Progress in Science Education, 2020
2020, Vol 3, No. 2, 1-9
ISSN 2405-6057
DOI.10.25321/prise.2020.967

Special Issue: Proceedings of the 11th DiNat Forum 2020

PROFESSIONAL DEVELOPMENT IN THE SCHOOL INSTITUTE "DISCOVERY EXPERIMENTATION" – FRAMEWORK AND FIRST RESULTS

Markus Emden[1], Armin Baur[2], Arne Bewersdorff[2]

[1]Zurich University of Teacher Education, Switzerland, markus.emden@phzh.ch (ORCID: 0000-0002-7111-3673)
[2]Heidelberg University of Education, Germany, baur@ph-heidelberg.de (ORCID: 0000-0001-5950-7613) and bewersdorff@ph-heidelberg.de (ORCID: 0000-0002-9725-268X)

STRUCTURED ABSTRACT

Background: Professional development (PD) in science education is understood to be influenced by personal dispositions as well as by the quality of formal learning opportunities for teachers. Some beneficial promotors of PD can be identified: duration of a PD programme, active learning of participants, content focus, coherence, collective participation. Respecting these promotors in PD programmes is expected to favourably influence teachers' professional knowledge, their beliefs about teaching, their teaching practices, and – in extension – student achievement. All these aspects (development of professional competence, promotors of PD, relevant goal variables of PD) can be merged into a coherent framework that can inform empirical studies as well as the design of PD.

Purpose of this study is to check the validity of one of the promotors by contrasting variant settings regarding "active learning" in two formats of PD. In one of these, participants are encouraged to intensively collaborate and coach each other (PD institute) while teachers' progress in the other format (personal PD) is left to their own disposal with the coaching function falling exclusively to the professional developers.

Sample/Setting: Forty-six teachers from eight secondary schools in Baden-Wuerttemberg (Germany) participated in the PD programme (PD institute: $n = 22$, personal PD: $n = 24$). The programme lasted for three consecutive semesters (1.5 academic years). Teachers were introduced to a novel approach to teaching through inquiry: "Discovery Experimentation" as a form of opened experimentation (semester 1). All the teachers were observed twice in their teaching (semesters 1 and 2) which formed the core of subsequent coaching sessions either in the teacher group (PD institute) or individually with professional developers (personal PD). The third semester served as a fade-out phase to still have professional developers available but without intensified personal engagement.

Design and Methods: This is a quasi-experimental study. Quantitative data were surveyed from teachers – over four points of measurement – on their pedagogical content knowledge (PCK) and their beliefs about teaching with opened experimentation. Paper-and-pencil-tests and -questionnaires prove to survey reliably (PCK: $\alpha = .853$, beliefs: average from four subscales $\alpha = .738$). Most teachers were video-taped twice (semesters 1 and 2); this is the focus of a separate video study on teaching practices the results of which are pending. Data survey has not yet been completed – thus, the reported data are provisional allowing, nonetheless, to identify general trends.

Results: Trends in teachers' developing beliefs about teaching with opened forms of experimentation suggest that the PD can contribute to advancing these. Regarding the experimental conditions, the PD institute appears more promising when it comes to improving an understanding of the significance of opened experimentation, and to decrease inhibitors to implementing opened experimentation. We suggest that this is due to increased discourse amongst teachers in the PD institute. PCK develops positively for the duration of the programme but without remarkable effect.

Conclusions: Professional developers should actively encourage teachers to collaborate and discuss content and implications from a PD programme. Left to their own impetus, teachers can easily miss (if not avoid) the development potentials of a formal learning opportunity. This might, ultimately, render any attempts at PD fruitless.

Keywords: *professional development; inquiry-based learning; experimentation; beliefs, intervention study; quasi experiment*

Received: *March 2020. Accepted: May 2020*

1 INTRODUCTION

Science education is a dynamic phenomenon, it is always *in flux*: be it due to reforms triggered by science education research (e.g., Anderson, 2002; Osborne, 2014), be it to – politically – react to novel challenges in a changing world (e.g., Kolisang, 2013; Moch, 2011; Ostermeier, Prenzel & Duit, 2010), or be it simply to account for innovation in a domain (e.g., McMorran & Warren, 2012; Schulz, 2009). Today's science education might share some features of the science education from past decades (Brotherton & Preece, 1995; DeBoer, 2006; Gagné, 1965) but in other features it is radically different (Köller & Parchmann, 2012; Schecker, 2012).

The "problem" in this is that paradigms in science education tend to change faster than the staff at schools, who are responsible for fulfilling science education's mission. For this reason, programmes of pre-service teacher training are augmented with in-service professional development (PD) that teachers are encouraged – if not compelled – to attend throughout their professional lives.

This is meant to allow schools to "quickly" react to novel challenges and to adopt to these. It is the educational system's lever to preserve a degree of dynamism instead of being obliged to wait for demographic transition of teacher generations for accomplishing change in classrooms. Little is known in how far PD programmes deliver on this expectation (Kunter, Kleickmann, Klusmann & Richter, 2013; Richter, 2011). Some evidence points disappointingly into directions of ineffectualness (Lipowsky, 2014; van Driel, Beijaard & Verloop, 2001, p. 140), some to more hopeful views of realized accountability (Desimone & Garet, 2015; Garet, Porter, Desimone, Birman & Yoon, 2001).

