

# Development and Changes in Student Teachers' Knowledge Concerning Diagnostic in Chemistry Teaching - A Longitudinal Case Study

Yannik Tolsdorf<sup>1</sup>, Silvija Markic<sup>2\*</sup>

<sup>1</sup> University of Bremen, Bremen, GERMANY

<sup>2</sup> Pädagogische Hochschule Ludwigsburg, Ludwigsburg, GERMANY

Received 18 February 2018 • Revised 29 May 2018 • Accepted 19 July 2018

## ABSTRACT

A knowledge about pedagogical diagnosis is important for professional teachers for teaching heterogeneous chemistry classes. For chemistry teachers, the knowledge about their students' perceptions, attitudes, knowledge, previous experience or their interest is crucial for planning lessons. Therefore, the diagnosis is a key component of teachers' knowledge. However, research in this area is still underrepresented. It is clear, that the development of teachers' knowledge about diagnosis in chemistry teaching and learning should be started during the university teacher education program. But, how does this knowledge develop or change during the teacher training program in chemistry education?

This paper attempts to investigate this development with a longitudinal interview case study. Two chemistry student teachers participated the study and were interviewed at seven different time points during their teacher education program. The interview and the analysis are based on the definition by Jäger with a focus on the four dimensions of the diagnostic competence: i. Competence Knowledge, ii. Conditional Knowledge, iii. Technological Knowledge and iv. Knowledge of Change. The analysis of the interviews follows documentary method. The results show a different development of student teachers, which allowed a contrasting view of the students. From this, implications for the teacher training will be presented.

**Keywords:** diagnostic competence, interview study, longitudinal study, teacher training

## INTRODUCTION

Since the heterogeneity and diversity in schools in general and chemistry classes in particular are increasing, diagnostic is a current topic in education research as well as in chemistry education (e.g. Klug, Bruder, Kelava, Spiel, & Schmitz; Ohle & McElvany, 2015; Tolsdorf & Markic, 2016a). The diagnostic describes the process to identify and promote different student skills (e.g. chemical knowledge or experimental skills). Thus, teachers in general and science teachers in particular need to possess knowledge and skills to identify and promote students in their classes (Brookhart, 2011). Finally, diagnosis isn't a new topic in the research field. Helmke and Schrader (1987) pronounce already 30 years ago that teachers must have abilities to diagnose students and use the results of the diagnosis for planning their lessons. Also, a knowledge about the diagnosis of different learning factors of students (e.g. pre-knowledge) is in the different constructs e.g. Pedagogical Content Knowledge (PCK) or Pedagogical Knowledge (PK) to be found (e.g. Shulman, 1987 or Park & Oliver, 2008). Especially, the assessment of conceptions and misconceptions (and the changing of these) are important topics in the diagnosis and support in chemistry teaching (Barke, Hazaari & Yitbarek, 2009). For teachers, different diagnostic instruments are developed to recognize misconceptions of students for chemistry lessons, which help to change or developed correct abilities (e.g. Chandrasegaran, Treagust & Mocerine, 2007; Taber, 2002). Additionally, teachers' diagnostic skills are linked

#### Contribution of this paper to the literature

- Student teachers' diagnostic competences change and develop during the teacher education program. However, this development is not linear.
- Internship and the mentor at school have a high influence on the development of student teachers' diagnostic competence.
- More focus should be put on reflection during the university teacher training program.

to different contexts: dealing with heterogeneity (Busch & Ralle, 2012; Grossenbacher, 2010), inclusion (Florian & Black-Hawkins, 2011), teaching and lesson planning (Morrison & Lederman, 2003; Vogt & Rogolla, 2009), linguistically-sensitive science teaching (Tolsdorf & Markic, 2016b) or teachers competences (Brookhart, 2011; Park & Oliver, 2008; Taber, 2005). These different contexts illustrate the relevance of diagnostic in past and today. But which knowledge do teachers and science teachers need for an appropriate diagnosis in their classes? What knowledge do teachers and science teachers possess? Finally, the development of a knowledge about the diagnosis for chemistry teaching and learning should be started during the university teacher training program in chemistry education. Therefore, it is interesting to know, how does the knowledge about diagnosis developed during their teacher training program in chemistry?

## THEORETICAL FRAMEWORK

In school, pedagogical-didactical diagnosis includes evaluation, assessment and/or observation of students' learning condition before and during the teaching and learning process aiming students' promotion (Ingenkamp & Lissmann, 2008). The diagnostic process starts with the assessment of the starting points (e.g. misconception, pre-knowledge) and the learning process, which help to plan and adapt the lessons to student' needs. This direction or type of a diagnosis is also described by Treagust (1988) or Barke, Hazaari, and Yitbarek (2009).

In the beginning, the diagnostician (here the teacher) sees a problem or a question which he or she wants to be answered through the diagnostic. For this, the teacher must choose the method and the instruments for the collection of the data. This is called as "preactional phase". The data collection at school is named "actional phase" (Klug et al., 2013). Fächter (2011) explains that the collected information helps to design the lesson and supporting students during the lesson. These implementations are carried out after the data collection in the "postactional phase" (Klug et al., 2013). The teacher re-evaluates and examines the impact of his or her support so that the diagnosis starts from the beginning (Heidemeier, 2008). Therefore, Klug and other (2013) describe the diagnosis process as a cyclical process, whereby one cycle is divided into three above described phases. This structure helps teachers to integrate diagnosis into science teaching. For all these steps, a science teacher needs a knowledge, either of methods, analyzing or instruments.

Schrader (2009) describes two directions of research in the field of diagnostic competence. The first direction focusses on the correctness of a diagnosis concerning diagnostic instrument, (e.g. Coladarci, 1986). Although the support is essential for a diagnosis through the diagnostic process, the focus here is only on the assessment or judgment. Dissimilar to this, Klug (2011) attempts to consider the support in the definition of diagnostic. She describes it as the "ability to interpret students' academic growth and their growth using learning strategies" (Klug, 2011, p. 14). This is the second direction, which is similar with the term of formative assessment (Bennet, 2011).

The named study of Coladarci (1989) describes a high correlation of exactness of assessment and the rank of students in the classroom. The same is described in the study by Feinberg and Shapiro (2003), but they found less correlation of the assessment to a research instrument regarding the level of students. A follow-up study shows that teachers overestimated the students' achievement (Feinberg & Shapiro, 2009), but teachers can also underestimate the students' abilities (Begeny, Eckert, Montarello, & Storie, 2008). Südkamp, Kaiser, and Möller (2012, 2014) showed in their study that students can be very well rated by the teacher and the level of the students was underestimate. Interesting are the results of this study for a second diagnosis after a teaching unit. They indicate that the second diagnosis is underestimated when in the first one a student was rated as low-achiever. Thus, the research in the field of exactness of diagnosis is still different, but it seems to be that teachers can assess students in the ranking. Hoge and Coladarci (1989) analyzed the correlation of teachers' work experience and their exactness of diagnostic. Not surprisingly, the working experience has a positive influence on the correct judgment (see also Bates & Nettelbeck, 2001). But Schrader (2009) describes that the number of years in a job doesn't correlate to teachers' diagnostic knowledge. He explains it by the fact that the numbers of years in a job are only an indicator for the experience but not for their diagnostic knowledge.

