


## Article

# Exploring Chemistry Student Teachers' Diagnostic Competence—A Qualitative Cross-Level Study

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

<sup>1</sup> Institute for Science Education—Chemistry Education, University of Bremen, Leobenerstraße NW2, Bremen 28334, Germany

<sup>2</sup> Institute for Science and Technology—Chemistry Education, Ludwigsburg University of Education, Reutealle 46, Ludwigsburg 71634, Germany; markic@ph-ludwigsburg.de

\* Correspondence: y.tolsdorf@uni-bremen.de; Tel.: +49-421-218-62834

Received: 23 August 2017; Accepted: 30 November 2017; Published: 2 December 2017

**Abstract:** Diagnostic competence is an important skill of professional teachers and is a part of Pedagogical Content Knowledge (PCK). Because increasing heterogeneity and diversity characterize our schools, diagnostic competence plays a prominent role in teacher pre- and in-service courses. At our university, two modules are focused on examining student teachers' diagnostic competence in chemistry. The present paper describes a cross-level case study analyzing three groups of student teachers in different phases of their university teacher training program. The study is based on a qualitative research study with a total of 108 participants. The results show both positive and negative developments in teacher trainees' diagnostic competence. They show that attitudes and beliefs about heterogeneity can be changed from a rather negative viewpoint to a more positive one. The results will be presented and discussed below. Suggestions for further development will be given.

**Keywords:** chemistry student teachers; cross-level study; development of diagnostic competence; university teacher training program

## 1. Introduction

The “PISA shock” in Germany is a recurring topic in education. It reveals that heterogeneity in our schools plays an increasingly important role in all school subjects. One possibility for properly dealing with the heterogeneity is diagnostics, which is currently a main field in education research, (e.g., [1–3]).

The term “diagnosis” is generally associated with medical care. A person goes to the doctor if he or she feels ill and the doctor questions the patient about relevant symptoms, so that a diagnosis can be reached. In addition, the doctor performs different tests (such as observing body temperature or ordering blood or urine tests) to get more information for an accurate diagnosis. A diagnosis does not end with recognition of the illness, but rather with a prescribed regimen of treatment. The doctor must continuously observe the patient and continue to react to inappropriate treatments. Without diagnostics, a medical expert has only limited possibilities for carrying out meaningful medical treatment. It should also be noted that a good diagnosis also includes corollary medical treatment. Despite the fact that diagnosis is normally perceived as an instrument in medicine, the idea can also be effectively used for teaching.

Science teachers need to understand their students' learning status. This is the idea behind constructivist lesson planning. Teaching is adapted to the learning condition and situation of the student (e.g., [4]). Because of the proximity to medicine, Taber (2005) [5] uses the term “learning doctor” in his article about courses for diagnosing students' misconceptions. However, diagnosis in

the science classroom is more complicated because the teacher must diagnose the whole class and not just one person.

Teachers need knowledge about diagnosis in order to assess, identify, and support learning processes and difficulties [6]. Thirty years ago, chemistry teacher Treagust (1988) [7] stated that diagnostic instruments help teachers to apply the research of misconceptions in lesson planning. Even today, diagnostic instruments are important components for identifying misconceptions and supporting performance among students (e.g., [8,9]). For a long time, students' misconceptions and differences in students' content knowledge have been the focus of the research. However, diagnoses are not just made in the context of misconceptions. Nowadays, chemistry classes are becoming increasingly heterogeneous. This is especially true in western European countries like Germany, Finland, France, the United Kingdom, and Ireland, where these changes are very noticeable. One explanation lies in the currently high levels of migration from one country to another due to worldwide economic changes ("globalization"). Students bring many different prerequisites to the classroom and teachers need to cope with each one. The differences are multifaceted and often overlap. In the literature in the USA, these differences are summarized using eight main dimensions. The so-called "Big 8" are age, gender, ethnicity, religion, race, sexual orientation, functional role, and mental/physical ability [10]. Thus, diagnostics are also connected to: (i) the handling of linguistic heterogeneity [11]; (ii) dealing with inclusion [12]; (iii) planning a teaching unit [3,13]; and (iv) evaluating teacher competency [5,6,14]. These different contexts in which diagnosis is important demonstrate that teachers (science teachers in particular) must take the knowledge, interest, motivation, skills, and abilities of their pupils in the classroom into account (e.g., [15–18]). However, it has also become clear that diagnosis is a basic tool in schools and, therefore, for chemistry lessons, too.

## 2. Theoretical Background

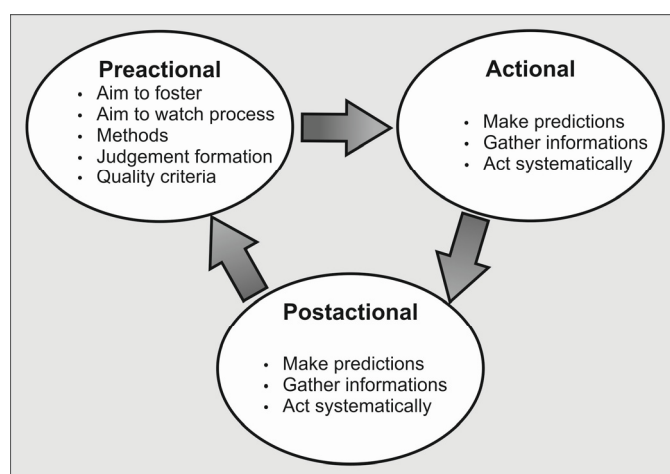
Ingenkamp and Lissmann (2008) [19] described diagnosis in school as being composed of all activities (mainly by the teacher) that help to understand and explain students' behavior, in order to optimize the overall learning process. This implies that identification of learning conditions, learner support, and recognition of possible learning barriers are all important aspects. Ingenkamp and Lissmann described it as follows: "A diagnosis includes all diagnostic activities which evaluate the conditions and the skill set of the individual or group learner, which can be observed during planned teaching and learning processes" [19] (p. 13). In the school context, a diagnosis without targeted work (support) is not profitable for the individual learning processes of the students. Two directions of pedagogical diagnosis are recognized in this wide understanding of diagnosis. On the one hand, pedagogical–psychological diagnosis includes the discovery of learning difficulties or lacking intellectual skills in the learner. To this end, experts are often involved or the teacher targets specific knowledge. Normally, a chemistry teacher does not perform such activities. On the other hand, pedagogical–didactical diagnosis means the determination of different starting points to learning and ends with assessment for grading at the end of a teaching unit [20]. Bell (2007) [21] has suggested that assessment should be integrated into lessons and should be an integral part of the learning process. Therefore, the traditional psychometric assessment should be rethought [21].