This article introduces the framework, design, and first results of an intervention study in a novel PD format on scientific inquiry as an instructional approach. The framework is derived from earlier research on professional competence (Kunter, Baumert et al., 2013) and on PD research in scientific inquiry settings (Capps & Crawford, 2013; Capps, Crawford & Constas, 2012). The study is a quasi-experimental intervention study investigating the PD's impact on teachers' cognition and attitudes as well as on students' cognition. First results will be presented primarily regarding teacher beliefs and PCK – this is to respect the manuscript's limitations in length as well as the data's provisional character as survey has not yet been completed, some schools are still participating in the programme.

2 FRAMEWORK

The suggested framework derives from two traditions of education research: (1) German COACTIV's research (Cognitive Activation in the Mathematics Classroom) and models of professional competence (Kunter, Baumert et al., 2013); (2) research findings from the US's Dwight D. Eisenhower Professional Development Program (Garet et al., 2001; Garet, Birman, Porter, Desimone & Herman, 1999). The framework merges essential findings from these approaches, thereby extending their primarily descriptive powers to account for more specific correlations. An in-depth introduction to the framework can be found in Emden and Baur (2017).

2.1 Professional Competence and its Development

Starting point for COACTIV is their model of professional competence (Baumert & Kunter, 2013). Teachers' professional competence has, for a long time, been viewed from two polar perspectives: competence as a natural gift vs. competence as acquired expertise – so, at one extreme, one would be "doomed" to be a good/bad teacher, at the other extreme, "anyone" could become a good teacher. Both extremes cannot satisfy a general idea of teachers-to-be who might build on their personal dispositions to develop into good teachers by studying and reflective practice.

COACTIV takes the latter perspective and assumes that professional competence can be purposefully developed. Professional competence serves as the head term for a complex interrelatedness of: (a) teachers' professional knowledge, (b) their beliefs, values, and goals regarding teaching, (c) their motivational characteristics, and (d) their self-regulation skills (Baumert & Kunter, 2013).

Professional knowledge – in itself a contested term – has been introduced by Shulman (1987). Most current frameworks agree on three distinct aspects of professional knowledge: teachers need (a) content knowledge (CK: know what to teach), (b) pedagogical knowledge (PK: know how to teach), and (c) pedagogical content-knowledge (PCK: know how to teach what) (e.g., Fischer, Borowski & Tepner, 2012). Shulman (1987, p. 8) refers to PCK as "that special amalgam of content and pedagogy that is uniquely the province of the teachers, their own special form of professional understanding". It is this knowledge – e.g., to know why learning about atoms is hard for students and how these difficulties might be overcome – that distinguishes the professional science teacher from the professional science researcher and from the professional pedagogue. For PD in science education, the nexus between CK and PCK appears to be especially crucial, with PK arguably being applicable to teaching generally irrespective of subject.

Developing professional competence is viewed as the result of a process in which a proposal for development is submitted to teachers, who take it up and put it to individual use, which eventually will affect their professional actions. The complete process, however, unfolds before the backdrop of personal, political, and institutional conditions and limitations, each of which can influence the teachers' acceptance and understanding of the proposal (see squared boxes in Fig. 1.; Kunter, Kleickmann et al., 2013). In this understanding, professional competence is to some degree acquired expertise (realization of proposal-uptake) and to some degree "fated" determination (influence of personal and contextual factors).

2.2 Promotors of Professional Development

Garet and colleagues (1999, 2001) investigated structural conditions of PD opportunities for mathematics and science teachers in the Dwight D. Eisenhower Professional Development Program (Improving America's Schools Act, 1994). They could identify five factors that promoted the effectiveness of PD on teacher understanding and ultimately on their practices: PD programmes benefitted from (1) extended duration of the programme (Adey, 2006; Lipowsky, 2014), (2) when active learning of the participants was encouraged, (3) when clear content focus was provided, (4) when content was coherent with teachers' working conditions (Davis, Janssen & van Driel, 2016; Timperley, Wilson, Barrar & Fung, 2007), (5) when participants learned in professional learning communities (van Driel et al., 2001). Arranging these factors in a structural equation model yielded an explained variance on teacher cognition of $R^2 = .517$, and in extension on changed teaching practice ($R^2 = .416$; Garet et al., 2001). These findings correspond with a compilation by Tinoca (2004) and some of the aspects for effective professional learning communities identified by Bolam, McMahon, Stoll, Thomas, and Wallace (2005).

2.3 A Framework for Planning Professional Development

These five promotors of PD can be used to underlay the COACTIV-model of PD (see rounded boxes in Fig. 1): e.g., the beneficial influence of extended programmes becomes self-evident when one realizes that learning takes time as it presupposes reflection and opportunities to discuss (Justi & van Driel, 2005). A "one-off" course is insensitive to this need by design: teachers come, listen, and leave, i.e. they are left to themselves before they can identify potential misunderstandings and problems, there is simply no room provided for clarification that reflection might necessitate. Likewise, active learning has been understood to be prerequisite for understanding since the advent of constructivism (e.g., Hood Cattaneo, 2017). Focussing on content that is coherent with teachers' needs and conditions will make them realize the benefits and, thus, trigger change (Gräsel & Parchmann, 2004). Integrating an innovation into a community of practice, again, requires joint learning and discussion so that learning in professional communities proves favourable almost naturally (Knight, 2002). Lastly, DiBiase (2014, p. 26) highlights that PD that does not "focus on specific content […] has little to no impact on teacher practice". Regarding professional knowledge, this model does not address PK expressly as this component is understood to be more general to teaching, while the model aims at supporting PD with subject-specific content.