Talking about the teacher's knowledge or the professionalization of teachers, the concepts of Pedagogical Content Knowledge (PCK) and Pedagogical Knowledge (PK) are widely used in science education research. These knowledge domains often refer to the constructs of Shulman (1987), for example by Loughran, Berry, and Mulhall

(2006). The PCK is a knowledge domain that specific is for the subject (Park & Oliver, 2008; Shulman, 1987). All further developed models of PCK (e.g. Hashweh, 2005; Loughran, Berry & Mulhall, 2006 or Park & Chen, 2012) include aspects of activities to observe students. Loughran and other (2006) expand this with knowledge of the diagnosis of students' prior knowledge and possible diagnostic instruments. Knowledge about misconceptions is also essential for a teacher to recognize and handle with the conceptions and misconceptions of students (Olszewski, 2010). The hexagon model by Park and Oliver (2008) described the PCK in six categories and one of these is the "Knowledge of Assessment", which have a big relationship to the pedagogical diagnosis. However, the problem is that the diagnosis is not a single category in these models but includes some activities which has also a relationship to other categories as well. Ohle and McElvany (2015) show the different description of the diagnostic knowledge in the literature. Krauss et al. (2004) stated that diagnostic knowledge or competence isn't only one competence or domain, but the diagnostic competence is a multiple facets of competencies. For this purpose, a knowledge domain is needed that can be assigned to some subcategories of the PCK (and PK). Thus, the question of the definition of diagnostic knowledge must be raised.

The psychologist Jäger (2006) pronounce knowledge domains which are relevant to diagnostician in general. He uses the term of "diagnostic competence", but he described it through knowledge domains. Jäger (2006) describes the diagnostic competence with the help of sixth knowledge domains, which he classifies from a comprehensive analysis of a diagnostic process. The description by Jäger (2006) is defined for the psychological diagnosis. However, four knowledge domains are important for the pedagogical diagnosis in schools (see Tolsdorf & Markic, 2016b).

1. *Conditional Knowledge*: defines teachers' knowledge about students' backgrounds which is important for chemistry teaching, including influences that effect teaching and learning. Additionally, the influences and effects of heterogeneity or diversity on chemistry teaching are also included.
2. *Technological Knowledge*: the ability to select the most appropriate data collection for the actional phase. Knowledge about methods and instruments is needed, including their advantages and disadvantages. This knowledge domain also includes methods for analyzing the obtained data.
3. *Knowledge of Change*: refers to the pre-actional phase, therefore the further development of students. It means the strategies to deal with changing the resulting experience or behavior of the students in chemistry teaching. For example, knowledge about dealing with misconceptions or about aspects of linguistically sensitive teaching is important.
4. *Competence Knowledge*: includes the awareness of and attitudes towards diagnostics. For schools, this means that teachers be able to integrate a diagnosis into their teaching and adapt the lesson plan. Jäger also described this knowledge as the ability to answer a question. If a teacher doesn't possess these skills, then his or her personal knowledge of the topic must be expanded or a more competent person must be sought out for assistance.

Starting from Jägers definition and the above named three steps of a diagnostic process by Klug and other (2013) parallels are to be seen. The different knowledge domains by Jäger describe the three steps by Klug et al. (2013). Thus, these three knowledge domains are important for the implementation of the pedagogical-didactical diagnosis into the lesson (Füchter, 2011).

Keeping this definition in mind, different studies are conducted, however, they are focusing on single knowledge domains by Jäger. Yeh (2009) particularly emphasize the potential for use of regular, simple and valid instruments (special for reading and arithmetic). According to this, the "technological knowledge" is obligatory. However, teachers' knowledge about diagnostic instruments isn't very pronounced, even though many diagnostic instruments have already been published, e.g. diagnosing misconceptions (Morrison and Lederman, 2003; Spinath, 2005). Turner, van der Heide, Fyneweaver and Shavelson (2011) or Gilbert, Justi, van Driel, de Jong and Treagust (2004) mention similar problems, that teachers don't have knowledge about misconception and research results about it, whereby the teachers cannot diagnose and support their students in the way it is needed. Furthermore, if teachers studied and developed tools and methods for a diagnosis, then these teachers are able to apply or adapt to other topics in teaching (Wiliam, Lee, Harrison & Black, 2004). In addition to the measurement, teachers need for an appropriate diagnosis a knowledge of interventions, this means a knowledge to adapt their teaching to students' needs (Black & Wiliam, 2009; Heward, 2003) and a knowledge to use the collected data (Capizzi & Fuchs, 2005). In the named study by Capizzi and Fuchs (2005) the focus is on how far teachers change or adapt the teaching when he or she receives comprehensive information from the students about their reading skills through standardized tests. It shows that teachers in this study don't use these information for their own lessons. Thus, a broad knowledge and a further knowledge domain must be available (see knowledge of change). Finally, Ohle, McElvany, Horz, and Ullricht (2015) mention the beliefs, motivation and attitudes to the pedagogical diagnosis as an important aspect for teachers. The beliefs affect the use of a diagnosis for chemistry teaching. This is described in the fourth domains, the "competence knowledge". A positive attitude towards diagnosis is needed.

To sum up, all four knowledge domains are important for chemistry teaching. The development needs to be started during the university teacher training program, but very few studies investigate the development of diagnostic competence in the chemistry teachers training program. Tolsdorf and Markic (2017) showed in their cross-level study that student teachers' diagnostic competence differs at different stages of their teacher training program. It is also to see that courses and internship influence their development in different ways. However, it is not clear in which way the different phases of university teacher training program influence on their knowledge domains about diagnostic.

## RESEARCH QUESTIONS

Some research studies examine the diagnostic competence, but these studies are limited to the exactness of a diagnostic. Few studies investigate the development and the support of diagnostic competence in the teacher training program, especially the teacher training program in chemistry education (above and Tolsdorf & Markic, 2016b).

Therefore, the present case study attempts to investigate the development of diagnostic competencies of student teachers in chemistry education. In addition, concepts and implications are to be derived that promotes the development of these competencies in the teacher training program, which are helpful for further teacher trainees. Thus, following research questions can be arising:

1. Which diagnostic competence according to Jägers' definition do student teachers possess at different time points during her or his teacher training program in chemistry education?
2. How does the diagnostic competency according to Jägers' definition develop in the teacher training program in chemistry education?
3. Are specific or individual knowledge domains of the diagnostic competence by Jäger influenced by seminars and/or internships of the teacher training program in chemistry education?
4. What factors encourage or hinder the development of the diagnostic competence by Jäger during the teacher training program in chemistry education?