Within pedagogical–didactical diagnosis, two different types can be classified [22]. The final assessment of learning processes measures learning success to give students a grade for gained knowledge, competencies, or experimental skills. These marks are necessary for certain school and university degrees. Therefore, this kind of pedagogical–didactical diagnosis can have long-term consequences for the learner [6]. This is similar to the concept of summative assessment. The second type of diagnostic focuses on the learning process. It is used to adapt teaching to student needs, just like formative assessment [23,24]. Teachers use assessment for planning the learning process. But grading should remain separate from the learning process, so that the students perceive the diagnosis as a part of the support or learning process. This does not correspond to the idea of formative assessment,

but teachers face this challenge with both diagnostic types. Therefore, they need to combine both types [25,26].

Diagnostic measurements can focus on larger teaching units and longer support times. This is called macro-adaption. But in everyday teaching, diagnosis always occurs in smaller sequences (one lesson), or so-called micro-adaption [22]. Black and William (2009) [27] described three actors (teacher, student, and peers) who play a role in the learning process and thus act differently during the diagnostic process. A student can self-diagnose, but he or she can also examine classmates (peers). Thus, the transfer of individual tasks may be helpful for teachers. For formative assessment, three main questions aid in designing and implementing diagnosis in school: “Where are you trying to go? Where are you now? How can you get there?” [28] (p. 27). These questions can be seen as a guide in which the learning process of students is important. Hattie and Yates (2014) [29] described similar questions in their feedback model, and these questions help to reduce the gap between the learning level and the final goal.

The diagnostic process model presented by Klug and other [2] divides this process into three phases. These phases include the three above questions, among other things. The first phase is the pre-actional phase, which includes describing the goal of diagnosis, identifying problems, suggesting possible methods, and learning objectives as per the first question. The second phase is described as the actional phase. This includes the systematic collection of information (data collection). The last phase is the post-actional phase (see Figure 1).



**Figure 1.** Presentation of the diagnostic process by Klug and others [2].

This involves further work such as supporting students to reach specific learning objectives beyond their detected learning level [2]. Heidemeier (2005) [30] mentions that implementation (support) must be re-evaluated to see whether or not learning success has been achieved. Diagnosis and process are not one-off occurrences, but belong to a recurring cycle (e.g., [30,31]). Von Aufschnaiter and others [32] warned that many different understandings of diagnostics still exist. They describe four different kinds of diagnostics in the school context, all of which differ in their aims, methods, and objects. These include (i) the status diagnostic (one specific time point); (ii) the process diagnostic (a time point within the solution process); (iii) the change diagnostic (the changes between two time points); and (iv) the course diagnostic (changes taking place between two treatments) [32].

As mentioned, students’ misconceptions considering different topics in chemistry are well evaluated. In the literature, some diagnostic instruments are used to explore or determine students’ misconceptions in chemistry or science teaching (e.g., [33,34]). Barke et al. (2009) [8] and Taber (2002) [9] both give an overview of misconceptions and diagnostic instruments for them. However, for other dimensions of the diversity wheel [10], diagnostic instruments are less developed in chemistry lessons [11].

To recognize differences in other dimensions on the diversity wheel, teachers and teacher trainees need to know about diagnosis and how to use it to develop their own diagnostic instruments. It should be noted here that these instruments do not correspond to scientific quality criteria. Although research and policy both tend towards the consensus that a wide variety of diagnostic instruments is important for formative assessment, this picture is somewhat misleading. Most often only content knowledge at school is evaluated [25,35].