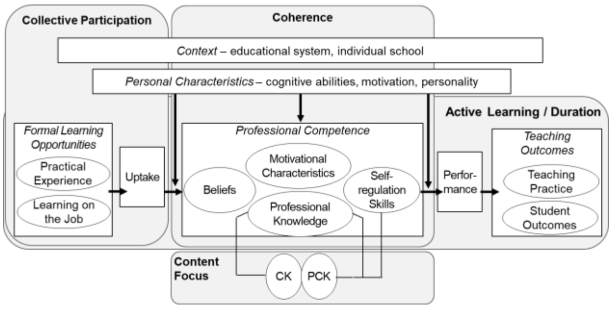

Fig. 1. Design- and Effect-Framework for Professional Development (translated from Emden & Baur, 2017)

3 PROFESSIONAL DEVELOPMENT INSTITUTE "DISCOVERY EXPERIMENTATION"

The framework (Fig. 1) served to design a PD format that aimed at respecting all of the cited promotors. This format was dubbed "professional development institute" to contrast it from a second format realized in the intervention study, the personal PD. These formats were designed as similar to each other as possible with a variation in only one of the promotors (Tab. 1).

Tab. 1. Setting of Promotors in the Study

	PD institute	personal PD
Active learning	lecture-style input plus discussions, development tasks for teachers;	
	collaboration between teachers is encouraged; lesson observations and discussions with colleagues and professional developers	individual work, lesson observations and discussions just with professional developers
Duration	three semesters (induction, supervision, fade-out), each of 0.5 academic years	
Coherence	reference to current curricula, consideration of school's/teacher's wishes and limitations (e.g., CK or infrastructure)	
Content focus	introducing "Discovery Experimentation" (see below)	
Collective Participation	PD formats address teachers of one school (professional learning community)	

3.1 Content Focus "Discovery Experimentation"

Teachers were introduced to a novel concept that aimed at informing their lessons on experimentation (inquiry as "an instructional approach"; Furtak, Seidel, Iverson & Briggs, 2012). Since the advent of German science education standards (KMK, 2005-c) there has been an increasing challenge on teachers to reform their implementation of science experiments from serving illustrative to serving sense-making purposes. Teachers are often unsure how to introduce their students to becoming self-reliant inquirers into nature (Capps, Shemwell & Young, 2016) as they themselves frequently do not know what scientific inquiry means (Furtak et al., 2012; Osborne, 2014) or what it entails (not just hands-on: Hodson, 2014). The proposed concept, first of all, sensitizes teachers to the conceptual difference between experiments and other forms of hands-on-experience by defining that an experiment is: (1) a deliberate investigation of nature, (2) in which variables are identified and controlled, usually including (3) control and (4) test set-ups, which (5) lead to reproducible observations. Teachers are given to understand that other hands-on approaches serve valuable functions, too, but that one must be precise when referring to an experiment.

The novel concept suggests teachers to consider five aspects when preparing to teach experimentation in their classes (Emden & Baur, 2017): (1) students execute the experiments, (2) they are made aware that inquiry is a structured process (e.g., question-hypothesis-investigation-conclusion), (3) they reflect the inner logic of this sequence, (4) they discover something "new" for them, (5) they are led gently to be fully responsible for the process (opening experimentation). Teachers are introduced to a matrix that can help them fading students into the inquiry process (Baur & Emden, 2020).

3.2 Sample, Design, Methods

The PD programme was administered in three overlapping waves to teachers from eight secondary schools (Tab. 2 and Fig. 2). Teachers of a school were assigned to either the PD institute- or the personal PD-condition based on their own estimation of how encultured collaboration was at their school. This approach to assigning experimental conditions was chosen to account for the probability that, if teachers who routinely collaborate with each other were assigned randomly to the personal PD condition, they could possibly not refrain from collaborating, nonetheless. At the same time, it is acknowledged that this might bias findings to some degree.

Tab. 2. Cross Table detailing the study's sample

	female	male	total
PD institute	21	1	22
personal PD	17	7	24
	38	8	46

Teachers were introduced to the concept, its background and practical suggestions in three 3-4 h-workshops over the course of 0.5 academic years (first semester). Twice, their lessons were observed and discussed in either group or personal consultations in the first and second semesters; the third semester served as a fade-out phase in which professional developers still were available but did not engage personally on a regular basis.

Data were gathered with respect to the change in four variables that Capps and Crawford (2013) have claimed to be decisive regarding the effectiveness of PD and that have not been surveyed coherently before (Capps et al., 2012): (1) teachers' cognition, (2) teachers' beliefs, (3) teachers' lesson practice, (4) students' cognition. Times of data sampling can be read from Figure 2. The survey is ongoing and, thus, portions of analyses are pending at the time of this report.