## CONTEXT OF THE STUDY

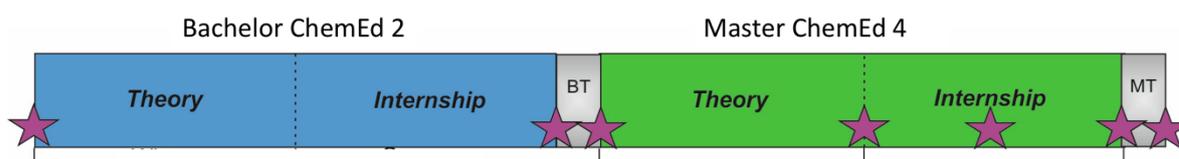
For answering the research questions, different time points of data collection are selected during the student teachers training program. Student teachers at our university visit two university modules regarding the topics of diagnosis, support, heterogeneity, diversity and lesson planning. The first module is called "Chemistry Education 2" (ChemEd 2) and consist of two seminars which start in the fifth-semester of bachelor and it is over one semester. One seminar "Diagnosing and planning of chemistry lessons" focuses on possibilities to diagnose students in chemistry lessons and the lesson on itself. However, the first step in this seminar is sensitization of student teachers for heterogeneity and diversity in chemistry classes. Further, student teachers learn diagnostic instruments for diagnosis of students' linguistic skills, their performance, and experimental skill. In the second seminar, different teaching and learning methods are discussed and student teachers experience these methods in the seminar. The student teachers learn these methods in the sense of a "pedagogic biplane" (translated from Wahl, 2006). First, the method is explained in a short introduction. Second, student teachers are learning details about the method by doing the method in their seminar, e.g. learning-at-station about learning-at-station or jigsaw classroom about jigsaw classroom. The aim is to teach the theoretical content and to let student teachers experience the work on the method from the view of their future students at the school.

Therefore, all knowledge domains by Jäger (2006) are the part of the seminars. Additionally, student teachers visit a short internship in teams of 10 hours teaching. At the beginning of this internship, student teachers observe their mentor's lessons for about 5 teaching hours and diagnose both the students and the classroom situation. In the second half of the internship, the students should plan the lessons based on their diagnosis. The teaching can be done in "team-teaching" or the two student teachers alternate the teaching time. The student teachers choose between these two options themselves together with their mentor at school. The second possibility has the advantage that the second student teachers can diagnose the learning process of individual students during the teaching of the first student teacher. After this module, student teachers decide whether they want to write their bachelor thesis in chemistry or chemistry education.

The master of education starts after the bachelor thesis with the second module on this topic. The second module about diagnosis, support and heterogeneity is called "Chemistry Education 4" (ChemEd 4). ChemEd 4 is divided in two phases over two semesters. The first phase is a theoretical seminar and it follows the long internship over five months. The focus of the seminar is on diagnosing individual students and adapting the own teaching according to the results of diagnosis. Concrete examples are the core of the seminar. Student teachers developed or adapted own instruments to different topics (e.g. misconception about chemical equilibrium) and discuss on

different materials and possibilities to support students in chemistry classes. Therefore, the interlinking of diagnosis and support is the main part of the present module. During the internship, the student teachers are visiting schools and are supervised by one mentor. First, they diagnose the lessons given by their mentors as well as the students of one class. The students conducted a comprehensive diagnosis on their own. While in the first small internship the diagnosis was on individual areas, now student teachers need to diagnose the breadth of heterogeneity. Diversity dimensions like students' linguistic skill, their experimental skills, content knowledge, misconceptions build the core. However, student teachers are free to choose any other dimension they see as important to diagnose. Student teachers are planning their own diagnostic process by choosing diagnostic instruments, collect and analyze data and use the results for their further teaching. Student teachers are teaching for about 12 lessons. During this period, they analyse (in the meaning of diagnostic process) their teaching and improvement in the classroom. Chemistry educator helps them while conducting the diagnostic process. Not seldom student teachers are having problems while implementing diagnosis and following support (see for more detail Tolsdorf & Markic, 2017). During their internship, student teachers have a consultation with the chemistry educator discussing the general issues during the internship and focusing on diagnosis and supporting student. Each consultation lasts between 30 and 60 minutes. After the internship student teachers do not attend any further modules in chemistry education. They end their university teacher training by a master thesis which can be supervised by chemistry educator.

Following the different stages of the university chemistry teacher training program at our university seven points in time are used for the data collection: (1) before ChemEd 2, (2) after the ChemEd 2 and before bachelor thesis, (3) after the bachelor thesis and before master phase starts, (4) after the first semester of ChemEd 4, (5) in the middle of the internship during the ChemEd 4, (6) after the internship during the ChemEd 4 and before master thesis, and (7) after the master thesis. The first interviews were in winter 2014 and the last data collection was done in March 2017. The overview is given in [Figure 1](#).



**Figure 1.** Overview of the data collection (stars represent the time points of collection)

## METHODS

The present case study analyses the development and changes of the chemistry student teachers' diagnostic competence during the chemistry teacher training in detail. For this reason and due to the gap in research, the case study is based on semi-structured interviews with open-ended questions. For answering the research questions, the questions are based on the four knowledge domains of the diagnostic competence by Jäger (2006). Because the individual domains aren't further differentiated, open-based questions are used to evaluate changes, development, and influences of and during the teacher training (see Miles, Huberman & Saldana, 2014). Open-based questions are useful for studies of attitudes, opinions, and beliefs (Converse & Presser, 1986), which are important for this study with regards to diagnostic process.

Data are collected in a period of about three years. To avoid the "Social-Desirability-Effects" (Weisberg, 2005) questions vary between the time points. The term of diversity is intentionally not used in the questions because the student teachers can be seldom distinguished this concept from the concept of heterogeneity. Some questions from the first-time point of the interview are listed below:

1. How would you define diagnosis?
2. Why should you diagnose in chemistry classes?
3. How did you perceive the heterogeneity until now?
4. What do you associate with heterogeneity?
5. Which diagnostic instruments do you know?
6. Describe instruments which are particularly important for you.
7. Do you know possibilities to adapt the teaching to the pupils?
8. How do you change the teaching while teaching chemistry in heterogeneous classes?

The questions one and two focus on Competence Knowledge and the following two questions on Conditional Knowledge. The Technological Knowledge is evaluated by questions five and six. The last domain, Knowledge of

Change, is addressed in the last two questions. An interview with the student teacher use this structure at any seven time-points and each interview lasted about 15-20 minutes.

## SAMPLE

Altogether, the data were collected from two chemistry student teachers while following them during their teacher training program. Both student teachers were participated in the study voluntary and were able to change their opinion any time. Both student teachers study chemistry as a subject with the aim of teaching at a comprehensive school. They are native German speakers and haven't a migration background.

Sandra (name given by the authors) was 24 years old at the beginning of the study. She visited grammar school and gained her university certificate there as well. She studies mathematics as the second subject. Already before the study, she gave different tutorial in chemistry and mathematics for high school students and she works voluntarily with children (also during the study). No further internships were completed additionally to the internship during her teacher training. The internships were all done at schools with high heterogeneity among students.

Lara (name given by the authors) is the second student teacher and her other subject is French. She was also at the grammar school. Her age is 24 at the beginning of the study and she offers volunteer sports lessons to young children. She didn't visit any further internships next to the obligatory one. Her internships were mainly at the grammar school, which are not characterized by large heterogeneity among students.

Both student teachers made similar school career and average grades in their study. They stand up by their willingness to learn and both wrote their bachelor and master thesis in chemistry education on topics of diagnosis – however with different foci.

