quality criteria as discussed in research papers and was deliberately designed for international use in different cultural settings. Naturally, in all 12 Primas countries the national context is different, but there are also many commonalities (e.g., IBL is included in all curricula, but in most countries not in assessment, for details see Dorier and Garcia, 2013), which support the idea of designing an international course.

All Primas countries implemented the PD courses on a large scale following the Cascade-Model. In all courses, the same principles (see 2.3) and the seven pre-designed modules were used. National adaptions mainly concerned organizational aspects or the use of additional modules, but did not dilute the major principles of the international PD concept (see 2.3).

The course was implemented in different countries with different contexts. Between the countries the sample was different as regards the percentage of male and female teachers, the science subjects taught and the age groups taught.

Against this background our results show that it is indeed possible to develop a PD course for use in different cultural settings and scale it up with the Cascade Model. In the following section we will discuss the results in detail.

5.2 Discussion

First, our results show that the Primas PD model had an impact on teachers and their perceived implementation of IBL as the mean of the estimated difference variable is significantly different from zero for all four core elements of IBL. Considering the number of participating countries and the variety of contexts in which the PD model was implemented this indicates the quality of the international PD course. This result is a relevant and new contribution to knowledge because so far we lack international studies on the effects of professional development courses (Roesken-Winter, Hoyles, et al., 2015) and studies on the effects of scaled-up professional development courses (Adler & Jaworski 2009).

Second, our results give insight into influence of teachers' prior knowledge on the effects of the PD course. They show that the pre-variable is an important predictor for the increase of IBL as perceived by the teacher. Apparently, the more the teachers in our sample seem to perceive doing IBL at the beginning of a PD the more difficult it is to increase this level. Here, we see two hypotheses to explain this result. First, it might be easier to do first steps in IBL but more complicated to get proficient. Second, teachers who know more about IBL might see their progress more critically. These results therefore are in accordance with Schoenfeld (2010) and Cohen (1990), but go beyond them by showing the concrete influence the prior knowledge has.

Third, our results very specifically and concretely outline which impediments teachers see in relation to the implementation of IBL.

Astonishingly, the more systemic restrictions teachers see at the beginning of the professional development course, the more change in relation to student-centredness (STC) they perceive at the end of the course. One explanation for this result could be that our Primas PD course actually helped them to overcome systemic restrictions they perceived in relation to student-centredness at the beginning. This would confirm the necessity of actually addressing challenges teachers face (Maass, 2011) during the PD course.

Additionally, the availability of resources (RES) predicts changes regarding hands-on activities (HON) and investigations (INV). Apparently, a lack of resources cannot be compensated through the course. This is understandable if we do not only think of classroom materials but of materials necessary for the experiments. This is in accordance with Schoenfeld (2010) who argues that what teachers do in classrooms is a function of their resources, goals and orientations (their beliefs, values, etc).

Problems in relation to "classroom management" do not seem to be a predictor of any of the difference values. Considering the variance of participants there are surely a number of explanations for this. However, if we take into account, that CLA is a significant predictor for the pre-value, this might also indicate that the professional development course supported the teachers in dealing with classroom management. This result therefore contributes to knowledge as it becomes evident that the demanded relevance of PD to teaching practice (Lipowsky & Rzejak, 2012) is important specifically for IBL.

Additionally, our results show an influence of the impediments on the pre-variable which we will outline now.

They show that CLA is a significant predictor for all pre-values of the four core elements and SYS for all except HON and consequently, that these impediments prevent teachers from implementing IBL without PD (Author, 2009, 2011). These results highlight the connection of teachers' beliefs and teaching as discussed in literature (Anderson, 2002; Clarke & Hollingsworth, 2002; Stipek, Givvin, Salmon, & MacGyvers, 2001). The fact that SYS does not predict the pre-variable of HON may indicate, at least in our sample, that hands-on activities - maybe by following recipes – are an important focus already.