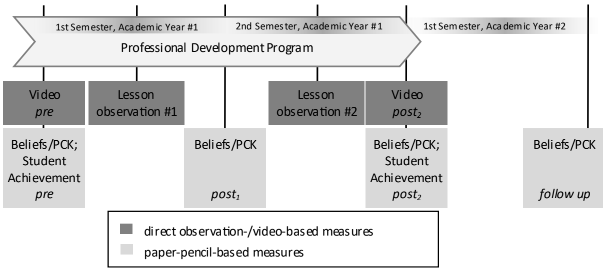

Fig. 2. Study Design

The design is meant to answer these research questions:

Does PD which is based on the identified promotors generally improve beliefs and PCK?

Does the variation in one of the promotors (active learning: collaboration vs. personal) lead to differential changes in beliefs and PCK?

Data for PCK and beliefs were surveyed with paper-and-pencil tests – the same tests and questionnaires were administered in each instance. Existing instruments were adapted for this purpose: Teachers' PCK on experimentation was surveyed with a Vignette-test (Schmitt, 2016) in which teachers ranked the appropriateness of teaching situations related to experimentation. Their rankings were scored with regard to aggregated expert rankings. Rankings that resembled the expert ranking scored higher than divergent rankings (Schmitt, 2016). The test relies on a three-dimensional test model (Tepner et al., 2012) which differentiates between (1) facets of PCK (here: inquiry-based experimentation), (2) types of PCK (declarative knowledge elucidating "What is [facet]?", procedural knowledge: "How to teach [facet]", conditional knowledge: "Which conditions need to be met with my students before some aspect of [facet] may be addressed?"), and (3) contexts of application (e.g., experiments on redox- or acid-base-reactions). The test has been in development and use since 2012 with Schmitt (2016) giving its reliability at $\alpha = .75$ and Anthofer (2017) at $\alpha = .77$, respectively. Even if the test investigates PCK on inquiry-based experimentation (not on "Discovery Experimentation"), it is assumed that Discovery Experimentation is sufficiently rooted in scientific inquiry as to detect changes in teachers' PCK due to the PD.

Teacher beliefs on inquiry-based learning were surveyed with items from the teacher questionnaire of project PRIMAS (Promoting Inquiry-based Learning in Mathematics and Science Education Across Europe; Engeln, Euler & Maass, 2013), which introduce personal statements regarding teachers' beliefs concerning inquiry-based learning. Teachers rate these statements on a four-point Likert scale ranging from 1 (disagree completely) to 4 (agree completely). For the present study, item packages 12-15 from the teachers' questionnaire were selected (Engeln, 2013) – item packages 13 and 14 were adapted to focus on opened experimentation to account for the PD's content focus on "Discovery Experimentation".

Regarding teacher practice, the development and evaluation of a coding scheme for lesson-videos is subject of a PhD-thesis (Bewersdorff, Baur & Emden, 2020); student performance was surveyed with an abbreviated version of Glug's (2009) achievement-test – students worked on the same test forms in both surveys. This presentation of first results will focus on teachers' beliefs and PCK as other analyses are pending.

4 RESULTS

(I) Beliefs: In order to investigate the test instrument's psychometric quality, data of all available questionnaires were combined irrespective of time of sampling (N = 144).

The items that were adopted from project PRIMAS yield four subscales (Tab. 3), which in their majority show acceptable if not good internal consistency values (Cronbach's α) – concessions have to be made with regard to "Significance of Opened Experimentation for Science Teaching and Learning" which often yields values .50 < α < .70, still allowing for group comparisons. In general, the questionnaire appears to measure reliably and validly – assuming the test developers attended to the latter aspect diligently. Objectivity is ensured by resorting to Likert scale items.

Tab. 3. Cronbach's α-reliabilities for Belief subscales

	global	pre	post1	post2	follow-up
Importance Assigned to PD (12 items)	*n*=139, *α*=.776	*n*=44, *α*=.805	*n*=34, *α*=.795	*n*=33, *α*=.733	*n*=28, *α*=.780
Significance of Opened Experimentation (8 items)	*n*=140, *α*=.665	*n*=44, *α*=.658	*n*=34, *α*=.523	*n*=32, *α*=.654	*n*=30, *α*=.857
Inhibitors to Implement Opened Experi-mentation (15 items)	*n*=132, *α*=.776	*n*=42, *α*=.797	*n*=34, *α*=.806	*n*=31, *α*=.769	*n*=25, *α*=.612
Self-Reported Imple-mentation of Inquiry-Based-Learning Elements (32 items)	*n*=130, *α*=.734	*n*=41, *α*=.728	*n*=31, *α*=.733	*n*=32, *α*=.750	*n*=26, *α*=.708

4.1 Development of Beliefs

In order to test for the PD's general potential to have an influence on teachers' beliefs, ANOVA with repeated measures was calculated. Keeping in mind that the data and results are provisional, potentially significant differences must not be overstressed, as neither must be missing significance.

Results for the aggregated sample can be read from Table 4. There is a general trend that the PD affects teachers' beliefs favourably regarding the Significance of Opened Experimentation and concerning Inhibitors to Implement Opened Experimentation. While we cannot conclusively rule out that the observed development might have different causes (as there is no control or waiting group), we find this scenario unreasonable to assume.

