Zitiervorschlag: Smit, R., Rietz, F., & Kreis, A. (2018). What are the effects of science lesson planning in peers? Analysis of attitudes and knowledge based on an actor–partner interdependence model. Research in Science Education, 48(3), 619-636. <https://doi.org/10.1007/s11165-016-9581-3>

Zur Verfügung gestellt auf PHIQ:

PHIQ-DOI: <https://doi.org/10.18747/PHSG-coll3/id/257>
Original-DOI: <https://doi.org/10.1007/s11165-016-9581-3>

Dokumentart: Journal Article

Version: accepted version

Copyright-Hinweis: This is a post-peer-review, pre-copyedit version of an article published in Research in Science Education. The final authenticated version is available online at: <https://doi.org/10.1007/s11165-016-9581-3>

Lizenz: Alle Rechte vorbehalten

What Are the Effects of Science Lesson Planning in Peers? Analysis of Attitudes and Knowledge Based on an Actor–Partner Interdependence Model

DOI 10.1007/s11165-016-9581-3

Robbert Smit¹ & Florian Rietz¹ & Annelies Kreis²

Abstract

This study focuses on the effects of collaborative lesson planning by science pre-service teachers on their attitudes and knowledge. In our study, 120 pre-service teachers discussed a preparation for a science inquiry lesson in dyads. The teacher with the lesson preparation had the role of the coachee, while the other was the coach. We investigated the following research questions: (1) Does learning occur between the two peers? and (2) Is the competency in lesson planning affected by the attitude and knowledge of coach and coachee? Based on an actor–partner interdependence model (APIM), we could clarify the relations of pedagogical content knowledge (PCK) and attitudes (ATT) between and within the dyads of coach and coachee, as well as their development over time. Furthermore, the APIM allowed the inclusion of a mediator (lesson planning competency). Both PCK and ATT increased slightly but significantly during our project. ATT and PCK seemed to converge between coach and coachee at the end of the project. However, we could not find any cross-lagged effects, meaning there was no effect of coach on coachee or vice versa over time. Further, preceding PCK showed a significant effect on the competency of lesson planning, but planning competency did not influence succeeding PCK or attitude. Finally, these results are discussed with respect to science teacher education.

Keywords: Peer coaching; Dyadic analysis; Pre-service teachers; Scientific inquiry teaching

1 University of Teacher Education, St. Gallen, Rorschach, Switzerland

2 Institute for Education, University of Zurich, Zurich, Switzerland

Introduction

Science education is currently going through a process of change (Neumann et al. 2010; Van Driel et al. 2001). One objective is to make students aware of the ways in which scientific knowledge is produced and developed, that is, to promote students' understanding of the nature of science. From a teaching perspective, this implicates a shift towards the teaching of inquiry skills, which is definitely more complex than the traditional training of practical skills (Schwartz et al. 2004; Van Driel et al. 2001). Teachers' beliefs and attitudes are deemed important for successful curriculum reform (Bryan 2012). Beginning teachers encounter several constraints when they start teaching science. Depending upon their prior beliefs, knowledge, and understanding, which may or may not have been reinforced during teaching education programmes, such constraints affect their teaching of inquiry (Roehrig and Luft 2004). To support beginning teachers in implementing science as inquiry, induction programmes that attend to these constraints are needed. For example, workshops focusing on inquiry-focused teaching strategies can offer pedagogical knowledge for those teachers lacking such tools for their preparation programmes (Luft et al. 2003).

As part of our study KUBEX (content-focused peer coaching in pre-service teacher education for teaching scientific inquiry), 120 pre-service teachers for secondary school first received short but ecologically valid training on science inquiry teaching. Ecological validity is the extent to which research findings would generalize to settings typical of everyday life (Wegener and Blankenship 2007). This means in our case that for many topics in pre-service teacher education in reality there is only limited training time. We could show that pre-service teachers' attitudes towards scientific inquiry teaching developed during the training programme (Smit 2015). The second part of the research design consisted of a collaborative lesson planning session between two pre-service teachers. One pre-service teacher (coachee) prepared a lesson plan before the session; the other teacher was assigned to be the coach. Lesson planning was considered as an important element of classroom practice and part of teacher professional knowledge (Gess-Newsome 2015). Britton and Anderson (2010) opine that peer coaching in teacher education is a helpful method to enable transfer of theoretical knowledge into teaching practice. The three major components of their successful peer coaching consisted of a pre-conference with peers, observation data from the classroom, and a post-conference with peers. Derived from a research synthesis, Lu (2010) concludes that there are only a few findings from research yet on the effects of peer coaching. This is true especially for studies which analyse the development of coach and coachee separately and which try to explain the reciprocal effects within each dyad. In collaborative learning settings, one peer might have greater knowledge or different beliefs than the other. Such interchanges are mutually beneficial because the less competent learns content from the more competent, and the knowledge of the more competent is strengthened by explaining. Also, changes in attitudes towards science inquiry teaching in one peer can be assumed to be accompanied by changes in the other. Dyadic dependence refers to the fact that the variable scores collected from individuals interacting within dyads are not independent, but are likely to be more correlated than scores from individuals in different dyads (Gonzalez and Griffin 2012; Peugh et al. 2013). Based on an actor-partner interdependence model (APIM) (Laursen et al. 2008), it is possible to clarify the relations between and within the dyads of coach and coachee. In addition, it is possible to analyse developments over time. Furthermore, the APIM allows for including mediators in the model (Ledermann et al. 2011). The question we aim at is whether pre-service teachers in peer learning situations can benefit from each other with respect to knowledge building or development of favourable attitudes towards science inquiry teaching.

Theoretical Background

Scientific Inquiry Teaching

The term inquiry is associated with 'good science teaching and learning' for over 50 years (Anderson 2002). This assumption is based on research that shows that pupils who participate in inquiry-based activities achieve higher scores on standardised measures of science learning while exhibiting increased interest in and understanding of the nature of science (Furtak et al. 2012; Schroeder et al. 2007). Science inquiry teaching refers to methods and activities that lead to the development of scientific knowledge (Schwartz et al. 2004). Recent research has also shown that teachers' attitudes on scientific inquiry teaching affect student learning outcomes (Fogleman et al. 2011; Keys and Bryan 2001). However, pre- and in-service teachers often carry limited views on scientific inquiry teaching. Such teachers associate, laboratory experience with knowledge verification and technical skill development (Gyllenpalm and Wickman 2011; Tesch and Duit 2004; Wallace and Kang 2004). Few teaching education programmes help pre-service teachers develop effective pedagogies for laboratory settings. In turn, some teachers do not know them well, have difficulty seeing the benefit of inquiry activities, or are not confident in their capacities to properly support students (Abrahams and Reiss 2012). As such, researchers have highlighted that it is still important to provide empirical support for inquiry-based learning, but as a precondition, they should ensure that the potential of such pedagogy is realised (Avery and Meyer 2012). Currently, in many countries, curriculum initiatives and national science standards are being imposed to guide and support science teachers in achieving these aims (Neumann et al. 2010; NGSS Lead States 2013). Beginning teachers encounter several constraints when they start teaching science. Depending upon their prior beliefs, knowledge, and understanding, which may or may not have been reinforced during teaching education programmes, such constraints affect their teaching of inquiry (Roehrig and Luft 2004). To support beginning teachers in implementing science as inquiry, induction programmes that attend to these constraints are needed. For example, workshops focused on inquiry-focused teaching strategies can offer pedagogical content knowledge for those lacking such tools for their preparation programmes (Luft et al. 2003).