Both student teachers participated voluntarily in this study during its time. Starting from a group of student teachers, these two student teachers were chosen based on their similar grads during the teacher training. Furthermore, their differences played a big role. A major difference are schools they visited during their internship. Another reason is their second subject. At the beginning, it was assumed that the student teacher with a language as a second subject is more sensitized to linguistic heterogeneity and the diagnostic competence in this area is more pronounced. The second student teacher has mathematics, which is often chosen by student teachers in conjunction with chemistry education.

## DATA ANALYSIS

Starting from the theoretical background, the research questions and collected data, documentary method is suitable to investigate the development of knowledge and emotional developments of student teachers (Asbrand, 2011; Ruhrig & Hötting, 2015), especially for a longitudinal study (Nohl, 2010; Trautrim, Grant, Cunliffe, & Wong, 2012). Bohnsack (2010) develops the method to analyze qualitative data, in which the analysis is from two perspectives. The first analysis of the data investigates the content ("what") and the emotions ("how") are examined in a second step (Bohnsack, 2010). The analysis can be divided into three parts: i. *Formulating interpretation* (or Rephrasing interpretation by Trautrim, Grant, Cunliffe, & Wong, 2012), ii. *Reflecting interpretation* and iii. *Comparative Analysis* (Bohnsack, 2010). The *Formulating interpretation* summarizes the content (meaning) and it reproduces the thematic structure of the interview. For this purpose, the text is divided step by step into sections and these sections are described with keywords, main topics and sub-topics (Bohnsack, 2010). The focus is only on the content of the interview. Therefore, the question is asked: What does the interviewee say? (Nohl, 2010). The second analysis asks how the content is said and described. The framework is to be reconstructed and explicated in which topics are elaborated. The analysis is done for the same section from the first formulating interpretation. The formal structure (exemplary, descriptive, contextually, experiences, ...) and the associated semantic aspects (word selection, repetitions, text structure, incomplete...) of the text are important to examine the interview (Trautrim, Grant, Cunliffe, & Wong, 2012). So, the reflective interpretation reconstructs the style or *modus operandi* of the statements in the interview (Bohnsack, 2010).

For the present study, at first the formal structure of the transcribed interviews was analyzed and, second, the audio file was investigated in term of emotions, speech and attitudes with the focus is on the "how". Bohnsack (2010) mention the traceability and documentation of the results and their interpretations for these two analysis steps, which must be discussed with other researchers. Therefore, the analysis of the interviews is done by two chemistry educational researcher from the field of heterogeneity, diagnosis, support and teacher training. Primarily, the two analysis steps should be examined separately and documented with a strong textual reference. Finally, the *comparative analysis* (Bohnsack, 2010) is to be conducted. Similarities and differences between cases (and in a case) need to be worked out using a case-by-case and case-by-cross comparison. This is a continuously circular process that combines the individualities of a case (Nohl, 2010). From this, types of cases can be generated by a

systematically and reasonably generalizing of cases (Trautrim, Grant, Cunliffe, & Wong, 2012). Thus, in a first step, an analysis was carried out focusing on single student teacher and comparing the results from different time points. Finally, in the second step, the chemistry student teachers were compared to one another.

## RESULTS

In the following, the results of the longitudinal interview study will be presented. To give a better overview, the results are summarized in **Table 1**.

**Table 1.** Development and comparison of the diagnostic knowledge during their studies

	<b>Sandra</b>	<b>Lara</b>
Before ChemEd 2	<ul style="list-style-type: none"> <li>o No awareness for diagnostic</li> <li>o Importance and meaning of diagnostic not seen</li> <li>o Heterogeneity as problem and challenge</li> <li>o Content knowledge as only diversity domain</li> </ul>	<ul style="list-style-type: none"> <li>o Diagnostic as „summative assessment “</li> <li>o Importance and meaning of diagnostic not seen</li> <li>o Reference back to her own time in school</li> <li>o Ignoring heterogeneity</li> </ul>
After ChemEd 2/ Before BT	<ul style="list-style-type: none"> <li>o Low awareness of a diagnostic process and diagnostic for chemistry teaching</li> <li>o Adaptive teaching is mentioned</li> <li>o Open/positive to heterogeneity with a focus on support of lower achievers</li> <li>o Written diagnostic tests as instruments</li> <li>o Differentiation, but no concrete description</li> </ul>	<ul style="list-style-type: none"> <li>o First ideas of a diagnostic process</li> <li>o Insecurity and fear to deal with heterogeneity</li> <li>o Fright from heterogeneity</li> <li>o Content knowledge and culture as diversity domains</li> <li>o Learner-centred methods</li> </ul>
After BT/ Before Chem Ed 4 – part 1	<ul style="list-style-type: none"> <li>o High awareness of diagnostic process</li> <li>o Aware of own lack in the diagnostic knowledge, but use technical terms</li> <li>o Self-confidence and positive to heterogeneity</li> <li>o Linguistic heterogeneity is described in detail (also with influences)</li> <li>o Description of instruments for chemistry lessons</li> </ul>	<ul style="list-style-type: none"> <li>o Low awareness of diagnostic process</li> <li>o See the importance of diagnostic</li> <li>o Heterogeneity is challenging for teachers</li> <li>o Recognizing of linguistic heterogeneity</li> <li>o Adaption of instruments from other subjects or</li> <li>o Own development helps for adaptive teaching</li> </ul>
After Chem Ed 4 -part 1	<ul style="list-style-type: none"> <li>o Heterogeneity as enrichment for all</li> <li>o Pupils are individual</li> <li>o Observation and interviews as instruments</li> <li>o Tests for language and content knowledge</li> <li>o Development of their own handling as teacher and planning of teaching</li> <li>o Possibilities of linguistic support</li> </ul>	<ul style="list-style-type: none"> <li>o Support is only goal of diagnostic process</li> <li>o Heterogeneity as positive however, afraid of dealing with it in classes</li> <li>o Content knowledge and performance are the main focus</li> <li>o Open-based instruments, e.g. drawing</li> <li>o Possibilities for dealing with misconceptions</li> </ul>
Middle of internship	<ul style="list-style-type: none"> <li>o Explanation of the diagnosis with examples and the use of technical terms</li> <li>o See the importance of diagnosis for chemistry teaching</li> <li>o Interest and motivation of pupils is important</li> <li>o Open-based instruments, e.g. drawing</li> <li>o possibilities to support pupils related to the collected information</li> </ul>	<ul style="list-style-type: none"> <li>o Diagnostic as data collection of information</li> <li>o Importance of the diagnostic decreases (e.g. it is too time consuming)</li> <li>o Less awareness of heterogeneity in practice</li> <li>o Own insecurity in the use of instruments</li> </ul>
After internship	<ul style="list-style-type: none"> <li>o Positive experiences in the internship</li> <li>o Fulfil students` needs</li> <li>o Focus on language, content knowledge, interest and migration background</li> <li>o Description of instruments with a discussion of (dis-)advantages</li> <li>o Differentiated description in detail and applications of support in chemistry teaching</li> </ul>	<ul style="list-style-type: none"> <li>o Single use of diagnosis</li> <li>o Big own insecurity in use or dealing with diagnosis and heterogeneity</li> <li>o To be fair with all pupils</li> <li>o Written tests are still important, but open-based instruments are gaining importance</li> <li>o Description of different possibilities for each dimension of the heterogeneity</li> </ul>
after MT	<ul style="list-style-type: none"> <li>o Starting point of a diagnosis is the support</li> <li>o Safe handling of diagnosis and a broad knowledge, but the routine is missing</li> <li>o Attitudes to adaption is uncertain, because she has a broad knowledge but the routines are missing</li> <li>o Broad and deep knowledge about “dealing” and a certainty for the use</li> <li>o Linking/reference of diagnosis and support</li> </ul>	<ul style="list-style-type: none"> <li>o Diagnosis as a “status diagnostic” for an adaptive teaching, similarities to summative assessment</li> <li>o Overchallenged with the topic and she hasn’t ideas for implementation</li> <li>o Contradictory attitude towards heterogeneity</li> <li>o Instruments for a selective diagnostic, with (dis-)advantages</li> <li>o Great uncertainty about support to considering the selected information</li> </ul>