However, what kind of knowledge do (future) chemistry teachers need to have to diagnose their classes and to plan their lessons considering the results of diagnosis. Whenever teachers' professionalism is spoken about, the concepts of content knowledge (CK), pedagogical knowledge (PK), and pedagogical content knowledge (PCK) are generally cited. Many models use these three dimensions, which were first proposed by Shulman [36,37]. These models extend or differentiate the "trichotomy" (e.g., [14,38–40]). In science educational research, particular models of teachers' professionalism are very widely used, for example, the hexagon model by Park and Oliver [40] or the model suggested by Loughran [14,41]. Park and Oliver assign the ideas of pedagogical–didactic diagnosis to four of their six overall categories [40]. Loughran, Berry, and Mulhall (2012) [42] also describe the idea of diagnosis when they talk about adapting teaching to the student: "In the development of our work in PCK, we have drawn on this constructivist perspective so that one aspect of PCK which we have paid particular attention to has been related to the nature of teachers' knowledge that helps them to develop and apply teaching approaches that promote student learning in ways other than 'teaching as telling' ... " [42] (p. 16). As early as 1998, Black [43] had already stated that teachers need a particular set of knowledge, abilities, and skills for formative assessment. Teachers require knowledge about possible goals, methods, and ways to change and refine their own teaching [43]. But exactly which knowledge and skillsets do teachers need to possess? In the context of teacher professionalism and diagnosis, the term "diagnostic knowledge" or "diagnostic skills" is most often used (e.g., [2,20,22,31]). Diagnostic competence includes all abilities "to interpret students' growth and their growth using learning strategies" [31] (p. 14). Regarding the diagnostic process in this definition, the word "and" is an important term. Diagnostic competence does not limit itself to knowledge about assessment. Klug (2011) [31] described the need for comprehensive knowledge and certain skills in diagnosing students. Schrader (2013) [22] described two areas of diagnosis (see above), which differentiate between formative and summative assessment. Summative assessment has a long-term effect, for example, students' choice of a further career. This may be one reason why the research in the field of diagnostic competence has been limited to the quality of teacher assessment. This has been investigated with respect to the accuracy or correctness of judgement (e.g., [44–47]). The research examined the correlation between teacher assessment and judgment using a standardized test instrument (e.g., [47–51]). However, Südkamp, Kaiser, and Möller (2012) [52] showed that the teacher judgment correlates only moderately with student performance. They viewed this result as positive and appropriate, but also described a desire for a higher correlation ratio. On the one hand, teachers tend to overestimate the abilities of their learners [44]. On the other hand, Begeny, Eckert, Montarello, and Storie (2008) [53] detected an underestimation of these abilities. Bates and Nettelback (2001) [54] examined diagnostic competence with regard to professional experience (e.g., years in the job). The study revealed a tendency for experience to promote diagnostic competence. But no direct correlation could be detected between work experience and competence (e.g., [22]). Schrader explains that the duration of professional work is one indicator of professional experience. Furthermore, teachers do not tend to change their lessons, even if they receive diagnostic results through a standardized test of linguistic competence [55]. Black and William (1998) [43] states that teachers need support in (formative) assessment, because practical assessment is risky and costs a great deal of time and energy.

Krauss and other [56] explain that the term diagnostic competence has been used for a long time in research literature. This "tradition" is difficult to deal with, since so much heterogeneity exists in descriptions of diagnosis [23]. Therefore, we chose the multidimensional description by Jäger [57]

in our study and named it diagnostic competence. The psychologist, Jäger, described a complex diagnostic process for a general diagnostic, which is too extensive for the school context. This makes it unsuitable for teachers with limited time and resources at their disposal. In this model, he classifies six knowledge domains [57]. However, not all of the dimensions are important for teachers' diagnostic competence. The important ones are as follows:

1. *Conditional knowledge*: defines teachers' knowledge about students' backgrounds that is important for chemistry teaching, including influences that effect teaching and learning [57]. On the one hand, the dimensions of the "Diversity Wheel" [10] are addressed here. On the other hand, the influences and effects of heterogeneity or diversity on chemistry teaching are also included.
2. *Technological knowledge*: defines 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 [57].
3. *Knowledge of change*: refers to the pre-actional phase, therefore the further development of students. It refers to the strategies to deal with changing the resulting experience or behavior of the students in chemistry teaching [57]. For example, knowledge about dealing with misconceptions or about aspects of linguistically sensitive teaching is important [11].
4. *Competence knowledge*: includes the awareness of and attitudes towards diagnostics. For schools, this means that teachers are 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 does not 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 [57].

The three phases of the diagnostic process by Klug et al. [2] are described by the first three knowledge domains. These domains are important for teachers in order to integrate diagnosis into teaching in general [20] and into chemistry teaching in particular [11]. Ohle, McElvany, Horz, and Ullricht (2015) [58] state that teacher beliefs, motivations, and attitudes influence the use of diagnostics in science lessons. This includes the fourth knowledge domain and is one reason that this domain is so important for education. Fächter (2011) [20] recommends training teachers in this fourth domain, because knowledge in this area reduces the excessive demands on teachers with regard to diagnostics.

In spite of the importance of diagnostic competence among chemistry teachers, widespread knowledge on this topic remains seldom among educators. Thus, the development of the competence in this field needs to be a part of university teacher training programs (e.g., [54,58]). This is not to say that it is totally lacking from such programs. However, the amount and quality of such instruction in diagnostics varies widely from one university to another. This is the reason that studies on the influence of university teacher training courses in this area are not well known. Nor is the efficiency of such programs in modifying or enhancing such diagnostic skills well researched or publicized.

### 3. Methods

#### 3.1. Research Question

As just stated, research into chemistry teachers' diagnostic competence remains a scarcely explored area in the literature. Even less research is targeted towards the development of diagnostic competence during teacher training. However, the personal development of such competence must begin at the teacher training program level at university. Thus, the focus of this study is on student teachers of chemistry at the university of Bremen in Germany. Our research efforts are aimed at the development and modification of diagnostic competence in chemistry teacher trainees. From this starting point, the following main research questions emerge:



1. What level of diagnostic competence do student teachers possess at different stages of their university teacher training program in chemistry?
2. How does diagnostic competence differ among student teachers in varying semesters?

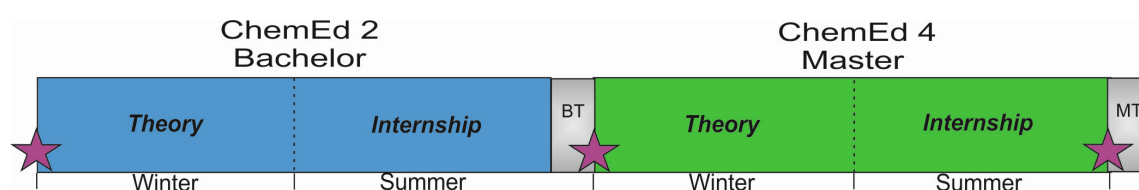
### 3.2. Context of the Research

In order to answer the above research questions, two university courses in chemistry education are required. They have diagnosis, heterogeneity, and diversity in chemistry education as central themes. Both courses were a part of the curriculum for a long time. The focus is now more stressed on the named topic, and new methods have been developed. As a cross-level study, this project chooses different points in time and separate groups of student teachers for data collection. One special aspect of both courses is the combination of theoretical learning units with practical phases (internships).