Pre-service Teacher Knowledge of and Attitudes Towards Scientific Inquiry Teaching

A synthesis of previous research by Davis et al. (2006) shows that numerous pre-service teachers possess unsophisticated understandings of inquiry teaching approaches and related skills, though, of course, individuals vary. With respect to beliefs on the nature of science, the authors present similar findings of relatively unsophisticated beliefs. In a study by Lemberger et al. (1999), it showed that secondary school teachers for biology who exited teacher preparation still struggled with the conflict between transmission beliefs about teaching and conceptual change teaching. Inquiry-focused teaching practices may be influenced by previous experience. The results from a teacher education project held at a public university in the northwest USA suggest that pre-service teacher uses of inquiry teaching approaches in the classroom are most strongly associated with previous research experiences gained via, and less so through experiences with the training project itself (Windschitl 2003). Prospective teachers may thus require access to continuing experiences with inquiry projects throughout their educational careers either in conjunction with scientific inquiry courses or in

contact with real scientists who can discuss their work practices and how they acquire new knowledge (Lotter et al. 2007; Windschitl 2003).

Pre-service teachers presenting negative feelings towards scientific inquiry instruction have less self-confidence and self-efficacy related to science teaching (Tosun 2000). Problems associated with low self-efficacy and low self-confidence beliefs seem to appear often among primary teachers (Avery and Meyer 2012; Forbes and Zint 2010). This may be due to a disregard for pre-service teacher self-confidence during science courses (Appleton 1995). Experience from science courses and high content knowledge (CK) do not seem to affect pre-service teacher self-efficacy beliefs (Tosun 2000). However, inquiry-based science content courses that actively involve primary school pre-service teachers in collaborative processes of learning and discovery can promote a positive change in attitudes (more confidence, enjoyment, and relevance), as a study by Riegle-Crumb et al. (2015) shows. Studies on beginning science teachers show that secondary school teachers struggle with inquiry-focused teaching as well. This is partly due to management issues, unmotivated or less competent students, and time constraints but also due to external impediments, related to class time, facilities, and/or administration (Crawford 2007; Wallace and Kang 2004). Teachers' rather positive beliefs with respect to inquiry teaching practices may differ from what external observers find (Brown and Melear 2006). Pre-service teachers with insufficient subject-matter knowledge seem to struggle more when teaching inquiry strategies, such as involving students in framing questions or in formulating explanations (Crawford 2007). Content knowledge alone does not appear to guarantee the successful implementation of inquiry-based lessons. However, strong (pedagogical) content knowledge combined with student-centred beliefs and a contemporary view on the epistemological underpinnings of science activities and on characteristics of resulting knowledge may increase the likelihood that inquiry will be taught in the classroom (Roehrig and Luft 2004). In the study of Roehrig and Luft (2004), inquiry based-instruction was modelled explicitly by science educators. A specific focus was on the role of 'scientifically oriented questions' in creating inquiry environments for student-centred learning.

Peer Learning in Collaborative Situations

Admiraal et al. (2012) recommend teacher communities, i.e. teachers working and learning together, as valuable for teacher education programmes. Teacher communities can be seen as a vehicle for discussing and reflecting on the professional domain and the process of becoming a teacher. Peer coaching can be described as a collegial approach to the analysis of teaching aimed at integrating new skills and strategies in classroom practice (Joyce and Showers 1995). As in McAllister and Neubert's (1995) model peer coaching generally involves a cycle of pre-lesson conference, lesson observation, and a post-lesson conference. It seems to be an effective learning opportunity for pre-service teachers in collaborative learning situations (Bowman and McCormick 2000). Robbins (2015) distinguishes between two forms of peer coaching: collaborative work and formal coaching. Collaborative work occurs between professional colleagues engaged in activities not related to classroom observation but not yet on classroom practices that affect student learning. Formal coaching requires mostly a coach invited by a teacher and a certain structure is part of the coaching process. Foltos (2013) characterises peer coaches as teacher leaders among peers who possess expertise in the field.

Vygotsky (1978) developed a social theory of cognitive bootstrapping, which was one of the first aimed at describing the role of social interaction in cognitive development. The theory holds that more advanced members of the peer group promote the cognitive development of their less

advanced counterparts by demonstrating effective learning strategies and skills. Through collective thinking and discussion, each participant interprets, transforms, and internalises new knowledge. In a study by DeLay et al. (2014), it showed that the greatest increases in computer programming knowledge occurred among confident partners who were paired with a friend who had relatively more initial computer programming knowledge. Next to a peer who possesses knowledge, mutual liking of each other also seems to promote development. In addition, establishing trust is a major component of peer coaching as it often facilitates learning (Costa and Garmston 1994). A non-threatening atmosphere in which partners perceive each other as equals is relevant for the perceived effectiveness of peer coaching (Zwart et al. 2007). However, creating a relationship based on trust and respect may not be enough to encourage innovation. Peer coaches should encourage their colleagues to leave the comfort zone in order to adopt more innovative practices (Foltos 2013).

Foltos (2013) further suggests that content knowledge is not necessarily needed to be a successful coach but might be acquired from additional training. Other coaching approaches such as content-focused coaching by West and Staub (2003) declare content knowledge as a core of (peer) coaching with a focus on designing effective learning environments. Kreis and Staub (2011) tested this approach as a scaffold for situated teacher learning for pre-service teachers. They report significantly positive differences related to the quality of the teaching practice for the intervention group. In a study by Jenkins et al. (2005), pre-service teachers developed PCK for physical education based on feedback by peer coaches who observed the lesson of their partners. Jang (2010) reports similar results for in-service science teachers observing each other teaching the concepts of heat and the process of floating with the help of whiteboards.