### **First Time Point (Before ChemEd 2)**

Both of the student teachers are starting with similar requirements. As well Sandra as Lara does not see the importance and the meaning of the diagnostic for teaching and learning of chemistry. Since Sandra does not mention the need for diagnostic, Lara describes it as a summative assessment. However, Lara seems to be insecure since she laughs when she talks or when she expresses a lack of knowledge. Both of them are referring their knowledge to their school experience as a student, e.g. "So actually I have not thought about it, that diagnostics in the teaching profession make sense. I was a student in a grammar school. Yes, and then we have written tests for diagnostic at the time." (Lara, translated by the authors) Though both were visiting similar schools (grammar schools), Sandra describes and express heterogeneity as a problem and challenge that need to be dealt with. She repeats this often, which shows her importance to this. Her focus is on differences in students' content knowledge, e.g. "So from my own school days, I would describe the differences in the performance. There were students who were very good at the topics." (Sanda, translated by the authors). In contrast, Lara ignores the fact that chemistry classrooms are heterogeneous. She describes that she does not know exactly how it feels because she didn't experience it as a student.

### **Second Time Point (After ChemEd 2 and Before the Bachelor Thesis)**

Good to see is that after ChemEd 2 (first module) both student teachers are describing first ideas of a diagnostic process, however, it is not explained in details. Still both are focusing on students' content knowledge as a dimension of heterogeneity. Additionally, Sandra is mentioning students' migration background and Lara differences in students' culture. Both student teachers experienced heterogeneous chemistry classes during their short internship. Sandra describes this positive with a focus on low-achievers. It is noticeable that Sandra recognizes the meaning, use and importance of diagnosis in order to adaptively teach (e.g. student-centred teaching and learning) in a process, but not specially for chemistry teaching. Different written diagnostic instruments are only mentioned from Sandra. She talks very self-confident and sovereign. Her voice sounds happy and optimistic when she speaks about heterogeneity. Compared to Sandra, Lara describes only an approach of a diagnostic process. In additional, she expresses a lack of the knowledge about diagnostic and own insecurity for dealing with the heterogeneity. In contrast to the first interview, Lara's phrasing changes to thoughtful voice. She begins to answer very slowly and thoughtfully. Unless the questions relate to the possibilities of support or change. Her voice is always sinking, as if she is sad, frustrated or without ambition. Apparently, she has bad experiences in the internship, which have influenced her attitudes and belief to these topics. Thinking about her internship Lara said: "In my first internship, I thought, oh my God [German speech]. I was completely shocked and I was not sure anymore if career as a teacher was the right decision (...) But somehow it is also nice to see, .....because if we were all the same and we do not want to experiment, then it would not be profitable for the chemistry class." (Lara, translated by the authors).

### **Third Time Point (After the Bachelor thesis and Before ChemEd 4)**

Both student teachers wrote their bachelor thesis with the focus on diagnostic. They developed different diagnostic instrument for diagnosing students' linguistic skills. Thus, it is not surprising that both Sandra and Lara described linguistic heterogeneity as a main dimension in heterogeneous classes. However, only Sandra was able to describe the exact influence of differences of students' linguistic skills on chemistry teaching and learning. Sandra seems also to be confident and positive about teaching heterogeneous chemistry classes, e.g. "So the language does not just mean how to talk or to write something. And here many factors play a role: grammar or vocabulary. These are very different from student to student. This heterogeneity has implications for chemistry learning in the classroom." (Sandra, translated by the authors). She reflects about this problem and thereby she develops possibilities for their own teaching. However, she is complaining about her lack of knowledge when it comes to diagnostic. As before, also here, Sandra is more positive about heterogeneity on chemistry classes while Lara sees it as a challenge for teachers. Lara is missing examples of good practice. A high awareness of diagnostic for science and chemistry teaching are described by Sandra. Lara has a less awareness of diagnostic process, but she recognizes the importance of diagnosis for an adaptive teaching, e.g. "Diagnosis is important in order to adapt the knowledge transfer to the students preknowledge. Here, the diagnosis for the students must remain undetected and thereby show the understanding problems in the students." (Lara, translated by the authors). At the same time, she describes problems with the adaptation and development of assessment instruments in the work of the bachelor thesis, e.g. use of C-Test from German education for chemistry teaching. For the first time, some instruments are mention, whereby Sandra describes the instruments much more detailed.

#### **Fourth Time Point (After ChemEd 4 – Seminar and Before the Internship)**

The next data collection is after visiting the seminar of the second module. The awareness of diagnosis is present among both student teachers, especially the importance of diagnosis for chemistry is the main focus in their descriptions. For Sandra, the main reason for diagnosis is to “pick up the students” where they are standing. Lara is going into the same direction and explains that diagnosis is maintained if supporting students in the right way. Overall, the attitudes towards diagnosis are positive at this point.

In the term of heterogeneity, there are differences between the student teachers. On the one side, Sandra sees the heterogeneity as very enriching for chemistry teaching. For example, Sandra says about the changes in the teaching from the planning about elementary particle: “And I believe that this product at the end of the lessons would not have been made if the heterogeneity had not been in the classroom. (p2) And this has helped all the pupils, not only the good or only bad pupils” (translated by the authors). Further, it is noticeable in her argumentation that she describes more and more pro and also contra arguments. She often speaks of problems or challenges with the implementation in practice, but she thinks further and develops her own solution. For these reasons, it can be assumed that Sandra speaks reflectively. Lara also feels the heterogeneity rather positive. She is looking forward to experience teaching in heterogeneous classes, but at the same time, she is afraid of it. This conflict of joy and insecurity can also be seen in other dimensions, e.g. for the use of diagnosis in chemistry teaching. Sandra distinguishes between her two main arts of diagnostic instruments. For diagnosing students` content knowledge, the observation and interviews are important. Written tests are important to diagnose the linguistic heterogeneity. Lara only mentions open-based instruments, which she wants to use for diagnosis of misconception.

#### **Fifth Time Point (Middle of the Internship)**

In the middle of the long internship, Sandra focuses in her description of diagnosis on scientific language (next to students` interest and their motivation) and explain it on example. She stressed that diagnosis is particularly important for planning and support, but also mentions that it is time-consuming. Lara mentions that diagnosis isn` t a process and the diagnosis only uses to collect information about students` knowledge. Since she doesn` t experience different dimensions of heterogeneity in practice, she focuses on the differences in students` content knowledge. The importance of diagnosis and the awareness is still in evidence by Sandra. For Lara, the importance of the diagnosis decreases, because she thinks that this topic is too complicated, time-consuming and she does not experience any support from the mentor in the internship. In addition, she says that she feels very insecure about handling and use of instruments in practice. Sandra mentions for the first time that open-based instruments are suitable in practice to better understand pupils` expertise about chemistry.