The first course is called “Chemistry Education 2” (ChemEd2). Participants in this course are Bachelor of Science students in their fifth semester. They have not had any other previous courses on this topic and have no internship experience. ChemEd2 combines two seminars and an internship, which is comprised of 12 h of team teaching in a school. The seminars focus on diagnosing and planning chemistry lessons and discuss various teaching methods. The focus during the internship is on diagnosing, planning, and teaching chemistry.

The second module is titled “Chemistry Education 4” (ChemEd4) and occurs in the first Master of Science semester of the program. No other courses in chemistry education occur between ChemEd2 and ChemEd4. It is a module that lasts for two semesters and begins with an introductory seminar. This seminar prepares teacher trainees for their four-month internship. It focuses on diagnosis skills and on how to deal with pupils’ misconceptions and perception of chemistry topics. In the second half of the module, student teachers are in a school and are required to analyze both lessons taught by a mentor at school and by other students. The data obtained helps them to plan a longer teaching unit in chemistry. Furthermore, student teachers are supported by science educators from the university. After ChemEd4, student teachers will not attend another seminar in chemistry education.

There are three time points of data collection, which are dictated by the overall structure of the seminars and the teacher training curriculum: (i) before ChemEd2 begins; (ii) after ChemEd2 ends but before ChemEd4 begins; and (iii) after ChemEd 4 is completed. Figure 2 gives an overview of the cross-level study.



**Figure 2.** Overview of the cross-level study (stars = data collection; BT = Bachelor’s thesis; MT = Master’s thesis).

### 3.3. Instrument and Evaluation Pattern

Student teachers are asked for their background information (age, sex, number of semesters, etc.). Their linguistic and migration backgrounds are collected as well, since this has a direct bearing on linguistic skills and the possible problems pupils during the school internship may be facing.

Since studies in science education that focus on the evaluation of diagnostic competence are rare, the following research was based on open-ended questions [59]. The questionnaire begins with the task “Write a short essay about diagnosis in chemistry lessons.” Thus, the participants are not influenced by any structured questions. This allows their a priori, first-hand knowledge, beliefs, and attitudes towards diagnostics and heterogeneity to be collected, all of which are influenced by social desirability effects [60]. In addition, this question is aimed at determining the participants’ level of knowledge.

The second part of the questionnaire was developed from Jägers' definition [57]. Each question focuses on a different knowledge domain: (i) conditional knowledge (first question); (ii) technological knowledge (second question); and (iii) knowledge of change (third and fourth questions). The final domain, "knowledge of change," is evaluated by two questions. The first focuses on strategies of chemistry teaching, and the second on strategies in planning. The participants are asked to answer the following questions:

1. How can learning group heterogeneity affect education?
2. What methods would you use for diagnosis?
3. What strategies would you use in the classroom to deal with heterogeneity?
4. How (if at all) would you include heterogeneity in your lesson planning?

The data were analyzed using an evaluation pattern developed by qualitative content analysis [61] with the help of the qualitative analyzing program MAXQDA. A detailed description of the development of the evaluation pattern is given by Tolsdorf and Markic [59].

Each of Jäger's [57] knowledge domains is represented and contains several categories. Seven of these subcategories deal with competence knowledge: (i) insecurity of knowledge about diagnosis; (ii) sensitivity for diagnostics; (iii) reasons for diagnostics; (iv) awareness of a diagnostic as a process; (v) a wish for more knowledge/skill; (vi) attitude towards heterogeneity; and (vii) the importance of diagnosis.

The second domain is conditional knowledge and includes the understanding of factors that influence student behavior and experience. Three categories were formed from the data, along with several subcategories: (i) the individual influence of students (linguistic heterogeneity; migration/immigration; learning difficulties; socioeconomic background; content knowledge; physical disability; motivation); (ii) administrative and organizational influences (lack of (effective) lesson time; number of students in the classroom); and (iii) the influence of the lesson and its planning (social behavior; suboptimal support; changes in the lesson plan; multiple or different ideas).

The third domain is called technological knowledge. This is the knowledge needed for preparing and implementing a diagnosis (the pre-actional and actional phases). The relevant categories here are (i) games; (ii) intuitive action; (iii) observation; (iv) communication with and between students; (v) presentation; (vi) reflection; (vii) writing (any kind of worksheets), and (viii) testing.

Within the grouping knowledge of change, four categories could be identified, in which the teachers must deal with heterogeneity and diversity in chemistry classrooms: (i) changes during lesson planning; (ii) changes in teacher behavior; (iii) changes in the teaching materials; and (iv) changes of the overall framework. Details about the questionnaire, its development, and the whole evaluation pattern are described by Tolsdorf and Markic [59,62].

Finally, the data was evaluated with the help of the evaluation pattern. Each questionnaire was rated to the different categories and subcategories. The coding was performed by two researchers in the field of chemistry education. Interagreement as defined by Swanborn (1996) [63] was reached with  $\alpha = 0.83$ . The codes for each subcategory were then calculated. Finally, the relative frequency for each code in each category for each group of chemistry student teachers was calculated.

### 3.4. Sample

A total of 108 chemistry student teachers at different stages of their university teacher training participated in the present cross-level study. Fifty-one of them were female and forty-seven were male.