In order to examine issues such as those outlined above, data from both partners (dyadic data) are needed. The statistical analysis of longitudinal data from pre-service teacher dyads is challenging because it requires us to take into account at least two types of non-independence simultaneously: autocorrelation and dyadic non-independence (Kenny et al. 2006). From a statistical point of view, it needs to be considered that if two individual pre-service teachers have interdependent beliefs and knowledge, it follows that the data describing these beliefs and knowledge cannot be independent. Unfortunately, most conventional statistical procedures assume independent data from participants. In this way, relational interdependence begets statistical non-independence (Gonzalez and Griffin 2012). With the help of a special structural equation model—the actor–partner interdependence model (see the methods section)—it is possible to analyse dyadic data under consideration of the dependence of the data.

We expect coaches to influence the coachees' PCK and beliefs with respect to the discussion about the lesson planning.

We investigate the following research questions:

1. Does learning occur between the two peers? Are there effects within the dyads with respect to attitudes towards science inquiry teaching and pedagogical content knowledge (PCK)?
2. Is there a mediation effect of the competency in lesson planning on the development of attitudes and PCK?

Methods

Design

The present study was part of the Internationale Bodensee Hochschule (IBH)-project 'KUBEX' (content-focused peer coaching in pre-service teacher education) and was conducted by four teacher education universities across Germany (1) and Switzerland (3). IBH projects aim at doing research in local networks around the Lake of Constance. Although teacher education is organised differently in Germany and Switzerland, the competence level of the pre-service teachers is comparable (cf. Baer et al. 2009). The project began in 2014 and occurred over 2 years.

The study design was quasi-experimental and longitudinal (Table 1), and its primary aim was to determine the effects of a pre-service teacher training module on content-focused peer coaching (Kreis et al. 2015). The intervention group was trained for coaching activities that foster co-constructive and critical elaboration during planning and reflection. Part of the training was the application of cognitive tools such as Core Issues of lesson planning (West and Staub 2003), which is a set of questions about four crucial dimensions of teaching that should help coaches to foster a reflective and transformative dialogue about the prepared lesson plan. For the content of the coaching, we used experimentation as a central means of acquiring biological knowledge involving all forms of scientific inquiry. In this respect, a secondary target involved exploring the state of pre-service teacher knowledge and their attitudes on scientific inquiry instruction.

We measured the participants' knowledge and attitudes before (T1) delivering a training service consisting of 2 × 2 lessons with theoretical input on scientific inquiry to bring all of the participants to a similar level of knowledge. We then measured their pedagogical content knowledge again (T2) and carried out the intervention. In a next step, again 2 × 2 lessons, the intervention group received information on peer coaching, while the control group discussed non-interfering biology-related content. All of the pre-service teachers were then asked to prepare a lesson focusing on process skills of scientific inquiry, such as: posing questions or hypotheses; planning, conducting, and evaluating experiments; or analysing scientific data. Visual perceptions were used as biological background content for the facilitation of scientific inquiry. Each pre-service teacher was required to present his or her lesson plans as a diagram to his or her coaching partner. These collaborative planning sessions were videotaped for subsequent analysis. Finally, their knowledge and attitudes were measured once again (T3). The entire training service (including the empirical component) lasted for roughly 1 month and was conducted at each university separately but standardized by the same research team.

Table 1 Research design (constrained to aspects of this article)

T1 test and questionnaire	Input	T2 test	Intervention	Videography	T3 test and questionnaire
<i>Beliefs on scientific inquiry</i>	IG and CG: PCK input teaching on	–	IG: peer coaching	<i>Coached/Peer-assisted collaborative lesson planning</i>	<i>Beliefs on scientific inquiry</i>
CK/PCK of teaching optics on scientific inquiry	scientific inquiry in biology	<i>PCK of teaching optics on scientific inquiry</i>	CG: independent input		<i>CK/PCK of teaching optics on scientific inquiry</i>

Measures set in italics are part of the APIM

Items

Attitudes

For the items that measured teachers' attitudes towards scientific inquiry, we referred to the work of Pell and Jarvis (2003) and their scales for cognitive beliefs. We used a 6-point Likert scale (6 = absolutely agree, 5 = agree, 4 = somewhat agree, 3 = somewhat disagree, 2 = disagree, and 1 = absolutely disagree). The survey items were summarised in two scales, and a reliability index (Cronbach's alpha) was computed. The appendix presents mean and standard deviation values for the scales. Exploratory factor analysis indicated that the two scales should be treated as one because only one factor could be extracted. Therefore, in the further analyses, we refer to one single scale for scientific inquiry attitudes.

PCK and CK

All items used to measure PCK and CK were newly developed for this study because close alignment with the teaching content was required. The items on PCK were based on two dimensions that are important components of the PCK model of Gess-Newsome (2015):

1. Knowledge of student understanding of science (e.g. student pre-conceptions of visual perception, student mistakes when planning, and conducting experiments or analysing data), and
2. Pre-service teacher knowledge of instructional strategies (e.g. the functioning of experiments in lessons on biology or the treatment of student misinterpretations of results from experimental results).

Park and Chen (2012), who examined the declarative dimensions of PCK, demonstrated that biology teachers tend to connect knowledge of student understanding and knowledge of instructional strategies and representation, and these two PCK components may comprise a target area for PCK improvement. The items that were related to the content requested knowledge (CK) in the field of visual perception. Item development for CK was guided by a research project of Dannemann and Krüger (2010), who developed items on the topic of visual perception for the diagnosis of student conceptions. The complete test battery consisted of 22 items for CK and 21 items for PCK. The test items for CK were single choice questions with four options, whereas 17 of the PCK items were true/false questions and four items were open-ended as shown in the appendix. For the analysis in the structural equation model, both PCK and CK were measured as one-dimensional constructs.

Planning Competency for Science Inquiry Teaching

To measure the competency of each student in planning a science inquiry lesson, we videotaped a lesson planning dialogue in a peer session. The lesson focused on science inquiry teaching in biology and the topic was accommodation and adaptation of the eye. As a basis for measurement, a rating-manual with 19 items was established. Although the lesson should address experimentation mainly, content knowledge always plays an important role as well. Therefore, the manual consists of two dimensions: experimentation as part of science inquiry teaching and visual perception as a topic. The 5 subdimensions for experimentation relate to student concepts of science inquiry, teachers' content knowledge with respect to science inquiry, didactical/lesson structuring, goal orientation and embedded lesson reflection. The dimension for teachers' content knowledge with respect to visual perception was measured with the help of two items: scientific correctness and scientific vocabulary.