Tab. 4. Repeated measures ANOVA data for the four beliefs subscales (sphericity may not be assumed; Greenhouse-Geisser corrections are reported)

	df (time)	*df* (error)	*F*	*p*
Importance Assigned to PD	2.322	62.683	.524	.622
Significance of Opened Experimentation	2.641	71.306	4.412	.009
Inhibitors to Implement Opened Experimentation	2.591	69.955	5.170	.004
Self-Reported Imple-mentation of Inquiry-Based-Learning Elements	2.724	73.557	2.465	.075

Regarding the second research question, the four belief scales were analysed with repeated measures ANOVA juxtaposing both the experimental conditions – PD institute vs. personal PD. As can be read from the line plots (Fig. 3–6), two of the four subscales' developments favour our hypotheses: (1) an awareness of the significance of experimentation increases in the institute and remains high, but almost levels in personal PD (Fig. 4); (2) the inhibitors to experimentation in classrooms decrease for the PD institute and remain low, but remain almost level for personal PD (Fig. 5). Regarding the importance assigned to PD not much changes in either condition (Fig. 3); considering self-reported implementation of inquiry-based learning elements both conditions show a parallel rise which drops markedly for personal PD in the follow-up survey (Fig. 6) – no levels of significance are breached.

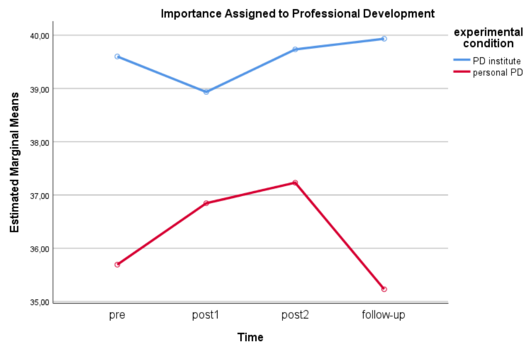

Fig. 3. Importance Assigned to PD (*n* = 28; no significant Within-Subjects Effects (WSE); Between-Subjects Effects (BSE): $F(1,26) = 9.519$, $p = .005$)

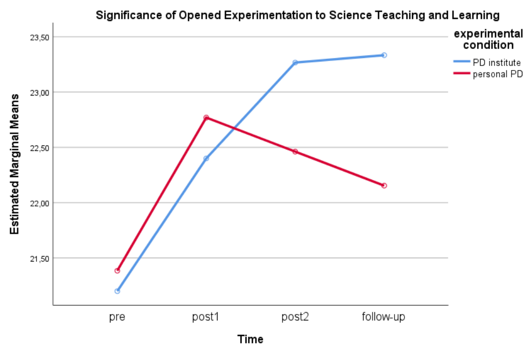

Fig. 4. Significance of Opened Experimentation for Science Teaching and Learning (*n* = 28; WSE: Main effect for Time, $F(2.667,69.352) = 4.223$; $p = .011$; no significant BSE)

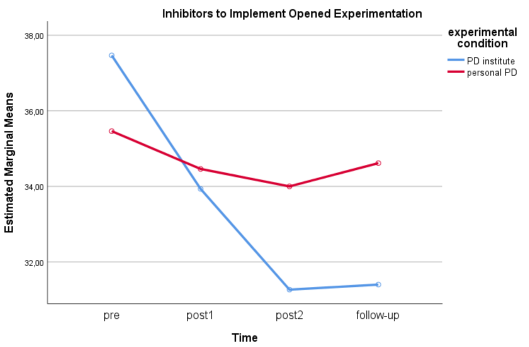

Fig. 5. Inhibitors to Implement Opened Experimentation (n = 28; WSE: Main effect for Time, $F(2.722,70.768) = 4.914$, $p = .005$; no significant BSE)

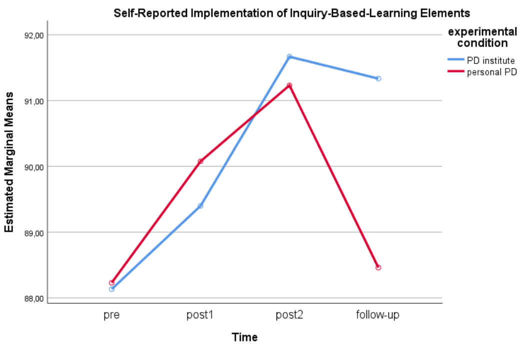

Fig. 6. Self-Reported Implementation of Inquiry-Based Learning Elements (n = 28; no significant WSE, nor BSE)

(II) PCK: Current findings on the development of teachers' PCK point into promising directions with the instrument proving to measure PCK reliably ($\alpha = .857$, over all points of measurement). Increases in PCK are detectable through repeated measures ANOVA and are significant between the pre- and post2-points of measurement (see. Tab. 5; $F(2,52) = 3.611$, $p = .034$) but prove to be unstable in the long run ($F(3,69) = 2.152$, $p = .102$). Data cannot yet support an assumed superiority of the PD institute in developing PCK ($F(1,22) = .159$, $p = .693$).