What is more, they show that RES predicts the pre-variable of HON. It is obvious that a lack of resources might hinder the implementation of hands-on activities. Apparently, before the course started, teachers did not perceive a lack of resources in relation to investigations (INV), student-centredness (STC) and authenticity (AUT). Here, different hypotheses can be set up: It might be, that teachers – before the intervention - would not do investigation, authentic connections or student-centredness anyway and therefore do not need materials or that they think they have sufficient materials (e.g., for authentic connections) or that they do not need particular materials (e.g., for student-centredness).

Altogether, our study shows that PD courses scaled up with the Cascade-Model can indeed have an impact on teachers' perceived classroom implementation regarding IBL across 12 different countries and settings. It also demonstrates that it is possible to design PD materials and principles for use in different countries and highlights factors that can influence the uptake of IBL for the individual teacher.

5.3 Limitations of the Study

This study of course also has limitations, which need to be into account when using the results: First of all, we had no control group. Considering, however, the large number of participating countries and the scale of the study, the results have a value on its own right.

Furthermore, we have no insight in what really happened in the classroom. Our results are based on teachers' perceptions and thus indicate their intentions. Classroom observations at this scale were of course not possible. For these reasons, we reported on teachers perceptions of their teaching in this paper. The design of the evaluation will however allow for more insights into what happened in the classrooms: We will report on students' perceptions of change based on students' questionnaires and our classroom observation of selected cases elsewhere.

As explained above, the study had a substantial drop-out rate. Considering the sample size and the data collection across 12 countries this seems understandable. Smaller studies of a shorter duration (within one school term) might have a smaller drop out rate but than cannot provide data from a long-term PD.

Additionally, our results do not provide in-depth results on which elements of the PD design actually where the ones which supported the change of teachers. It would be the task of another research study to look in more detail at this question.

References

Adler, J., & Jaworski, B. (2009). Public writing in the field of mathematics teacher education. In R. Even & D. L. Ball (Eds.), *The professional education and development of teachers of mathematics - the 15th ICMI study* (pp. 249-254). New York: Springer. https://doi.org/10.1007/978-0-387-09601-8_26

Akker, J. V. D., Gravemeijer, K., McKenney, S., & Nieveen, N. (2006). Introducing educational design research. In J. v. d. Akker, K. Gravemeijer, S. McKenney & N. Nieveen (Eds.), *Educational design research* (Vol. 1, pp. 3-7). Oxford: Routledge Chapman & Hall.

Anderson, R. D. (2002). Reforming science teaching: What research says about inquiry. *Journal of Science Teacher Education, 13*(1), 1-12. https://doi.org/10.1023/A:1015171124982

Artigue, M., & Blomhøj, M. (2013). Conceptualizing inquiry-based education in mathematics. *ZDM, 45*(6), 797-810. https://doi.org/10.1007/s11858-013-0506-6

Baumert, J., & Kunter, M. (2013). The COACTIV model of teachers' professional competence. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers: Results from the COACTIV project*. Berlin: Springer. https://doi.org/10.1007/978-1-4614-5149-5_2

Baumert, J., Kunter, M., Brunner, M., Krauss, S., Blum, W., & Neubrand, M. (2004). Mathematikunterricht aus Sicht der PISA-Schülerinnen und Schüler und ihrer Lehrkräfte. In P.-K. Deutschland (Ed.), *PISA 2003 - Der Bildungsstand der Jugendlichen in Deutschland - Ergebnisse des zweiten internationalen Vergleichs* (pp. 314-354). Münster: Waxmann.

Berman, P., & McLaughlin, M. W. (1978). Federal programs supporting educational change, Vol. VIII: Implementing and sustaining innovations.