Sample

Our sample included 121 teacher education students of three Swiss universities ($n = 53$) and one German ($n = 68$) university in the Lake of Constance area; 3 students dropped out over the course of the project. 120 students took part just after the videography and 118 terminated all of the three questionnaires. All of the students were preparing to teach the subject of biology in secondary schools (grades 5–9 in Germany, grades 7–9 in Switzerland). Participants were required to have completed laboratory or scientific inquiry instruction courses. Thus, most of the participants were in more advanced study semesters. While the range was between the 2nd and 8th semesters, the modus and median were the 6th semester. In total, 75 % of the students were female and 25 % were male with a mean age of 22.9 years ($SD\ 3.4$). Ten students possessed previous work experience, and 15 had previously studied another subject. Only 29 of the students had not taught biology as part of their practical education over short periods in schools, and many (55) had never used experiments for their own classes prior to this study.

Analysis

Test Analysis Based on Item Response Theory

Mplus 7.0 (Muthén and Muthén 2012) was used for our calculations. The test data were first scaled for PCK and CK separately using either a three- or two-dimensional 2PL-IRT model with one dimension for each measurement occasion. Item parameters were estimated using the probit regression and the WLSMV-estimator for categorical data. Item parameters for each anchor item were fixed as equal across measurement occasions. To determine whether the developed test instruments fit the IRT model, model fit indices on the items were reviewed (Hu and Bentler 1999). After checking the item characteristic curves, we conducted an exploratory factor analysis to determine whether the items loaded on their intended factors for the unsatisfactory fit-indices. Non-fitting items were removed until adequate fit indices were reached. Upon completing the IRT analyses, final person measures based on Bayesian plausible values (Von Davier et al. 2009) were computed for the CK and PCK tests. Plausible values were also calculated for persons with missing data. Mplus provides imputed data sets using Rubin's method (Rubin 2004).

Bayesian Estimation

Especially for small sample sizes and non-normal distributed samples, Bayesian analyses are an attractive alternative to ML estimation (Muthén 2010; Muthén and Asparouhov 2012). Bayesian estimates are obtained as the means, modes, or medians of their posterior distributions. Priors can help optimise small variance parameters. Mplus uses a series of default priors. Bayesian explorations of model fit can be performed in a flexible way via posterior predictive checking. An excellent fitting model is expected to have a posterior predictive p value (PPP) of approximately 0.5 and an f statistic difference of zero that falls close to the middle of the confidence interval. A positive lower limit is in line with a low PPP and denotes poor fit. A 95 % confidence interval is produced for differences in the f statistic for real and replicated data.

Video Rating

Two raters independently assessed the first 26 videos of a total 120 until reaching an inter-rater reliability of 0.80. The next following videos were single rated. Each item had a scale with four descriptive categories from 1 (is not present at all), 2 (is present partly), 3 (is present largely), to 4 (is present completely). Inter-rater reliability was measured with the help of the software Facets (Many-Facet Rasch Measurement) Version No. 3.71.4 (Linacre 2013). The Facets programme used the 26 first double ratings to estimate the competency of each student with respect to science inquiry lesson planning and the severity of each rater, as well as the difficulty of each item. The MFRM approach to analysing rating data logistically transforms raters' ordinal ratings to an equal-interval logit scale of measures. The Facets programme reports several fit indices for each rater, student, and item. The rating data showed reasonably sound psychometric properties and had a Cronbach's alpha of 0.84. One item possessed insufficient item discrimination and was omitted for further analysis. Generally, higher ratings were much less frequent than lower ratings. Rating averages were between 1.12 and 1.87. Next, the complete rating data (N = 120) was analysed to check for the two-dimensional structure of the construct competency for science inquiry lesson planning. We compared a categorical SEM model with two correlated dimensions with a second-order model. A Bayesian estimator with a Gibbs algorithm was applied in order to obtain plausible values for the pre-service teachers' competence. The model with two correlated dimensions showed better fitting:

Fig. 1 The actor-partner interdependence model. Variables X_1 , X_2 , Y_1 , and Y_2 indicate measured variables; E_1 and E_2 denote errors

After conducting estimations for different iterations to determine convergence and PSR values, the outputs of the SEM showed stable results. The PPP value amounted to 0.39 and to an f difference of 15.98, which should be positive and strive towards 1. The number of free parameters was 71. All 18 factor loadings were significant and the correlation between the two dimensions was $r = 0.31$ (standardized).

Actor-Partner Interdependence Model (APIM)

The standard APIM for distinguishable members, shown in Fig. 1, consists of four measured variables (represented by rectangles) and two latent error terms (represented by circles). In this model, measured variables vary between and within dyads and are termed mixed variables (Kenny and Ledermann 2010). The variables X_1 and X_2 represent the causal or predictor variables of persons 1 and 2 of a dyad, respectively, and Y_1 and Y_2 represent the outcome variables for the two members. The model contains two actor effects, a , represented by horizontal arrows, and two partner effects, p , represented by diagonal arrows. The curved, double-headed arrow on the left represents the covariance between the two causal variables, and the one on the right represents the correlation between the two error terms. The latter indicates that the errors co-vary between dyad members because of unmeasured common causes.

In an actor-partner interdependence model with a longitudinal design X_1 and X_2 represent the same variable at two points of measurement while Y_1 and Y_2 stands for the other person in the dyad but with the same variable. The actor effect a is interpreted as the stability effect for X and Y over time, while the partner effect p represents the influence that each partner has over the other. For example, the effect of the coach's attitude on the coachee's attitude over the training session. Such a longitudinal APIM is similar to a cross-lagged model for longitudinal analysis. An actor-partner interdependence model can be extended with mediators (Ledermann et al. 2011). Such a model is called actor-partner interdependence mediation model or APIMeM. It is important to note that time 1 correlation is between raw scores, whereas the time 2 estimates correlations between residual terms. Residual correlation (between E_1 and E_2) represent any remaining non-independence in the dyad after partialing out individual stability and partner influence effects including error and unobservable variables. To overcome this limitation, correlations between raw scores can be computed. Of course, correlations between raw scores inflate similarity because they include stability and partner influence effects, but they do provide a common metric for comparing results.