Tab. 5. Descriptive Statistics for PCK, after exclusion of three outliers with $Score_{fu}–Score_{pre}<–30$

Score PCK (theor. max. score: 100)	*n*	min.	max.	Mean	*SD*
pre	42	21.98	80.60	64.87	13.82
post$_1$	31	22.41	89.66	67.14	15.07
post$_2$	29	43.10	84.48	67.61	10.76
follow up	26	25.86	82.76	65.75	12.84

(III) Lesson practice and students' cognition: Regarding teachers' practice, results are pending due to extensive coding – here, too, measures of inter-rater reliability are suggestive of a valid and reliable instrument (Bewersdorff, Baur & Emden, 2020). Unfortunately, the student test for inquiry (Glug, 2009) does not survey student performance reliably, which cannot be excused by too small a sample (n = 550) but might be an effect of the adaptation process.

4.2 Limitations and Discussion

The study's character as work in progress and its small sample sizes speak clearly of its limitations. With 46 teachers in a non-randomized quasi-experimental design, we cannot rule out that, e.g., personal motivation to participate in a PD plays an interfering part. With respect to those teachers who claim to routinely collaborate and who were, thus, assigned to the PD institute there remains a probability of bias favouring the PD institute. Moreover, small sample sizes typically incriminate the reliability readings of tests; it stands to be investigated, how well each of the instruments performs for each of the sampling points separately, i.e., with smaller sample sizes. Considering developments in beliefs and PCK, so far, absolute gains/losses have been investigated. However, should there already be substantial differences between the teachers' results for the pre-surveys, resorting to residual gains/losses might be more informative. These analyses will be conducted once the data collection is completed.

Considering all justified reservations and caveats towards these preliminary results, it appears that the PD institute can develop teachers' professional competence in some of the relevant variables in the expected direction.

Teachers' attitudes on the significance of opened experimentation for teaching and learning in science education develop more favourably in the PD institute than in the group that is largely left to themselves to reflect the PD content and the professional developers' coaching. We argue that more active learning through collaboration ensures that the abstract concept is checked for its curricular and ecological validity through teacher conversation. Teachers are thus enabled to identify and resolve potential transfer problems more easily than teachers who do not share these sense-making processes with their colleagues. Consequently, teachers from the PD institute increasingly abolish their views on potential inhibitors to experimentation. They realize that experimentation can be incorporated into regular classes as their peers show them probate solutions and "work-arounds" to arising challenges. All these aspects remain largely unchanged for the personal PD-group which suggests that these teachers do not involve themselves as thoroughly with the PD's aims and concepts. This argument appears to be well aligned with other research on professional learning communities (Bolam et al., 2005; Bonsen & Rolff, 2006) and collaborative PD (Gräsel, Fußangel & Parchmann, 2006; Gräsel, Pröbstel, Freienberg & Parchmann, 2006).

To find that teachers in both experimental groups increasingly feel that they incorporate aspects of inquiry-based learning might be attributed to the PD's aim of sensitizing teachers to the concept. Some degree of social desirability in the teachers' response patterns must, however, not be dismissed prematurely.

Regarding the development of PCK, the jury is still out at the time of this report. The tendencies that surface are sobering and need to be investigated in more detail – a less than moderate increase (max ≈ 3%) in PCK might put the effort in administering the PD into a harsh perspective. For the final evaluation, a complete data set is mandatory as each response contributes to the final choice of items from the PCK-test (Anthofer, 2017; Schmitt, 2015); it is hoped to ultimately arrive at a test that sensitively detects relevant changes in PCK. Conclusive analyses will commence after data collection is complete.

5 CONCLUSION AND IMPLICATIONS

Science teaching needs to adapt to ever changing challenges; it needs to do so rather quickly and cannot wait for changes in teacher education to slowly trickle down into the classroom. For this reason, effective PD programmes are needed to bring in-service teachers up to speed with science education theory. What little research has been done in this field (e.g., Hofmann, 2015; Schmitt, 2016) consistently points into the direction of what we have merged in the suggested framework to inform PD in science education (Fig. 1).

Considering one of the promotors of effectiveness of a PD through shades of "active learning", first results promise that there can be substantial benefit from strengthening participants' active involvement, also in delivering a PD by lesson observations and coachings in groups. The framework encourages to investigate other promotors in a similar fashion: One could contrast a PD that addresses scientific inquiry as a meta-method with one that clearly focuses on one exemplary method (e.g., experimentation) (promotor Content Focus). Similarly, PD programmes that suggest "one-size-fits-all" instead of respecting local conditions, e.g., lacking equipment, should be investigated (Coherence).

In any case, it appears imperative to find an instrument that reliably, economically and ecologically surveys student performance in experimentation. This is indispensable if the complete impact of a PD is to be estimated – in this assessment, we whole-heartedly concur with Capps et al. (2012).

ACKNOWLEDGEMENT

We are indebted to the German Research Association (DFG) who have funded this research under reference number BA 5653|2-1, EM 267|2-1.

REFERENCES

Improving America's Schools Act of 1994. Title II - Dwight D. Eisenhower Professional Development Program (1994). Public Law 103-382. Washington, D.C.: 103rd United States Congress.

Adey, P. (2006). A model for the professional development of teachers of thinking. Thinking Skills and Creativity, 1(1), 49–56.

Anderson, R. D. (2002). Reforming Science Teaching: What Research Says About Inquiry. Journal of Science Teacher Education, 13(1), 1–12.