All of the teacher trainees were studying to become chemistry teachers at the secondary school level. The German educational system also requires teachers to have a second teaching subject. In addition to chemistry, our participants primarily chose either biology ( $n = 51$ ) or mathematics ( $n = 33$ ) as a second subject. Other secondary subjects included either German or a foreign language ( $n = 7$ ), geography ( $n = 6$ ), physics ( $n = 3$ ) or politics ( $n = 2$ ). Most students were under 25 years old; only seven of them were older than 30. All of the participants were native German native speakers and

had more-or-less fluent knowledge of another language. Fourteen student teachers had a migration background (Russian, Polish, or Turkish). Six of them spoke mainly German. The other eight spoke German for most everyday situations, but their second language made up about 50% of their daily language use. Table 1 gives an overview of the three groups of student teachers at the different stages of their teacher training.

**Table 1.** Overview of the information on the participants in the study.

		Group 1 Pre CemEd2	Group 2 Pre CemEd4	Group 4 Post CemEd4
Number		43	37	28
Sex	Female	19	19	13
	Male	24	18	15
Age	Under 21	12	0	0
	21–24	21	24	19
	25–29	6	11	8
	More than 30	4	2	1
Migration background		8	6	0
Visited school	Grammar school	37	35	25
	Comprehensive school	5	1	3
Second subject	Biology	22	16	13
	Mathematics	13	12	8
	Physic	3	0	0
	Geography	0	3	3
	Language	2	3	2
	Politics & History	2	1	0
Music		0	2	2

#### 4. Results

Diagnostic competence as defined by the four domains discussed above varied widely from one group of student teachers to another. We could identify a broad understanding of diagnosis. Predominantly, the largest differences occurred in knowledge of change. Competence knowledge and conditional knowledge showed the highest differences when student teachers started their Master's program. In general, differences between student teachers with a different second subject, their sex, or their knowledge of foreign language were not found in the present sample. Further data will be presented in detail below for all of the four knowledge domains.

##### 4.1. Competence Knowledge

Student teachers who had already taken both chemistry education modules (Group 3) consistently mentioned the importance of diagnostics in chemistry teaching and learning at a higher level than the other two groups. Only about 5% of student teachers in Group 1 and about 20% of Group 2 made this connection. The biggest difference between the groups was in their attitude toward heterogeneity and their perceived need of diagnostics. Group 1 did not express a need for diagnostics. Members of this group tended to describe heterogeneity as something negative, which has a negative influence on chemistry teaching and learning. This was especially true considering the work and behavior during the laboratory work. This attitude differed from one group to the next. However, only 3% of the more experienced student teachers in Group 3 still viewed heterogeneity as negative. The others in this group saw it as more of an advantage for chemistry teaching. One student wrote: *“Heterogeneity can be a positive influence on chemistry teaching in my opinion. Particularly, for example, on open inquiry experiments, creativity, ideas, knowledge and discussion. Students bring different content knowledge and learning levels into the classroom. This should not be changed negatively, but should instead be used to positive effect. The students*



*can support each other and it helps to consider different topics from different perspectives. These raise interest and increase fun in chemistry"* (PAAE5).

Student teachers also positively described the need for diagnosis in chemistry classes. In most cases, the focus was on the teaching and learning of scientific language and the influence of heterogeneity during laboratory sessions. Fortunately, even 50% of the student teachers in Group 1 already knew why diagnosis should be performed in chemistry classes. In Group 3, a total of 80% of the teachers said the same thing. The latter group also described diagnosis as a cyclical process and used an example of chemistry teaching to make their point.

#### 4.2. Conditional Knowledge

Conditional knowledge requires student teachers to be sensitive to heterogeneity and diversity. It also demands that they be constantly aware of these factors and be in a position to identify their effects on chemistry teaching. This knowledge includes all possible influences and effects on teaching. Within the three categories of this knowledge domain, differences could be detected among the groups. While the first group mainly focused on laboratory work and group work, the second and third groups went deeper into understanding chemistry and gaining content knowledge. In general, administrative and organizational influences were mentioned by all groups (ranging between 2% and 15%). All three groups of student teachers named the individual influence of pupils (e.g., students pre-knowledge and misconceptions), as the most important factor influencing chemistry teaching and learning. The most-mentioned aspects by all three groups were linguistic heterogeneity considering the German language as well as the language of chemistry and personal differences in content knowledge. However, Group 2 compared to Groups 1 and 3, had the lowest coding number. The largest inter-group difference was found to be in the participant's culture, migration background, and motivation. The longer the student teachers studied and the more time they spent in internships, the three sub-categories were more mentioned by the participants in this study (and this was presented and finally written in their answer). Dimensions of the diversity wheel such as physical requirements or socioeconomic status were mentioned by only a few student teachers in all of the three groups.

#### 4.3. Technological Knowledge

All three groups of student teachers named different methods for data collection during the analysis. But the most predominate methods listed by all of the participants were tests and questionnaires. Group 1 focused on written data collection (any kind of worksheets) and interviews with students. Classroom observation was also mentioned as a diagnostic tool by 40% of student teachers in this group. Group 2 followed the same general pattern, but not as much as the first group. It is worth noting that Group 3 mentioned interviews much less frequently. For this group, the focus was mainly on tests and questionnaires, especially considering diagnostics during group work in the laboratory.

#### 4.4. Knowledge of Change

Knowledge of change was the most prominent dimension for all groups of student teachers in this study. The differences between the groups could be identified primarily in the subcategories "changes in the teaching materials" and "change of the overall framework". Only the participants of Group 1 mentioned changes of the (school) framework as a possibility for dealing with heterogeneity and diversity in chemistry lessons. None of the other participants in the present study mentioned it. In the first group, one student teacher wrote: *"The concept of teaching in the school must be generally changed and adapted to physically handicapped students, e.g., less movement during the lesson, thinking about the amount of experiments"* (CJAN1).