Boaler, J. (2008). Bridging the gap between research and practice: International examples of success. In M. Menghini, F. Furinghetti, L. Giarcardi & F. Arzarella (Eds.), *The first century of the International Commission on Mathematics Instruction (1908–2008): Reflecting and shaping the world of mathematics education*. Roma: Instituto della Enciclopedia Italiana foundata da Giovanni Treccani.

Broad, K., & Evans, M. (2006). *A review of literature on professional development content and delivery modes for experienced teachers*: University of Toronto, Ontario Institute for Studies in Education.

Bruder, R., & Prescott, A. (2013). Research evidence on the benefits of IBL. *ZDM, 45*(6), 811-822. https://doi.org/10.1007/s11858-013-0542-2

Clarke, D., & Hollingsworth, H. (2002). Elaborating a model of teacher professional growth. *Teaching and Teacher Education, 18*(8), 947-967. https://doi.org/10.1016/S0742-051X(02)00053-7

Cohen, D. K. (1990). A revolution in one classroom: The case of Mrs. Oublier. *Educational evaluation and policy analysis, 12*(3), 311-329. https://doi.org/10.3102/01623737012003311

Colburn, A. (2006). *What teacher educators need to know about inquiry-based instruction*. Paper presented at the Annual meeting of the Association for the Education of Teachers in Science Akron, OH.

Dewey, J. (1938). *Experience and education*: Kappa delta Pi.

Dorier, J. L., & Garc á, F. J. (2013). Challenges and opportunities for the implementation of inquiry-based learning in day-to-day teaching. *ZDM, 45*(6), 837-849. https://doi.org/10.1007/s11858-013-0512-8

Dorier, J. L., & Maass, K. (2014). Inquiry-based mathematics education *Encyclopedia of Mathematics Education* (pp. 300-304). Heidelberg: Springer.

Engeln, K. (2013). WP9: Report about the implementation of IBL in those schools who took part in the professional development *PRIMAS: Promoting inquiry-based learning in mathematics and science education across Europe*: European Commision.

Engeln, K., Euler, M., & Maass, K. (2013). Inquiry-based learning in mathematics and science: A comparative baseline study of teachers' beliefs and practices across 12 European countries. *ZDM, 45*(6), 823-836. https://doi.org/10.1007/s11858-013-0507-5

Engeln, K., Euler, M., & Maass, K. (2013). Inquiry-based learning in mathematics and science: a comparative baseline study of teachers' beliefs and practices across 12 European countries. *ZDM 45*(6), 1-14. https://doi.org/10.1007/s11858-013-0507-5

Fullan, M. G., & Miles, M. B. (1992). Getting reform right: What works and what doesn't. *Phi delta kappan, 73*(10), 745-752.

Geiser, C. (2011). *Datenanalyse mit MPLUS [Analysing data using MPLUS]*. Wiesbaden: VS Verlag. https://doi.org/10.1007/978-3-531-93192-0

Guskey, T. R. (2000). *Evaluating Professional Development*. Thousand Oaks, CA: Cirwin Press.

Hmelo-Silver, C. E. (2004). Problem-Based Learning: What and How do Students Learn? *Educational Psychology Review, 16*(3), 235-266. https://doi.org/10.1023/B:EDPR.0000034022.16470.f3

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist, 42*(2), 99-107. https://doi.org/10.1080/00461520701263368

Hodson, D. (1996). Laboratory work as scientific method: three decades of confusion and distortion. *Journal of Curiculum Studies, 28*(2), 115-135. https://doi.org/10.1080/0022027980280201

Hofstein, A., & Kind, P. M. (2012). Learning in and from science laboratories *Second international handbook of science education* (pp. 189-207): Springer. https://doi.org/10.1007/978-1-4020-9041-7_15

Hudson, S. B., McMahon, K. C., & Overstreet, C. M. (2002). The 2000 national survey of science and mathematics education: Compendium of tables. *Chapel Hill, NC: Horizon Research*.