Anthofer, S. (2017). Förderung des fachspezifischen Professionswissens von Chemielehramtsstudierenden. Berlin: Logos.

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 (pp. 25–48). Boston, MA: Springer US.

Baur, A. & Emden, M. (2020). How to open inquiry teaching? An alternative teaching scaffold to foster students' inquiry skills. Chemistry Teacher International, 2, 1–12.

Bewersdorff, A., Baur, A. & Emden, M. (2020). Untersuchung der Wirksamkeit einer Lehrkräftefortbildung zum Experimentieren. In S. Habig (Ed.), Naturwissenschaftliche Kompetenzen in der Gesellschaft von morgen (pp. 860–863). Essen: Universität Duisburg-Essen.

Bolam, R., McMahon, A., Stoll, L., Thomas, S. & Wallace, M. (2005). Creating and sustaining effective professional learning communities (Research report / Department for Education and Skills No. RR637). Annesley.

Bonsen, M. & Rolff, H.-G. (2006). Professionelle Lerngemeinschaften von Lehrerinnen und Lehrern. Zeitschrift Für Pädagogik, 52(2), 167–184.

Brotherton, P. N. & Preece, P. F.W. (1995). Science Process Skills: Their Nature and Interrelationships. Research in Science & Technological Education, 13(1), 5–11.

Capps, D. K. & Crawford, B. A. (2013). Inquiry-Based Professional Development and Nature of Science? What does it take to support teachers in learning about inquiry? International Journal of Science Education, 35(12), 1947–1978.

Capps, D. K., Crawford, B. A. & Constas, M. A. (2012). A Review of Empirical Literature on Inquiry Professional Development: Alignment with Best Practices and Critique of the Findings. Journal of Science Teacher Education, 23(3), 291–318.

Capps, D. K., Shemwell, J. T. & Young, A. M. (2016). Over reported and misunderstood? A study of teachers' reported enactment and knowledge of inquiry-based science teaching. International Journal of Science Education, 38(6), 934–959.

Davis, E. A., Janssen, F. J. J. M. & van Driel, J. H. (2016). Teachers and science curriculum materials:

Where we are and where we need to go. Studies in Science Education, 52(2), 127–160.

DeBoer, G. E. (2006). Historical Perspectives on Inquiry Teaching in Schools. In L. B. Flick & N. G. Lederman (Eds.), Scientific Inquiry and Nature of Science: Implications for Teaching, Learning, and Teacher Education (pp. 17–35). Dordrecht, Boston, London: Kluwer Academic Publishers.

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.

DiBiase, D. (2014). Formative Assessment Professional Development: Impact on Teacher Practice (Dissertation). Johnson & Wales University, Providence, RI. http://search.proquest.com/docview/1548006514

Emden, M. & Baur, A. (2017). Effektive Lehrkräftebildung zum Experimentieren: Entwurf eines integrierten Wirkungs- und Gestaltungsmodells. Zeitschrift für Didaktik der Naturwissenschaften, 23(1), 1–19.

Engeln, K. (2013). IBL implementation survey report: The PRIMAS project: Promoting inquiry-based learning (IBL) in mathematics and science education across Europe. Deliverable No. 9.3. Retrieved from https://primas-project.eu/wp-content/uploads/sites/323/2017/11/PRIMAS_D-9.3_IBL-Implementation-survey-report.pdf

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 Mathematics Education, 45(6), 823–836.

Fischer, H. E., Borowski, A. & Tepner, O. (2012). Professional Knowledge of Science Teachers. In B. J. Fraser, K. Tobin, & C. J. McRobbie (Eds.), Second International Handbook of Science Education (pp. 435–448). Dordrecht: Springer Netherlands

Furtak, E. M., Seidel, T., Iverson, H. & Briggs, D. C. (2012). Experimental and Quasi-Experimental Studies of Inquiry-Based Science Teaching: A Meta-Analysis. Review of Educational Research, 82(3), 300–329.

Gagné, R. M. (1965). The Psychological Bases of Science - A Process Approach (AAAS Miscellaneous Publication No. 65-8). Washington, D.C.

Garet, M. S., Birman, B. F., Porter, A. C., Desimone, L. M. & Herman, R. (1999). Designing Effective Professional Development: Lessons from the Eisenhower Program (No. ED/OUS99-3). Washington, D.C.

Garet, M. S., Porter, A. C., Desimone, L. M., Birman, B. F. & Yoon, K. S. (2001). What Makes Professional Development Effective? Results From a National Sample of Teachers. American Educational Research Journal, 38(4), 915–945.

Glug, I. (2009). Entwicklung und Validierung eines Multiple-Choice-Tests zur Erfassung prozessbezogener naturwissenschaftlicher Grundbildung. (Dissertation). Christian-Albrechts University, Kiel. Retrieved from http://eldiss.uni-kiel.de/macau/receive/dissertation_diss_00003649

Gräsel, C., Fußangel, K. & Parchmann, I. (2006). Lerngemeinschaften in der Lehrerfortbildung: Kooperationserfahrungen und -überzeugungen von Lehrkräften. Zeitschrift Für Erziehungswissenschaft, 9(4), 545–561.

Gräsel, C. & Parchmann, I. (2004). Implementationsforschung: oder: der steinige Weg, Unterricht zu verändern. Unterrichtswissenschaft, 32(3), 196-214.