It is interesting to see the differences between naming the "changes in teaching materials" between the three groups. While only about 70% of student teachers in Group 1 mention this aspect, almost 97% of participants in Group 3 listed different ideas for how teaching materials could be effectively

changed. They described concrete ways of designing materials or adding content help and linguistic aids. Their list included tools such as differentiation, cooperative learning, the use of class experts, and considering different learning styles.

Important for this category is the design of the experiments. This issue differs from one group of student teachers to another. The highest focus on this was given by Group 2. Only a few student teachers in Group 3 mentioned any changes during experimentation phases.

Finally, “changes in teacher behavior” was mentioned by all of the groups. Group 1 exclusively focuses on teacher language, while Groups 2 and 3 mentioned increased cooperation with colleagues as one method of dealing with heterogeneity and diversity in science classes. For example, one student wrote about the development of teaching: *“It is also important that his or her teaching will analyze mistakes/problems/barriers by other teachers or persons”* (KNSO9).

## 5. Discussion and Implications

One important result for our future work and the development of our university seminars was the fact that so many differences in teacher trainees’ diagnostic competence between the three groups could be identified. In the sense of a cross-level study, these differences pinpoint the explicit influence of university teacher training programs on future teachers.

Student teachers’ attitudes and beliefs in this study differed widely from one group to another. We can now say that the longer student teachers study at university, the more positive their views and attitudes towards heterogeneity become. The same holds true for their awareness of the importance of diagnostics in the classroom. This is especially noticeable when comparing Group 3 to the other, less-experienced groups. Group 3 student teachers were just finishing up their four-month internship in school (15 h per week). In this phase of the teacher training program, they were required to diagnose chemistry classes and lessons and then use the collected data to better their own teaching. It was also interesting that most of the participants visited schools that are characterized by high levels of heterogeneity with regard to their pupils’ migration backgrounds and linguistic skills. Because this group had the highest coding in all of the categories (although the smallest group sizes), one can assume that this phase of their teacher training influenced them the most.

Student teachers in Group 1 mentioned different methods of diagnosis. The most-mentioned were written work (e.g., worksheets), tests, and interviews with the students. As mentioned above, this differed from one group to another. Group 3 was mainly focused on tests as the method of data evaluation. This development of important diagnostic instruments is similar to development results during the pilot study of the questionnaire [59,62]. This tool seems to be the most consciously present among the student teachers in this study. The third group analyzed their classes during the course of their internship and had had the chance to experience and work with different diagnostic instruments. We can assume that they found other methods of assessment to be excessively time-consuming and sometimes (depending on the size of the class) even impossible for a teacher with a full workload. Furthermore, the variety of tools they selected did not match their own beliefs and attitudes about chemistry teaching and learning and about their beliefs and knowledge of diagnostics. Again we can assume that this discrepancy points to one reason why student teachers cannot (or will not) adapt different instruments to their present knowledge level. Finally, we need to think briefly about the type of cooperation occurring with the mentors at school. Student teachers and mentors work together for a total of four months (15 h per week). Mentors help inexperienced student teachers to organize diagnosis, manage their time and efforts, come to grips with a full-time job as a teacher, and become acquainted with the school system from the administrative, instead of the student, perspective. The mentors and trainees teach together in teams and eventually plan and evaluate lessons which are taught solo by the student teachers. They work together intensely for quite a long time.

Knowledge of change was the most pronounced knowledge domain during all three time periods, since it most closely affects the participants in this study. Here the coding values were highest. Student teachers in the first group mentioned more general changes like group work and differentiation.

Similarity can be seen with Group 2. Student teachers in Group 3 mentioned more concrete ideas for changing their materials and teaching in order to deal with heterogeneity (e.g., tools for adapting worksheets or experiments). Different methods and tools for dealing with heterogeneity were a part of both of the ChemEd modules. However, the results of this study seem to indicate that the methods and tools learned became more noticeably present after the internship phase. We can assume that this topic becomes more important for student teachers after they have to use it in their own classes. Furthermore, we need to examine our seminars. Perhaps student teachers were not sensitive or experienced enough to see the need of internalizing the different methods and tools. This is the point that should be evaluated more deeply in further research. Additionally, this change was not noticeable after the first short internship within the framework of ChemEd2. One reason could be the length of the internship, coupled with the gathering of first work experience by the students. Therefore, student teachers needed to be mentored and supported by the university during this short time, so that they can better structure the knowledge gained during the seminars. Additionally, Morrison and Ledermann (2003) [13] have identified a lack of knowledge or insufficient knowledge of diagnostic instruments among teacher trainees. Capizzi and Fuchs (2005) [55] also showed that knowledge about diagnostic instruments does not automatically lead to positive changes in teaching. Therefore, the knowledge of change must relate to diagnostic instruments, as was done in the present chemistry education courses. The study shows similarities in the student teachers' diagnostic competence and in the content of the university courses. But differences could also be identified. The practical experience provided by the internships and the mentors seems to have influenced student teachers' diagnostic competence on a high level. These effects were not examined directly in this study. It was apparent from the questionnaire that student teachers answered in more concrete and differentiated ways during the later teacher training program.

The idea of a process diagnostic develops its potential only in adaptive teaching environments [52]. In 1986, Corno and Snow [64] wrote that the success of education depends upon adaptation of the teaching to the individual pupils. Teaching is more successful if teachers can assess their students. Accordingly, the teacher can then structure teaching towards the learners. Adaptive teaching cannot be taught in detail at university courses, but initial ideas can be presented to the student teachers. Though the courses show success, for our future development of the courses, few methods in this direction are planned.

Using the results from above, we should ask the question of how much a school internship, as compared to theoretical seminars, actually influences future chemistry teachers. This study shows that both have an effect on future educators. However, it seems that the internship more heavily influences the participants. The topic for further research must be to evaluate more closely the influence of the different phases of teacher training on teachers. This should especially compare theoretical and practical phases. In order to approach these questions step by step and to recognize any changes in detail, a qualitative longitudinal interview study needs to be conducted.