Results

Pre-Service Teacher Attitudes and PCK for Science Inquiry Instruction

Descriptive statistics on the pre-service teachers' attitudes and knowledge are presented in Table 2. In relation to the Pell and Jarvis (2003) sample, our prospective secondary school teachers presented slightly more positive scientific inquiry teaching beliefs. At the end of our project at t3, the students showed slightly higher attitudes than at t1 (ANOVA, $p < 0.001$; $\eta^2 = 0.12$) and they possessed a little more PCK than before the lesson planning session at t2 (ANOVA, $p < 0.05$; $\eta^2 = 0.05$). The students'

Table 2 Descriptive statistics and correlations of individual pre-service teacher attitudes towards science instruction, PCK, and science inquiry planning competence

	Mean	S.D.	1	2	3	4	5
1. Scientific inquiry attitudes t1	5.02	0.74					
2. Scientific inquiry attitudes t3	5.26	0.50	0.51**				
3. PCK t1	0.01	0.50	0.14	0.03			
4. PCK t2	0.59	0.87	0.19*	0.16	0.83**		
5. PCK t3	0.53	0.88	0.23*	0.18	0.84**	0.95**	
6. Science inquiry planning competence	-0.11	1.29	0.13	0.02	0.27**	0.33**	0.35**

Attitude scale 1–6, 6 fully agree, PCK and science inquiry planning in logits

$N = 120$; ** $p < 0.01$, * $p < 0.05$

mean planning competence was -0.11 logits, indicating that the average student is a little below the test mean difficulty. Planning competence and PCK correlates significantly before and after the collaborative planning session, but there is no significant correlation between planning competence and inquiry attitudes. To check for interaction effects with respect to the intervention and the control group over time, we applied an analysis of variance with repeated measurements, each for attitudes and PCK. However, we could not find any significant effects for the two groups and over time. For planning competence, we could not find a significant difference between the two groups, either. Therefore, we continued our analysis based on the complete sample of 60 dyads. As the two student teachers swapped roles and each of them was once the coach and once the coachee, we have data for all individuals in both roles. This produced a total of 120 datasets, where a pre-service teacher presents his or her lesson planning.

Dyadic Effects on PCK and Attitudes

The dyads started collaborating after the PCK-input on scientific inquiry teaching (t2) and before the lesson planning sequence. Thus, we integrated PCK t2 and t3 in our model but not PCK t1. Attitudes were only measured twice: at the beginning (t1) and at the end of our project (t3). Both, PCK and attitudes towards scientific inquiry teaching (ATT) have increased slightly but significantly during our project. The ICC values (above 0.20, except ATT t1 = 0.06) for PCK and ATT indicate that a relevant proportion of the variance lies on the level of the dyads and not on the individual level. This suggests that the individual development of the pre-service teacher depends on the coaching partner and his or her PCK and ATT, as well. Noteworthy is a higher amount of shared variance for ATT at t2.

In Table 3, the correlations for ATT, PCK, and lesson planning competence between the peers in each dyad are presented. The correlations inflate similarity because they include stability and partner influence effects, but they do provide a common metric for comparing results. For example, correlations at each time point (Kenny et al. 2006). While correlations within the dyad remained the same for PCK, for ATT, changes can be noticed: Before the planning discussion, there was no significant correlation apparent for ATT, but afterwards, there is a significant correlation.

Table 3 Correlations between peers' attitudes towards science instruction, PCK, and science inquiry planning competence

	ATT t1 coachee	PCK t2 coachee	ATT t3 coachee	PCK t3 coachee	Planning coachee
ATT t1 coach	0.04	0.01	0.03	0.00	-0.02
PCK t2 coach	0.07	0.24*	0.15	0.23*	-0.02
ATT t3 coach	0.04	0.04	0.27**	0.06	0.00
PCK t3 coach	0.05	0.23*	0.16	0.24*	0.01

N = 120; ***p* < 0.01, **p* < 0.05

The stability and the effects of the peers can be tested in a longitudinal APIM model with the help of a structural equation model (SEM). A basic APIM without covariate (Fig. 2) shows good fit-values: $\chi^2 = 672.94$, *df* = 22, *p* = 0.00, CFI = 1.00, TLI = 1.00, RMSEA = 0.00, and SRMR = 0.00.

As already mentioned, there is a high correlation of PCK and ATT for each peer before the lesson planning session. On the dyad level, PCK correlates significantly, but not ATT. The actor effects that indicate stability for each pre-service teacher show high stability for PCK ($\beta = 0.94$) and medium stability effects for ATT ($\beta = 0.50$). There are two small but significant cross-lagged effects within the coachee and the coach each: ATT at t1 predicts PCK at t3 ($\beta = 0.05$). However, there are no significant partner effects, meaning there is no effect of coach on coachee or vice versa over time related to PCK and ATT. At the end of the project, there is no correlation between PCK and ATT within each peer any longer, but PCK and ATT correlate each within the dyads. For the coachee's PCK and ATT, there was not a correlation before. As a conclusion, we could state that knowledge and attitudes have converged between coach and coachee towards the end of the project. Change has happened more clearly for ATT as the actor effects are less high and there are different correlations before the discussion and after. The reason for this change in attitudes seems not to lay in the interaction of coach and coachee but before, supposedly in the course input on scientific inquiry teaching (see Table 1).

Fig. 2 APIM Model for the interplay of PCK and attitude towards science inquiry teaching; $N = 120$, all effects are standardized, *bold lines*: sign., $p < 0.01$; *dotted lines* are non-significant

In a next step, we added the variable lesson planning competence of the coachee as a mediator to the actor-partner interdependence mediation model or APIMeM. Mediation occurs when an independent variable affects an outcome variable through a third variable, called mediator (Wu and Zumbo 2008). We assumed that the lesson planning discussion has an effect on the development of the coachee's PCK over time. Moreover, we expected the coachee's attitude and knowledge to be a significant predictor of the lesson planning competency of the coachee.

The APIMeM with lesson planning as a covariate (Fig. 3) shows good fit values as well: $\chi^2 = 690.46$, $df = 30$, $p = 0.00$, CFI = 1.00, TLI = 1.01, RMSEA = 0.00, and SRMR = 0.04. Compared to the first APIM, there has been only one cross-lagged effect left between ATT at t1 and PCK at t2 for the coach. The other paths are similar to those in Fig. 1. The lesson planning competence has no mediating effect for the development of PCK over time but PCK2 is a significant predictor of the coachee's lesson planning competency. In addition, neither the coach's attitudes nor PCK helps to predict the lesson planning competency of the coachee.