#### **Sixth Time Point (After the Internship)**

After the long internship, Sandra again mentions the important of diagnosis for chemistry teaching. She also explains in detail that diagnosis is a cyclical and repetitively process. During the internship, she has gained good or positive experience while teaching and diagnosing heterogeneous classes. In contrast, Lara describes diagnosis as isolated from the teaching. Lara doesn` t reports good experiences about heterogeneity from the internship. As in at the beginning she describes differences in students` content knowledge as main dimension. The language of Lara contains many interruptions and it is unstructured. Her statements are always repeated. While repeating, her voice is more confident and she speaks faster. Therefore, it seems as she as less reflected on these issues. It seems that she didn` t reflect her experience very deeply and thus she couldn` t elaborate good solutions. This is also evident on her “unhappy” way of expression, e.g. “I think, yes, I think, yes, exactly, that`s just positive, but also negative, (p2) also negative. Yes altogether (um) it is just positive to see, especially what social skills means. (...) but the heterogeneity in subject teaching is something else again. (...) And from a student`s point of view, this can also be very exhausting.” (Lara, translated by the authors). For the technological knowledge, written test as instruments for the assessment of the content knowledge are important for both participants. Additionally, open-base instruments are useful for chemistry teaching but these instruments are difficult to integrate in the lesson. Sandra names different instruments describes advantages and disadvantages of those.

#### **Seventh Time Point (After the Master Thesis)**

Similar results are obtained at the last time point of data collection. According to Sandra, supporting students is the main goal of diagnosis. Lara mentions that diagnosis can only be used at certain points and her description has similarities to a summative assessment. She felt overchallenged to integrate a diagnostic process in her teaching. Her insecurity is also mentioned for use of instruments and the changing of the lesson for a support. For both student teachers, the content knowledge is the main dimension of heterogeneity in chemistry teaching. Only Sandra

lists the importance of further dimension, which must be considered in the lessons. But, she also expresses the lack of routines which makes her feel insecure.

At all-time points, it is noticeable that both student teachers don't describe the influences of the heterogeneity on chemistry teaching. Also, they don't name any reasons for the heterogeneity in chemistry teaching. Only Sandra rarely mention this, but a structure or regularity couldn't be established. Further, both student teachers don't mention any methods or possibilities of analyzing diagnostic instruments during the whole study.

## DISCUSSION AND IMPLICATIONS

Overall, the developments and changes of both student teachers are very different, but similarities can be identified as well. Strongest changes are noticeable during and after the internships. For the competence knowledge, Sandra is developing very well and her awareness of an appropriate use of a diagnosis for chemistry lessons increases. Sandra's attitudes towards heterogeneity and diversity are positive during her study and she always describes potentials for chemistry teaching. In the end, a positive view is still noticeable, even she describes challenges. On the one hand, Sandra is social engagement and works voluntary with children, which seems as an advantage for the development of her attitudes. On the other hand, the school and the mentor seem to encourage her attitudes, because she emphasizes the mentor's support for handling the heterogeneity in the interview. Until the last and long internship, Lara's development is positive, but it is slower compared to Sandra. Starting from her development and the interviews on itself, it seems that Lara didn't reflect deeply on this topic. The reason here are only to be assumed. Since Lara expresses insecurity and being afraid at the beginning, it is to assume that she didn't want to be confronted with those feelings. Lara's development could be seen as stopped because she repeats view of heterogeneity as challenge and problems. In addition, her language sounds not so positive and reflective as Sandra's. It seems that her internship and especially her mentor were not well chosen for her development and that here better matching and the analysis of the conditions is needed. It seems that Lara needed more supportive mentor instead of one which offers her possibility to try what she wants. Sandra also talks about challenges and problems with implementation in practice, but she seems to reflect a lot about her experiences. It is to conclude that the role of the mentor but also the personality of the student teachers plays a high role in this phase. Thus, chemistry educator should have short interviews with their student teacher before their internship on their expectations and needs during the internship. The individual matching of student teacher and mentor should be in focus before the internship as well. Furthermore, Sandra reported that she had problems by the implementation and her mentor talks about difficulties, but the mentor helped her anyway and wants to learn himself. Here, the mentor has embarked on the new topic. Lara complains that the teachers couldn't help her and these have to be trained before. Thus, mentors need to be trained on these topics. Some universities offer extra in-service training for the future mentors. Here also the aims and the procedure of the internship is more detail discussed with the mentors as well.

In general, both participants in this study learn more heterogeneity dimensions over time, but they focus mainly on differences in students' content knowledge and their linguistic skill. Especially Lara perceives other dimensions of heterogeneity rather unimportant (similar to Gilbert, 2004; Tolsdorf & Markic, 2017). A similar development tendency can also be found in the technological knowledge. Sandra's knowledge expands and she recognizes the respective (dis-)advantages during the internship. Lara expands her knowledge as well, but during the internship, she feels an uncertainty and previously mentioned instruments are unimportant. Similar is to be found in the study by Turner, Van der Heide, Fynewever, and Shavvelson (2011) where a lack of knowledge of diagnostic instruments and the implementation of instruments is found among chemistry teachers. The importance of the summative assessment by teachers in the school practice is already mentioned by other researchers (e.g. Bennet, 2011; Ingenkamp & Lissmann, 2008). But this importance can only be emphasized in the study with Lara.

Both student teachers describe own insecurity toward teaching in heterogeneous chemistry classes. This development can be related to the greater knowledge. Sandra described this challenge or problem in the adaptation of materials and instruments to the teaching topic (Bleas et al., 2012). Lara describes many areas of problems, but frequently she speaks of more practice in her study and conduction of diagnostic process. This result is similar to the study by Capizzi and Fuchs (2005), where teachers couldn't adapt the teaching to a diagnostic.

The reflection with the collected experience during student teachers' internship is important, in order to promote or further development of student teachers and their attitudes. This is particularly obvious in the second step of the analysis. Sandra speaks much more reflected and always describes advantages and disadvantages, before making a decision. Lara's descriptions are less reflected. One of the reasons for the deeper reflection of Sandra in term of diagnosis and heterogeneity might be the second subject, mathematics because the mathematics education has also the focus on heterogeneity, diagnosis, and support in the teacher training program. Thus, student teachers need more space and opportunities so that they can be guided reflect on their experiences with the diagnosis, support, and heterogeneity in chemistry teaching during the whole internship. It is to suggest that

deeper reflection on the different topics is needed. Since the lack of time during the seminar, one option would be designing the seminars in the meaning of flipped classroom.

Finally, we should not forget that our future chemistry teachers in our seminars are heterogeneous groups as well. Maybe they start with similar requirements, but they differ in their personality, need for support and ability to reflect. The question must be arisen, how much do we plan our seminars and the internships considering the heterogeneity of our groups. More focus on these issues should be put in the university teacher training program.