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

**Author Contributions:** Yannik Tolsdorf and Silvija Markic have contributed equally to the study in the context of the dissertation project of Yannik Tolsdorf. Yannik Tolsdorf wrote the first text and Silvija Markic revised this text.

**Conflicts of Interest:** The authors declare no conflict of interest. The German Telekom Foundation had no role in the design of the study (data collection, analyses, interpretation of data), in the writing of the manuscript, or in the decision to publish the results.

## References

1. Ohle, A.; McElvany, N. Teachers' diagnostic competences and their practical relevance, special issue editorial. *J. Educ. Res. Online* **2015**, *7*, 5–10.

2. Klug, J.; Bruder, S.; Kelava, A.; Spiel, C.; Schmitz, B. Diagnostic competence of teachers: A process model that accounts for diagnosis learning behaviour tested by means of a case scenario. *Teach. Teach. Educ.* **2013**, *30*, 28–46. [CrossRef]
3. Vogt, F.; Rogalla, M. Developing adaptive teaching competency through coaching. *Teach. Teach. Educ.* **2009**, *25*, 1051–1060. [CrossRef]
4. Harrison, A.G.; Treagust, D.F. The Particulate Nature of Matter: Challenges in Understanding the Submicroscopic World. In *Chemical Education: Towards Research-Based Practice*; Gilbert, J.K., Jong, O.D., Justi, R., Treagust, S.F., van Driel, J., Eds.; Kluwer: Dordrecht, The Netherlands, 2002; pp. 189–212, ISBN 1402011121.
5. Taber, K. Developing Teachers as Learning Doctors. *Teach. Dev.* **2005**, *9*, 219–235. [CrossRef]
6. Brookhart, S.M. Educational assessment knowledge and skills for teachers. *Educ. Meas. Issues Pract.* **2011**, *30*, 3–12. [CrossRef]
7. Treagust, D.F. Development and use of diagnostic test to evaluate students' misconceptions in science. *Int. J. Sci. Educ.* **1988**, *10*, 159–169. [CrossRef]
8. Barke, H.-D.; Hazari, A.; Yitbarek, S. *Misconceptions in Chemistry: Addressing Perceptions in Chemical Education*; Springer: Berlin, Germany, 2009; ISBN 3540709886.
9. Taber, K.S. *Chemical Misconceptions—Prevention, Diagnosis and Cure: Theoretical Background*; RSC: London, UK, 2002; Volume 1, ISBN 0854043861.
10. Johns Hopkins University Diversity Leadership Council. Diversity Wheel. Available online: [https://www.web.jhu.edu/dlc/resources/diversity\\_wheel/index.html](https://www.web.jhu.edu/dlc/resources/diversity_wheel/index.html) (accessed on 22 August 2017).
11. Tolsdorf, Y.; Markic, S. Dealing language in science classroom—Diagnosing student' linguistic skills. In *Science Education towards Inclusion*; Markic, S., Abels, S., Eds.; Nova: New York, NY, USA, 2016.
12. Florian, L.; Black-Hawkins, K. Exploring inclusive pedagogy. *Br. Educ. Res. J.* **2011**, *37*, 813–828. [CrossRef]
13. Morrison, J.A.; Lederman, N.G. Science Teachers' Diagnosis and Understanding of Students' Preconceptions. *Sci. Educ.* **2003**, *87*, 849–867. [CrossRef]
14. Loughran, J.; Mulhall, P.; Berry, A. Exploring pedagogical content knowledge in science teacher education. *Int. J. Sci. Educ.* **2008**, *30*, 1301–1320. [CrossRef]
15. Pellegrino, J.W.; Chudowsky, N.; Glaser, R. *Knowing What Students Know: The Science and Design of Educational Assessment*; National Academic Press: Washington, DC, USA, 2001; ISBN 0309072727.
16. Ruiz-Primo, M.A.; Li, M.; Wills, K.; Giamellaro, M.; Lan, M.C.; Mason, H.; Sands, D. Developing and evaluating instructionally sensitive assessments in science. *J. Res. Sci. Teach.* **2012**, *49*, 691–712. [CrossRef]
17. Shepard, L.A. The role of classroom assessment in teaching and learning. In *Handbook of Research on Teaching*, 4th ed.; Richardson, V., Ed.; American Educational Research Association: Washington, DC, USA, 2001; pp. 1066–1101, ISBN 0935302263.
18. Black, P. Formative and summative aspects of assessment: Theoretical and research foundations in the context of pedagogy. In *Sage Handbook of Research on Classroom Assessment*; McMillan, J.H., Ed.; Sage: Thousand Oaks, CA, USA, 2013; pp. 167–178, ISBN 1412995876.
19. Ingenkamp, K.; Lissmann, U. *Lehrbuch der Pädagogischen Diagnostik*, 6th ed.; Beltz: Weinheim, Germany, 2008; ISBN 3407255039.
20. Fächter, A. Pädagogische und didaktische Diagnostik—Eine schulische Entwicklungsaufgabe mit hohem Professionalitätsanspruch. In *Diagnostik und Förderung, Teil I: Didaktische Grundlagen*; Fächter, A., Moegling, K., Eds.; Prolog: Immenhausen, Germany, 2011; pp. 45–83, ISBN 3934575560.
21. Bell, B. Classroom assessment of science learning. In *Handbook of Research on Science Education*; Abell, S.K., Lederman, N.G., Eds.; Lawrence Erlbaum: Mahwah, NJ, USA, 2007; pp. 965–1006, ISBN 0805847146.
22. Schrader, F.-W. Diagnostische Kompetenz von Lehrpersonen. *Beiträge zur Lehrerinnen- und Lehrerbildung* **2013**, *31*, 154–165.
23. Bennett, R.E. Formative assessment: A critical review. *Assess. Educ.* **2011**, *18*, 5–25. [CrossRef]
24. Nitko, A.J.; Brookhart, S.M. *Educational Assessment of Students*, 5th ed.; Prentice-Hall: Upper Saddle River, NJ, USA, 2007; ISBN 0131719254.
25. Shwartz, Y.; Dori, Y.J.; Treagust, D.F. How to outline objectives for chemistry education and how to assess them. In *Teaching Chemistry—A Studybook*; Eilks, I., Hofstein, A., Eds.; Sense: Rotterdam, The Netherlands, 2013; pp. 37–65, ISBN 9462091382.