Discussion

The results show that exchanging knowledge and attitudes in peer coaching situations triggers development. However, the results are not as clearly as expected. Knowledge and attitudes of the coach have no effect on the competency for lesson planning or on PCK and ATT of the coachee. The exchange leads to shared attitudes between the peers

Fig. 3 APIMeM for the interplay of PCK, attitude towards science inquiry teaching and competency of lesson planning, $N = 120$ dyads, all shown effects are standardized and significant, $p < 0.01$. Non-significant paths are omitted

towards science inquiry teaching and to shared PCK only. Nevertheless, in the socio-constructivist concept of grounding, such shared understandings are seen as essential in order to maintain the collaborative discourse (Clark and Brennan 1991). A little irritating, at the end of the project, ATT and PCK within each peer do not correlate anymore. Actually, they still do, but not so strong anymore that the values become significant. The collaborative learning session (lesson planning) did not lead to an effect on PCK or ATT either. While the theoretical input on science inquiry teaching lead to an increase in pre-service teachers' PCK (see Table 2), the practice-oriented transfer (lesson planning) itself seemed not to have added much to the development of professional knowledge despite the postulated effects of reflection as part of successful teacher training in other studies (Van Driel and Berry 2012). An analysis of the content of the planning dialogues in our study shows a focus on general pedagogical knowledge such as time management and less on PCK. Moreover, we assume that the quality of the peer coaching sessions in our study varies considerably, and that not all of the dialogues reached the level of co-construction which is crucial in the learning process (Rytivaara and Kershner 2012). This matter is something which we will look into further.

It is likely that the interaction frequency and time was too short to produce clearer reciprocal effects between coach and coachee. In addition, an effective coach needs to have coaching expertise acquired by broad experience. Our pre-service students were not used to plan lessons collaboratively. They probably need more practice in peer coaching. It might be that our pre-service teachers found it difficult in the role of the coach to be a critical peer and to push the coachee to become more innovative with respect to science inquiry teaching (Foltos 2013). In addition, scientific inquiry teaching is a particularly difficult and complex teaching topic (Harlen 2013). Lotter et al. (2009), who conducted an intensive teacher training programme, likewise report that their secondary science pre-service teachers struggled with incorporating NOS instruction and inquiry-

based instructional practices into their unit plans. With respect to teacher education, it can be concluded from our study that even if pre-service teachers understand the aim of scientific inquiry teaching and even if they possess the supporting attitudes, implementation in the classroom is not guaranteed at all. Teacher training should offer more opportunities for transfer of pre-service teacher professional knowledge to practice. Lesson planning is such an opportunity to build up personal PCK (Gess-Newsome 2015) during teacher training.

A limitation of our research is that we did not include a variable indicating the affective connections between coach and coachee. In a study by DeLay et al. (2014), partner effects were only significant for partners who were friends and not for non-friends. However, the quality of interaction (coachee feels understood by the coach) was generally rated high in our survey. We did no study on any subgroups, such as pre-service teachers with high PCK working together with peers with low PCK. Perhaps, according to theory, partner effects might be stronger in such dyads. This would require a larger sample size; nevertheless, the study by Hartl et al. (2015) about learning in dyads is based on a similar sample size.

Acknowledgments

The presented project was funded by the International University of Lake Constance (IBH) (585113). The KUBeX project research team included Christine Bieri, Robert Blank, Christian Brühwiler, Eva Engeli, Annelies Kreis, Stephanie Musow, Florian Rietz, Stefanie Schnebel, Robbert Smit, Josiane Tardent, Sandra Wagner, Holger Weitzel, and Corinne Wyss.

Appendix: Scales and sample test items

Scale 'General scientific method' (Pell and Jarvis 2003) – Scientific inquiry attitudes part I

Item	An effective biology teacher...	<i>M</i>	<i>SD</i>
gsm2	Tries to encourage pupils to think for themselves.	5.41	0.72
gsm5	Encourages pupils to make careful observations.	5.48	0.62
gsm7	Shows pupils how to reason and make deductions from facts and observations.	4.78	0.89
gsm10	Explains ways to recognize and control variables in an experiment.	4.94	0.95

Answers included the following: 1 (absolutely disagree), 2 (largely disagree), 3 (rather disagree), 4 (rather agree), 5 (largely agree), 6 (absolutely agree); $N = 120$, $\alpha = 0.64$.

Scale 'Pupil-centred science' (Pell and Jarvis 2003) – Scientific inquiry attitudes part II

Item	An effective biology teacher...	<i>M</i>	<i>SD</i>
pcs1	Encourages pupils to guess and speculate.	5.13	0.95
pcs3	Can help pupils generalise from observations.	4.68	1.01
pcs4	Encourages pupils to plan their own investigations.	4.91	1.13
pcs6	Shows pupils how classroom lessons relate to the real world.	5.49	0.97
pcs9	Through careful planning, guides pupils in making their own 'discoveries'.	5.17	1.00

Answers included the following: 1 (absolutely disagree), 2 (largely disagree), 3 (rather disagree), 4 (rather agree), 5 (largely agree), 6 (absolutely agree); $N = 120$, $\alpha = 0.84$.

'PCK' sample item (please tick the appropriate boxes)

D4. You would like to plan your lessons on 'seeing' topics in such a way that concept changes by students are explicitly enhanced. Which of the following strategies do you think are appropriate based on your current knowledge on teaching biology?

	appr.	not appr.
a. At the start of a lesson, the teacher explains how seeing works and then repeats what was explained as part of classroom discourse.	☐	☐
b. The teacher assesses students' concepts with the help of appropriate tools.	☐	☐
c. The teacher does not provide feedback on incorrect student responses and instead presents correct concepts as part of class instruction.	☐	☐
d. The teacher allows students to evaluate their partly incorrect concepts with the help of appropriate tasks.	☐	☐
e. The teacher inspires student dissatisfaction in their incorrect knowledge on seeing, e.g., with the help of learning tasks that show issues associated with these inadequate concepts.	☐	☐