Jackson, K., Cobb, P., Wilson, J., Webster, M., Dunlap, C., & Appelgate, M. (2015). Investigating the development of mathematics leaders' capacity to support teachers' learning on a large scale. *ZDM, 47*(1), 93-104. https://doi.org/10.1007/s11858-014-0652-5

Joubert, M., & Sutherland, R. (2008). A perspective on the literature: CPD for teachers on mathematics. *National Centre for Excellence in the Teaching of Mathematics*.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educational Psychologist, 41*(2), 75-86. https://doi.org/10.1207/s15326985ep4102_1

Krainer, K. (2011). Teachers as stakeholders in mathematics education research. In B. Ubuz (Ed.), *Proceedings of the 35th conference of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 47-62). Ankara: Middle East Technical University.

Krainer, K., & Wood, T. (2008). *The International Handbook of Mathematics Teacher Education. Participants in Mathematics Teacher Education* (Vol. 3). Rotterdam: Sense.

Kunter, M., & Baumert, J. (2006). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environments Research, 9*(3), 231-251. https://doi.org/10.1007/s10984-006-9015-7

Lipowsky, F. (2004). Was macht Fortbildungen für Lehrkräfte erfolgreich. *Die Deutsche Schule, 96*(4), 462-479.

Lipowsky, F., & Rzejak, D. (2012). Lehrerinnen und Lehrer als Lerner–Wann gelingt der Rollentausch? Merkmale und Wirkungen wirksamer Lehrerfortbildungen. *Schulpädagogik heute, 3*(5), 1-17.

Loucks-Horsley, S., Stiles, K. E., Mundry, S., Love, N., & Hewson, P. W. (2009). *Designing professional development for teachers of science and mathematics*: Corwin Press.

Maass, K. (2011). How can teachers' beliefs affect their professional development? *ZDM, 43*(4), 573-586. https://doi.org/10.1007/s11858-011-0319-4

Maass, K., & Artigue, M. (2013). Implementation of inquiry-based learning in day-to-day teaching: a synthesis. *ZDM, 45*(6), 779-795. https://doi.org/10.1007/s11858-013-0528-0

Maass, K., & Doorman, M. (2013). A model for a widespread implementation of inquiry-based learning. *ZDM, 45*(6), 887-899. https://doi.org/10.1007/s11858-013-0505-7

Minner, D. D., Levy, A. J., & Century, J. (2010). Inquiry-based science instruction—what is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching, 47*(4), 474-496. https://doi.org/10.1002/tea.20347

Muthén, L. K., & Muthén, B. O. (1998-2010). *Mplus Users' Guider's guide* (Vol. 4). Los Angeles, CA: Muthén&Muthén.

National Research Council (NRC). (1996). *National Science Education Standards*. Washington, DC: National Academy Press.

National Research Council (NRC). (2000). *Inquiry and the National Science Education Standards: A Guide for Teaching and Learning*

OECD. (2009a). *Creating Effective Teaching and Learning Environments: First Results from TALIS*: OECD Publishing.

OECD. (2009b). *Technical Report- PISA 2006*. Paris: OECD Publishing.

Pehkonen, E., & Törner, G. (1996). Mathematical beliefs and different aspects of their meaning. *Zentralblatt für Didaktik der Mathematik, 28*(4), 101-108.

Pehkonen, E., & Törner, G. (1999). Teachers' Professional Development: what are the key change factors for mathematics teachers? *European Journal of Teacher Education, 22*(2-3), 259-275.

Prince, M., & Felder, R. (2007). The many faces of inductive teaching and learning. *Journal of College Science Teaching*.