## ACKNOWLEDGEMENTS

We gratefully acknowledge the funding and support of the German Telekom Foundation (Deutsche Telekom Stiftung).

## REFERENCES

- Asbrand, B. (2011). The meaning of peer culture for learning at school: the example of a student company. *EDT – Educacao Tematica Digital*, 12(2), 53-76. <https://doi.org/10.20396/etd.v12i2.1188>
- Barke, H.-D., Hazaari, A., & Yitbarek, S. (2009). *Misconceptions in chemistry: Addressing perceptions in chemical education*. Berlin: Springer.
- Bates, C., & Nettelbeck, T. (2001). Primary school teachers' judgments of reading achievement. *Educational Psychology*, 21, 177-187. <https://doi.org/10.1080/01443410020043878>
- Begeny, J. C., Eckert, T. L., Montarello, S. A., & Storie, M. S. (2008). Teachers' perceptions of students' reading abilities: An examination of the relationship between teachers' judgments and students' performance across a continuum of rating methods. *School Psychology Quarterly*, 23, 43-55. <https://doi.org/10.1037/1045-3830.23.1.43>
- Bennett, R. E. (2011). Formative assessment: a critical review. *Assessment in Education: Principles, Policy & Practice*, 18(1), 5-25. <https://doi.org/10.1080/0969594X.2010.513678>
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5-31. <https://doi.org/10.1007/s11092-008-9068-5>
- Bohnsack, R. (2010). Documentary Method and Group Discussions. In R. Bohnsack, N. Pfaff & W. Weller (Eds.), *Qualitative analysis and documentary method in international educational research* (pp. 99-124). Opladen: Barbara Budrich.
- Brookhart, S. M. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice*, 30(1), 3-12. <https://doi.org/10.1111/j.1745-3992.2010.00195.x>
- Busch, H., & Ralle, B. (2012). Fachsprachliche Kompetenzen prüfen und fördern [Diagnosing and promoting scientific language competences]. In S. Bernholt (Ed.), *Konzepte fachdidaktischer Strukturierung für den Unterricht* (pp. 578-580). Berlin: LIT.
- Capizzi, A. M., & Fuchs, L. (2005). Effects of curriculum-based measurement with and without diagnostic feedback on teacher planning. *Remedial and Special Education*, 26(3), 159-174. <https://doi.org/10.1177/07419325050260030401>
- Chandrasegaran, A. L., Treagust, D. F., & Mocerino, M. (2007). The development of a two-tier multiple-choice diagnostic instrument for evaluating secondary school students ability to describe and explain chemical reactions using multiple levels of representation. *Chemistry Education Research and Practice*, 8(3), 293-307. <https://doi.org/10.1039/B7RP90006F>
- Coladarci, T. (1986). Accuracy of teacher judgment of students' responses to standardized test items. *Journal of Educational Psychology*, 78(2), 141-146.
- Converse, J. M., & Presser, S. (1986). *Survey Questions. Handcrafting the standardized questionnaire*. Sage University Paper series on Quantitative Applications in the Social Sciences (Series/Number 07-063). Beverly Hills, CA: Sage.
- Feinberg, A. B., & Shapiro, E. S. (2003). Accuracy of teacher judgments in predicting oral reading fluency. *School Psychology Quarterly*, 18(1), 52-65. <https://doi.org/10.1521/scpq.18.1.52.20876>
- Feinberg, A. B., & Shapiro, E. S. (2009). Teacher Accuracy: An Examination of Teacher-Based Judgments of Students' Reading with Differing Achievement Levels. *The Journal of Educational Research*, 102(6), 453-462. <https://doi.org/10.3200/JOER.102.6.453-462>
- Florian, L., & Black-Hawkins, K. (2011). Exploring inclusive pedagogy. *British Educational Research Journal*, 37(5), 813-828. <https://doi.org/10.1080/01411926.2010.501096>

- Füchter A. (2011). Pädagogische und didaktische Diagnostik: Eine schulische Entwicklungsaufgabe mit hohem Professionalitätsanspruch [Pedagogical and didactical diagnostic: a developmental task of schools with demand on professionalisation]. In A. Füchter & K. Moegling (Eds.), *Diagnostik und Förderung Teil I: Didaktische Grundlagen* (Vol. 14)( pp. 45-83). Kassel: PROLOG.
- Gilbert, J. K., Justi, R., van Driel, J. H., de Jong, O., & Treagust, D. F. (2004). Securing a Future for Chemical Education. *Chemistry Education Research and Practice*, 5(1), 5-14. <https://doi.org/10.1039/B3RP90027D>
- Grossenbacher, S. (2010). Kompetenz und Professionalität entwickeln [Developing competences and professionalism]. In A. Buholzer & A. Kummer Wyss (Eds.), *Alle gleich - alle unterschiedlich! Zum Umgang mit Heterogenität in Schule und Unterricht* (pp. 162-168). Seelze: Kallmeyer.
- Hashweh, M. (2005). Teacher pedagogical constructions: a reconfiguration of pedagogical content knowledge. *Teachers and Teaching*, 11(3), 273-292. <https://doi.org/10.1080/13450600500105502>
- Heidemeier, H. (2005). *Self and supervisor ratings of job-performance: Meta-analyses and a process model of rater convergence* (Doctoral dissertation). Retrieved from <https://opus4.kobv.de/opus4-fau/frontdoor/index/index/docId/143>
- Helmke, A., & Schrader, F. W. (1987). Interactional effects of instructional quality and teacher judgement accuracy on achievement. *Teaching and Teacher Education*, 3(2), 91-98. [https://doi.org/10.1016/0742-051X\(87\)90010-2](https://doi.org/10.1016/0742-051X(87)90010-2)
- Heward, W. L. (2003). Ten faulty notions about teaching and learning that hinder the effectiveness of special education. *The Journal of Special Education*, 36(4), 186-205. <https://doi.org/10.1177/002246690303600401>
- Hoge, R. D., & Coladarci, T. (1989). Teacher-based judgments of academic achievement: A review of literature. *Review of Educational Psychology*, 59(3), 297-313. <https://doi.org/10.3102/00346543059003297>
- Ingenkamp, K., & Lissmann, U. (2008). *Lehrbuch der Pädagogischen Diagnostik* (6<sup>th</sup> edition). Weinheim: Beltz.
- Jäger, R. S. (2006). Diagnostischer Prozess [Diagnostic Process]. In F. Petermann & M. Eid (Eds.), *Handbuch der psychologischen Diagnostik* (pp. 89-96). Göttingen: Hogrefe.
- Klug, J. (2011). *Modelling and training a new concept of teachers' diagnostic competence* (Unpublished doctoral dissertation). Technische Universität Darmstadt, Darmstadt, Germany.
- Klug, J., Bruder, S., Kelava, A., Spiel, C., & Schmitz, B. (2013). Diagnostic competence of teachers: A process model that accounts for diagnosing learning behaviour tested by means of a case scenario. *Teaching and Teacher Education*, 30(2), 38-46. <https://doi.org/10.1016/j.tate.2012.10.004>
- Krauss, S., Kunter, M., Brunner, M., Baumert, J., Blum, W., & Neubrand, M. (2004). COACTIV: Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz [COACTIVE: Professional knowledge of teachers, cognitive active mathematics classes and the development of mathematics competence]. In J. Doll & M. Prenzel (Eds.), *Die Bildungsqualität von Schule: Lehrerprofessionalisierung, Unterrichtsentwicklung und Schülerförderung als Strategien der Qualitätsverbesserung* (pp. 31-53). Münster: Waxmann.
- Loughran, J., Berry, A., & Mulhall, P. (2006). *Professional learning: understanding and developing science teachers' pedagogical content knowledge*. Rotterdam: Sense.
- Miles, M., Huberman, A. M., & Saldana, J. (2014). *Qualitative Data Analysis. A Methods Sourcebook* (3<sup>rd</sup> edition). Thousand Oaks: Sage.
- Morrison, J. A., & Lederman, N. G. (2003). Science Teachers' Diagnosis and Understanding of Students' Preconceptions. *Science Education*, 87(6), 849-867. <https://doi.org/10.1002/sce.10092>
- Nohl, A.-M. (2010). Narrative Interview and Documentary Interpretation. In R. Bohnsack, N. Pfaff, & W. Weller (Eds.), *Qualitative analysis and documentary method in international educational res.* Opladen: Barbara Budrich.
- Ohle, A., & McElvany, N. (2015). Teachers' diagnostic competences and their practical relevance, Special issue editorial. *Journal for Educational Research Online*, 7(2), 5-10.
- Ohle, A., McElvany, N., Horz, H., & Ullricht, M. (2015). Text-picture integration - Teachers' attitudes, motivation and self-related cognitions in diagnostics. *Journal for Educational Research Online*, 7(2), 11-33.
- Olszewski, J. (2010). *The impact of physics teachers' Pedagogical content knowledge on teacher action and student outcomes*. Berlin: Logos.
- Park, S., & Chen, J. S. (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), 922-942. <https://doi.org/10.1002/tea.21022>
- Park, S., & Oliver, J. S. (2008). Revisiting the conceptualisation of pedagogical content knowledge (PCK): PCK as a conceptual tool to understand teachers as professionals. *Research in Science Education*, 38(3), 261-284. <https://doi.org/10.1007/s11165-007-9049-6>