References

- Abrahams, I., & Reiss, M. J. (2012). Practical work: its effectiveness in primary and secondary schools in England. *Journal of Research in Science Teaching*, 49(8), 1035–1055. doi:10.1002/tea.21036.
- Admiraal, W., Akkerman, S., & de Graaff, R. (2012). How to foster collaborative learning in communities of teachers and student teachers: introduction to a special issue. *Learning Environments Research*, 1–6.
- Anderson, R. D. (2002). Reforming science teaching: what research says about inquiry. *Journal of Science Teacher Education*, 13(1), 1–12. doi:10.1023/A:1015171124982.
- Appleton, K. (1995). Student teachers' confidence to teach science: is more science knowledge necessary to improve self-confidence? *International Journal of Science Education*, 17(3), 357–369.
- Avery, L. M., & Meyer, D. Z. (2012). Teaching science as science is practiced: opportunities and limits for enhancing preservice elementary teachers' self-efficacy for science and science teaching. *School Science and Mathematics*, 112(7), 395–409. doi:10.1111/j.1949-8594.2012.00159.x.
- Baer, M., Doerr, G., Fräfel, U., Kocher, M., Küster, O., Larcher, S., & Wyss, C. (2009). Competences and standards in teacher education in Switzerland and Germany—do prospective teachers become more competent through teacher training. In F. Achtenhagen, F. Oser & U. Renold (Eds.), *Teachers' professional development: Aims, Modules, Evaluation* (pp. 145–174). Rotterdam: Sense.
- Bowman, C. L., & McCormick, S. (2000). Comparison of peer coaching versus traditional supervision effects. *The Journal of Educational Research*, 93(4), 256–261.
- Britton, L. R., & Anderson, K. A. (2010). Peer coaching and pre-service teachers: examining an underutilised concept. *Teaching and Teacher Education*, 26(2), 306–314. doi:10.1016/j.tate.2009.03.008.
- Brown, S. L., & Melear, C. T. (2006). Investigation of secondary science teachers' beliefs and practices after authentic inquiry-based experiences. *Journal of Research in Science Teaching*, 43(9), 938–962. doi:10.1002/tea.20110.
- Bryan, L. A. (2012). Research on science teacher beliefs. In J. B. Fraser, K. Tobin, & J. C. McRobbie (Eds.), *Second international handbook of science education* (pp. 477–495). Dordrecht: Springer Netherlands.
- Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. *Perspectives on Socially Shared Cognition*, 13(1991), 127–149.
- Costa, A.L., & Garmston, R.J. (1994). Cognitive coaching: A foundation for renaissance schools: ERIC.
- Crawford, B. A. (2007). Learning to teach science as inquiry in the rough and tumble of practice. *Journal of Research in Science Teaching*, 44(4), 613–642. doi:10.1002/tea.20157.
- Dannemann, S., & Krüger, D. (2010). Evaluation eines Aufgabeninventars zur Ermittlung von Schülervorstellungen zum Sehen [Evaluation of a task inventory for assessing students' concepts of eyesight]. In U. Harms & I. Mackensen-Friedrichs (Eds.), *Lehr- und Lernforschung in der Biologiedidaktik, Band 4* (pp. 137–154). Innsbruck: Studienverlag.
- Davis, E. A., Petish, D., & Smithey, J. (2006). Challenges new science teachers face. *Review of Educational Research*, 76(4), 607–651. doi:10.3102/00346543076004607.

DeLay, D., Hartl, A. C., Laursen, B., Denner, J., Werner, L., Campe, S., & Ortiz, E. (2014). Learning from friends: measuring influence in a dyadic computer instructional setting. *International Journal of Research & Method in Education*, 37(2), 190–205.

Fogleman, J., McNeill, K. L., & Krajcik, J. (2011). Examining the effect of teachers' adaptations of a middle school science inquiry-oriented curriculum unit on student learning. *Journal of Research in Science Teaching*, 48(2), 149–169. doi:10.1002/tea.20399.

Foltos, L. (2013). *Peer coaching: Unlocking the power of collaboration*. Thousand Oakes: Corwin Press.

Forbes, C. T., & Zint, M. (2010). Elementary teachers' beliefs about, perceived competencies for, and reported use of scientific inquiry to promote student learning about and for the environment. *The Journal of Environmental Education*, 42(1), 30–42.

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.

Gess-Newsome, J. (2015). A model of teacher professional knowledge and skill including PCK: results of the thinking from the PCK Summit. In A. Berry, P. Friedrichsen, & J. Loughran (Eds.), *Re-examining pedagogical content knowledge in science education* (pp. 14–27). New York: Routledge.

Gonzalez, R., & Griffin, D. (2012). Dyadic data analysis. *APA handbook of research methods in psychology*, 3, 439–450.

Gyllenpalm, J., & Wickman, P.-O. (2011). BExperiments^ and the inquiry emphasis conflation in science teacher education. *Science Education*, 95(5), 908–926. doi:10.1002/sce.20446.

Harlen, W. (2013). *Assessment & inquiry-based science education: Issues in policy and practice*. Trieste: Global Network of Science Academies.

Hartl, A. C., DeLay, D., Laursen, B., Denner, J., Werner, L., Campe, S., & Ortiz, E. (2015). Dyadic instruction for middle school students: liking promotes learning. *Learning and Individual Differences*, 44, 33–39.

Hu, L.-t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Structural Equation Modeling: a Multidisciplinary Journal*, 6(1), 1–55.

Jang, S.-J. (2010). Integrating the interactive whiteboard and peer coaching to develop the TPACK of secondary science teachers. *Computers & Education*, 55(4), 1744–1751. doi:10.1016/j.compedu.2010.07.020.

Jenkins, J. M., Garn, A., & Jenkins, P. (2005). Preservice teacher observations in peer coaching. *Journal of Teaching in Physical Education*, 24(1), 2–23.

Joyce, B., & Showers, B. (1995). *Student achievement through self-development: Fundamentals of school renewal*. White Plains: Longman.

Kenny, D. A., & Ledermann, T. (2010). Detecting, measuring, and testing dyadic patterns in the actor–partner interdependence model. *Journal of Family Psychology*, 24(3), 359–366. doi:10.1037/a0019651.

Kenny, D.A., Kashy, D.A., & Cook, W.L. (2006). *Dyadic data analysis*: Guilford Press.