Gräsel, C., Pröbstel, C., Freienberg, J. & Parchmann, I. (2006). Anregung zur Kooperation von Lehrkräften im Rahmen von Fortbildungen. In M. Prenzel & L. Allolio-Näcke (Eds.), Untersuchung zur Bildungsqualität von Schule (pp. 310–329). Münster: Waxmann.

Hodson, D. (2014). Learning Science, Learning about Science, Doing Science. Different goals demand different learning methods. International Journal of Science Education, 36(15), 2534–2553.

Hofmann, J. (2015). Untersuchung des Kompetenzaufbaus von Physiklehrkräften während einer Fortbildungsmaßnahme. Berlin: Logos.

Hood Cattaneo, K. (2017). Telling Active Learning Pedagogies Apart: From theory to practice. Journal of New Approaches in Educational Research, 6(2), 144–152.

Justi, R. & van Driel, J. H. (2005). The development of science teachers' knowledge on models and modelling: promoting, characterizing, and understanding the process. International Journal of Science Education, 27(5), 549–573.

Knight, P. (2002). A Systemic Approach to Professional Development: Learning as Practice. Teaching and Teacher Education, 18(3), 229–241.

Kolisang, C. (Ed.) (2013). Research. Bundesweiter Bildungsstreik 2009: Protestbewegung - Aktionismus - Reformen der Reformen. Wiesbaden: Springer VS.

Köller, O. & Parchmann, I. (2012). Competencies: The German notion of learning outcomes. In S. Bernholt, K. Neumann & P. Nentwig (Eds.), Making it tangible: Learning outcomes in science education (pp. 151–171). Münster: Waxmann.

Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S. & Neubrand, M. (Eds.) (2013). Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers. Boston, MA: Springer US

Kunter, M., Kleickmann, T., Klusmann, U. & Richter, D. (2013). The Development 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 (pp. 63–77). Boston, MA: Springer US.

Lipowsky, F. (2014). Theoretische Perspektiven und empirische Befunde zur Wirksamkeit von Lehrerfort- und -weiterbildung. In E. Terhart, H. Bennewitz & M. Rothland (Eds.), Handbuch der Forschung zum Lehrerberuf (pp. 511–541). Münster: Waxmann.

McMorran, D. & Warren, D. (2012). Taking Chemistry Out of the Lab: Perspectives on Chemistry Outreach at Otago. Chemistry in New Zealand, 76(2), 56–61.

Moch, P. (2011). The Race to Space: Launching Fear and Blasting Off to Reform. In M. L. Kysilka (Ed.), Studies in the history of education. Critical times in curriculum thought: People, politics, and perspectives (pp. 165–179). Charlotte, N.C: Information Age Pub.

Osborne, J. (2014). Teaching Scientific Practices: Meeting the Challenge of Change. Journal of Science Teacher Education, 25(2), 177–196.

Ostermeier, C., Prenzel, M. & Duit, R. (2010). Improving Science and Mathematics Instruction: The SINUS Project as an example for reform as teacher professional development. International Journal of Science Education, 32(3), 303–327.

Richter, D. (2011). Lernen im Beruf. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Eds.), Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV (pp. 317–325). Münster: Waxmann.

Schecker, H. (2012). Standards, competencies and outcomes: A critical view. In S. Bernholt, K. Neumann & P. Nentwig (Eds.), Making it tangible: Learning outcomes in science education (pp. 219–234). Münster: Waxmann.

Schmitt, A. K. (2016). Entwicklung und Evaluation einer Chemielehrerfortbildung zum Kompetenzbereich Erkenntnisgewinnung. Berlin: Logos.

Schulz, R. M. (2009). Reforming Science Education: Part I. The Search for a Philosophy of Science Education. Science & Education, 18(3-4), 225–249.

Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland [KMK] (2005a). Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss. München: Luchterhand.

(2005b). Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss. München: Luchterhand.

(2005c). Bildungsstandards im Fach Physik für den Mittleren Schulabschluss. München: Luchterhand.

Shulman, L. S. (1987). Knowledge and Teaching: Foundations of the New Reform. Harvard Educational Review, 57(1), 1–22.

Tepner, O., Borowski, A., Dollny, S., Fischer, H. E., Jüttner, M., Kirschner, S., . . . Wirth, J. (2012). Modell zur Entwicklung von Testitems zur Erfassung des Professionswissens von Lehrkräften in den Naturwissenschaften. Zeitschrift für Didaktik der Naturwissenschaften, 18, 7–28.

Timperley, H., Wilson, A., Barrar, H. & Fung, I. (2007). Teacher professional learning and development: Best evidence synthesis iteration (BES). Wellington, N.Z.: Ministry of Education.

Tinoca, L. F. (2004). From Professional Learning for Science Teachers to Student Learning in Science (Dissertation). University of Texas, Austin, TX. Retrieved from http://www.lib.utexas.edu/etd/d/2004/-tinocalf042/tinocalf042.pdf

Van Driel, J. H., Beijaard, D. & Verloop, N. (2001). Professional Development and Reform in Science Education: The Role of Teachers’ Practical Knowledge. Journal of Research in Science Teaching, 38(2), 137–158.