Putnam, R. T., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? *Educational researcher, 29*(1), 4-15. https://doi.org/10.3102/0013189X029001004

Roesken-Winter, B., Hoyles, C., & Blömeke, S. (2015). Evidence-based CPD: Scaling up sustainable interventions. *ZDM, 47*(1), 1-12. https://doi.org/10.1007/s11858-015-0682-7

Roesken-Winter, B., Schüler, S., Stahnke, R., & Blömeke, S. (2015). Effective CPD on a large scale: examining the development of multipliers. *ZDM, 47*(1), 13-25. https://doi.org/10.1007/s11858-014-0644-5

Savery, J. R. (2006). Overview Of PBL: Definition and Distinctions. *The Interdisciplinary Journal of Problem-based Learning, 1*(1), 57-75. https://doi.org/10.7771/1541-5015.1002

Schafer, J. L., & Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychological methods, 7*(2), 147-177. https://doi.org/10.1037/1082-989X.7.2.147

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of psychological research online, 8*(2), 23-74.

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics *Handbook of research on mathematics teaching and learning* (pp. 334-370).

Schoenfeld, A. H. (2010). *How we think: A theory of goal-oriented decision making and its educational applications*. Nwe York: Routledge.

Shulman, L. S. (1986). Paradigms and research programs in the study of teaching: A contemporary perspective. In M. C. Wittrock (Ed.), *Handbook of research in teaching* (pp. 3-36). New York: Macmillan.

Skott, J. (2013). Understanding the role of the teacher in emerging classroom practices: Searching for patterns of participation. *ZDM, 45*(4), 547-559. https://doi.org/10.1007/s11858-013-0500-z

Sörbom, D. (1975). Detection of correlated errors in longitudinal data. *British Journal of Mathematical and Statistical Psychology, 28*(2), 138-151. https://doi.org/10.1111/j.2044-8317.1975.tb00558.x

Staver, J. R., & Bay, M. (1987). Analysis of the Project Synthesis Goal Cluster Orientation and Inquiry Emphasis of Elementary Science Textbooks. *Journal of Research in Science Teaching, 24*(7), 629-643. https://doi.org/10.1002/tea.3660240704

Steyer, R., Eid, M., & Schwenkmezger, P. (1997). Modeling true intraindividual change: True change as a latent variable. *Methods of Psychological Research Online, 2*(1), 21-33.

Stipek, D. J., Givvin, K. B., Salmon, J. M., & MacGyvers, V. L. (2001). Teachers' beliefs and practices related to mathematics instruction. *Teaching and teacher education, 17*(2), 213-226. https://doi.org/10.1016/S0742-051X(00)00052-4

Swan, M. (2005). *Improving learning in mathematics: Challenges and strategies*. Sheffield: Teaching and Learning Division, Department for Education and Skills Standards Unit.

Swan, M. (2007). The impact of task-based professional development on teachers' practices and beliefs: A design research study. *Journal of Mathematics Teacher Education, 10*(4-6), 217-237. https://doi.org/10.1007/s10857-007-9038-8

Swan, M., Pead, D., Doorman, M., & Mooldijk, A. (2013). Designing and using professional development resources for inquiry-based learning. *ZDM, 45*(7), 945-957. https://doi.org/10.1007/s11858-013-0520-8

Swan, M., & Swain, J. (2010). The impact of a professional development programme on the practices and beliefs of numeracy teachers. *Journal of further and Higher Education, 34*(2), 165-177. https://doi.org/10.1080/03098771003695445

Tirosh, D., & Graeber, A. O. (2003). Challenging and changing mathematics teaching classroom practices *Second international handbook of mathematics education* (pp. 643-687): Springer.

Walker, M. D. (2007). *Teaching inquiry-based science - A guide for middle and high school teachers*. LaVergne, TN: Lightning Source.

Weiss, I. R., Pasley, J. D., Smith, P. S., Banilower, E. R., & Heck, D. J. (2003). Looking inside the classroom. *Chapel Hill, NC: Horizon Research Inc*.

Zehetmeier, S., & Krainer, K. (2011). Ways of promoting the sustainability of mathematics teachers' professional development. *ZDM, 43*(6-7), 875-887. https://doi.org/10.1007/s11858-011-0358-x