- Keys, C. W., & Bryan, L. A. (2001). Co-constructing inquiry-based science with teachers: essential research for lasting reform. *Journal of Research in Science Teaching*, 38(6), 631–645. doi:10.1002/tea.1023.
- Kreis, A., & Staub, F. C. (2011). Fachspezifisches Unterrichtscoaching im Praktikum. *Zeitschrift für Erziehungswissenschaft*, 14(1), 61–83.
- Kreis, A., Engeli, E., Wagner, S., & Musow, S. (2015). Qualities of interaction in co-planning dialogues of student teachers - An intervention study. Paper presented at the Presentation in the symposium BPeer Mentoring as a mediator for pre-service teacher qualification[^]. EARLI-Conference 2015, Limassol, Cyprus, 26th August 2015.
- Laursen, B., Popp, D., Burk, W. J., Kerr, M., Stattin, H., & Card, N. (2008). Incorporating interdependence into developmental research: examples from the study of homophily and homogeneity. In N. A. Card, J. P. Selig, & D. L. Todd (Eds.), *Modeling dyadic and interdependent data in the developmental and behavioral sciences* (pp. 11–37). New York: Routledge.
- Ledermann, T., Macho, S., & Kenny, D. A. (2011). Assessing mediation in dyadic data using the actor-partner interdependence model. *Structural Equation Modeling: A Multidisciplinary Journal*, 18(4), 595–612. doi:10.1080/10705511.2011.607099.
- Lemberger, J., Hewson, P. W., & Park, H. J. (1999). Relationships between prospective secondary teachers' classroom practice and their conceptions of biology and of teaching science. *Science Education*, 83(3), 347–371.
- Linacre, J.M. (2013). A user's guide to FACETS Rasch-model computer programs (version 3.71.0) Retrieved from <http://www.winsteps.com>.
- Lotter, C., Harwood, W. S., & Bonner, J. J. (2007). The influence of core teaching conceptions on teachers' use of inquiry teaching practices. *Journal of Research in Science Teaching*, 44(9), 1318–1347. doi:10.1002/tea.20191. Lotter, C., Singer, J., & Godley, J. (2009). The influence of repeated teaching and reflection on preservice teachers' views of inquiry and nature of science. *Journal of Science Teacher Education*, 20(6), 553–582. doi:10.1007/s10972-009-9144-9.
- Lu, H.-L. (2010). Research on peer coaching in preservice teacher education—a review of literature. *Teaching and Teacher Education*, 26(4), 748–753. doi:10.1016/j.tate.2009.10.015.
- Luft, J. A., Roehrig, G. H., & Patterson, N. C. (2003). Contrasting landscapes: a comparison of the impact of different induction programs on beginning secondary science teachers' practices, beliefs, and experiences. *Journal of Research in Science Teaching*, 40(1), 77–97.
- McAllister, E. A., & Neubert, G. A. (1995). *New teachers helping new teachers: Preservice peer coaching*. Bloomington: ERIC Clearinghouse on Reading, English, and Communication and EDINFO Press.
- Muthén, B.O. (2010). Bayesian analysis in Mplus: A brief introduction. Retrieved from <https://www.statmodel.com/download/IntroBayesVersion3.pdf>.
- Muthén, B. O., & Asparouhov, T. (2012). Bayesian structural equation modeling: a more flexible representation of substantive theory. *Psychological Methods*, 17(3), 313–335. doi:10.1037/a0026802.
- Muthén, L. K., & Muthén, B. O. (2012). *Mplus: statistical analysis with latent variables: User's Guide* (7th ed.). Los Angeles: Muthén & Muthén.

Neumann, K., Fischer, H. E., & Kauertz, A. (2010). From PISA to educational standards: the impact of large-scale assessments on science education in Germany. *International Journal of Science and Mathematics Education*, 8(3), 545–563.

NGSS Lead States. (2013). *Next generation science standards: For states, by states*. Washington, DC: The National Academies Press.

Park, S., & Chen, Y.-C. (2012). Mapping out the integration of the components of pedagogical content knowledge (PCK): examples from high school biology classrooms. *Journal of Research in Science Teaching*, 49(7), 922–941. doi:10.1002/tea.21022.

Pell, A., & Jarvis, T. (2003). Developing attitude to science education scales for use with primary teachers. *International Journal of Science Education*, 25(10), 1273–1295. doi:10.1080/0950069022000017289.

Peugh, J. L., DiLillo, D., & Panuzio, J. (2013). Analyzing mixed-dyadic data using structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 20(2), 314–337.

Riegle-Crumb, C., Morton, K., Moore, C., Chimonidou, A., Labrake, C., & Kopp, S. (2015). Do inquiring minds have positive attitudes? The science education of preservice elementary teachers. *Science Education*, 99(5), 819–836. doi:10.1002/sce.21177.

Robbins, P. (2015). *Peer coaching to enrich professional practice, school culture, and student learning*. Alexandria: ASCD.

Roehrig, G. H., & Luft, J. A. (2004). Constraints experienced by beginning secondary science teachers in implementing scientific inquiry lessons. *International Journal of Science Education*, 26(1), 3–24. doi:10.1080/0950069022000070261.

Rubin, D. B. (2004). *Multiple imputation for nonresponse in surveys*. Hoboken: Wiley.

Rytivaara, A., & Kershner, R. (2012). Co-teaching as a context for teachers' professional learning and joint knowledge construction. *Teaching and Teacher Education*, 28(7), 999–1008. doi:10.1016/j.tate.2012.05.006.

Schroeder, C. M., Scott, T. P., Tolson, H., Huang, T. Y., & Lee, Y. H. (2007). A meta-analysis of national research: effects of teaching strategies on student achievement in science in the United States. *Journal of Research in Science Teaching*, 44(10), 1436–1460.

Schwartz, R. S., Lederman, N. G., & Crawford, B. A. (2004). Developing views of nature of science in an authentic context: an explicit approach to bridging the gap between nature of science and scientific inquiry. *Science Education*, 88(4), 610–645.

Smit, R. (2015). Attitudes and knowledge of pre-service biology teachers planning science inquiry lessons. Paper presented at the Earli Conference, Limmasol, Cyprus, 25.-29.8.15.

Tesch, M., & Duit, R. (2004). Experimentieren im Physikunterricht - Ergebnisse einer Videostudie [Experiments as part of physics instruction - results of a video study]. *Zeitschrift für Didaktik der Naturwissenschaften*, 10, 51–69.

Tosun, T. (2000). The impact of prior science course experience and achievement on the science teaching self-efficacy of preservice elementary teachers. *Journal of Elementary Science Education*, 12(2), 21–31.

- Van Driel, J. H., & Berry, A. (2012). Teacher professional development focusing on pedagogical content knowledge. *Educational Researcher*, 41(1), 26–28.
- 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.
- Von Davier, M., Gonzalez, E., & Mislevy, R. (2009). What are plausible values and why are they useful. IERI Monograph Series, 2, 9–36.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge: Harvard University Press.
- Wallace, C. S., & Kang, N.-H. (2004). An investigation of experienced secondary science teachers' beliefs about inquiry: an examination of competing belief sets. *Journal of Research in Science Teaching*, 41(9), 936–960. doi:10.1002/tea.20032.
- Wegener, D. T., & Blankenship, K. L. (2007). Ecological validity. In R. F. Baumeister & K. D. Vohs (Eds.), *Encyclopedia of social psychology* (Vol. 1, pp. 275–277). Thousand Oaks: Sage.
- West, L., & Staub, F. C. (2003). *Content-focused coaching: Transforming mathematics lessons*. Portsmouth: Heinemann.
- Windschitl, M. (2003). Inquiry projects in science teacher education: what can investigative experiences reveal about teacher thinking and eventual classroom practice? *Science Education*, 87(1), 112–143.
- Wu, A., & Zumbo, B. (2008). Understanding and using mediators and moderators. *Social Indicators Research*, 87(3), 367–392. doi:10.1007/s11205-007-9143-1.
- Zwart, R., Wubbels, T., Bergen, T. C., & Bolhuis, S. (2007). Experienced teacher learning within the context of reciprocal peer coaching. *Teachers and Teaching: Theory and Practice*, 13(2), 165–187.