- Ruhrig, J., & Höttecke, D. (2015). Components of Science Teachers' Professional Competence and Their Orientational Framework when Dealing with Uncertain Evidence in Science Teaching. *International Journal of Science and Mathematics Education*, 12(2), 447-465. <https://doi.org/10.1007/S10763-015-9628-3>
- Schrader, F. W. (2009). The Diagnostic Competency of Teachers. *German Journal of Educational Psychology*, 23(3/4), 237-245. <https://doi.org/10.1024/1010-0652.23.34.237>
- Shulman, L. S. (1987). Knowledge and Teaching: Foundations of the New Reform. *Harvard Educational Review*, 57(1), 1-23. <https://doi.org/10.17763/haer.57.1.j463w79r56455411>
- Spinath, B. (2005). Akkuratheit der Einschätzung von Schülermerkmalen durch Lehrer/innen und das Konstrukt der diagnostischen Kompetenz [Accuracy of Teacher Judgements on Student Characteristics and the Construct of Competence]. *Zeitschrift für Pädagogische Psychologie*, 19, 85-95. <https://doi.org/10.1024/1010-0652.19.12.85>
- Südkamp, A., Kaiser, J., & Möller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *Journal of Educational Psychology*, 104(3), 743-762. <https://doi.org/10.1037/a0027627>
- Südkamp, A., Kaiser, J., & Möller, J. (2014). Teachers' Judgement of Students' Academic Achievement. Results from Field and Experimental Studies. In S. Krolak-Schwerdt, S. Glock & M. Bähmer (Eds.), *Teachers' Professional Development. Assessment, Training, and Learning* (Volume 3). (pp. 5-25). Rotterdam: Sense.
- Taber, K. (2002). *Chemical misconception - Prevention, diagnosis and cure: Volume I: Theoretical background*. London: Royal Society of Chemistry.
- Taber, K. (2005). Developing Teachers as Learning Doctors. *Teacher Development*, 9(2), 219-236. <https://doi.org/10.1080/13664530500200249>
- Tolsdorf, Y., & Markic, S. (2016a). Exploring student teachers' knowledge concerning diagnostics in science lessons. In J. Lavonen, K. Juuti, J. Lampiselkä, A. Uitto & K. Hahl (Eds.), *Electronic Proceedings of the ESERA 2015 Conference. Science education research: Engaging learners for a sustainable future, Part 13* (pp. 2002-2009). Helsinki, Finland: University of Helsinki.
- Tolsdorf, Y., & Markic, S. (2016b). Dealing language in science classrooms - Diagnosing students' linguistic skills. In: S. Markic & S. Abels (Eds.), *Science Education towards Inclusion* (pp. 23-42). New York: Nova Publishing.
- Tolsdorf, Y., & Markic, S. (2017). Exploring chemistry student teachers' diagnostic competence - a qualitative cross-level study. *Education Sciences*, 7(4), 86-100.
- Trautrimis, A., Grant, D. B., Cunliffe, A. L., & Wong, C. (2012). Using the "documentary method" to analyse qualitative data in logistics research. *International Journal of Physical Distribution & Logistic Management*, 42(8/9), 828-842. <https://doi.org/10.1108/09600031211269776>
- Treagust, D. F. (1988). The development and use of diagnostic instruments to evaluate students' misconceptions in science. *International Journal of Science Education*, 10(2), 159-169. <https://doi.org/10.1080/0950069880100204>
- Turner, M., Van der Heide, K., Fynewever, H., & Shavelson, R. J. (2011). Motivations for and barriers to the implementation of diagnostic assessment practice - a case study. *Chemistry Education Research and Practice*, 12(2), 142-157. <https://doi.org/10.1039/C1RP90019F>
- Vogt, F., & Rogalla, M. (2009). Developing adaptive teaching competency through coaching. *Teaching and Teacher Education*, 25(8), 1051-1060. <https://doi.org/10.1016/j.tate.2009.04.002>
- Wahl, D. (2006). *Lernumgebungen erfolgreich gestalten. Vom trägen Wissen zum kompetenten Handeln* (2<sup>nd</sup> Ed.) [Successful development of learning settings. From lazy knowledge to competent action]. Klinkhardt: Bad Heilbrunn.
- Walter, J. (2011). Die Messung der Entwicklung der Lesekompetenz im Dienste der systematischen formative Evaluation von Lehr- und Lernprozessen [The measurement of development of reading competence in service of systematical formative assessment of teaching and learning processes]. *Zeitschrift für Heilpädagogik*, 6, 204-217.
- Weisberg, H. F. (2005). *The total survey error approach. A guide to the new science of survey research*. Chicago: The University of Chicago press.
- William, D., Lee, C., Harrison, C., & Black, P. (2004). Teachers developing assessment for learning: impact on student achievement. *Assessment in Education*, 11, 49-65. <https://doi.org/10.1080/0969594042000208994>
- Yeh, S. S. (2009). Class size reduction or rapid formative assessment?: A comparison of cost-effectiveness. *Educational Research Review*, 4(1), 7-17. <https://doi.org/10.1016/j.edurev.2008.09.001>