26. The Organisation for Economic Co-Operation and Development. *Formative Assessment: Improving Learning in Secondary Classrooms*; OECD Publishing: Paris, France, 2005; Available online: <https://www.oecd.org/edu/ceri/35661078.pdf> (accessed on 22 August 2017).
27. Black, P.; Wiliam, D. Developing the theory of formative assessment. *Educ. Assess. Eval. Account.* **2009**, *21*, 5–31. [\[CrossRef\]](#)
28. Atkin, J.M.; Black, P.; Coffey, J. *Classroom Assessment and the National Science Education Standards*; National Academy Press: Washington, DC, USA, 2001; ISBN 030906998X.
29. Hattie, J.; Yates, G. *Visible Learning and the Science of How We Learn*; Routledge: London, UK, 2014; ISBN 0415704995.
30. Heidemeier, H. Self and Supervisor Ratings of Job-Performance: Meta-Analyses and a Process Model of Rater Convergence. Ph.D. Dissertation, Friedrich-Alexander-Universität, Erlangen-Nürnberg, Germany, 10 May 2005. Available online: <https://opus4.kobv.de/opus4-fau/frontdoor/index/index/docId/143> (accessed on 22 August 2018).
31. Klug, J. Modeling and Training a New Concept of Teachers' Diagnostic Competence. Ph.D. Dissertation, Technische Universität Darmstadt, Darmstadt, Germany, 25 August 2011. Available online: [http://tuprints.ulb.tu-darmstadt.de/2838/1/16.01.2012\\_Dissertation\\_Julia\\_Klug.pdf](http://tuprints.ulb.tu-darmstadt.de/2838/1/16.01.2012_Dissertation_Julia_Klug.pdf) (accessed on 22 August 2018).
32. Von Aufschnaiter, C.; Cappell, J.; Dübbelde, G.; Ennemoser, M.; Mayer, J.; Stiensmeier-Pelster, J. Diagnostische Kompetenz: Theoretische Überlegungen zu einem zentralen Konstrukt der Lehrerbildung. *Z. Pädagogik* **2015**, *5*, 738–758. [\[CrossRef\]](#)
33. Tan, K.-C.D.; Taber, K.S.; Goh, N.-K.; Chia, L.-S. The ionisation energy instrument: A two-tier multiple-choice instrument to determine high school students' understanding of ionisation energy. *Chem. Educ. Res. Pract.* **2005**, *6*, 180–197. [\[CrossRef\]](#)
34. Peterson, R.F.; Treagust, D.F.; Garnett, P.J. Development and Application of a Diagnostic Instrument to Evaluate Grade 11 and 12 Students' Concepts of Covalent Bonding and Structure following a Course of Instruction. *J. Res. Sci. Teach.* **1989**, *26*, 301–314. [\[CrossRef\]](#)
35. Tamir, P. Assessment and evaluation in science education—Opportunities to learn and outcomes. In *International Handbook of Science Education*; Fraser, B.J., Tobin, K.G., Eds.; Kluwer: Dordrecht, The Netherlands, 1998; pp. 762–789, ISBN 9780792335313.
36. Shulman, L.S. Those Who Understand: Knowledge Growth in Teaching. *Educ. Res.* **1986**, *15*, 4–14. [\[CrossRef\]](#)
37. Shulman, L.S. Knowledge and Teaching: Foundations of the New Reform. *Harv. Educ. Rev.* **1987**, *57*, 1–22. [\[CrossRef\]](#)
38. Abell, S.K. Research on science teacher knowledge. In *Handbook of Research on Science Education*; Abell, S.K., Lederman, N.G., Eds.; Lawrence Erlbaum Associates: Mahwah, NJ, USA, 2007; pp. 1105–1149, ISBN 0805847146.
39. Abell, S.K. Twenty years later: Does pedagogical content knowledge remain a useful idea? *Int. J. Sci. Educ.* **2008**, *30*, 1405–1416. [\[CrossRef\]](#)
40. Park, S.; Oliver, J.S. Revisiting the Conceptualization of Pedagogical Content Knowledge (PCK): PCK as a Conceptual Tool to Understand Teachers as Professionals. *Res. Sci. Educ.* **2008**, *38*, 261–284. [\[CrossRef\]](#)
41. Loughran, J.; Berry, A.; Mulhall, P. *Professional Learning: Understanding and Developing Science Teachers' Pedagogical Content Knowledge*; Sense: Rotterdam, The Netherlands, 2006; ISBN 9789087903657.
42. Loughran, J.; Berry, A.; Mulhall, P. *Understanding and Developing Science Teachers' Pedagogical Content Knowledge*, 2nd ed.; Sense: Rotterdam, The Netherlands, 2012; ISBN 9789460917882.
43. Black, P.; William, D. Assessment and classroom learning. *Assess. Educ.* **1998**, *5*, 7–74. [\[CrossRef\]](#)
44. Feinberg, A.B.; Shapiro, E.S. Teacher Accuracy: An Examination of Teacher-Based Judgments of Students' Reading with Differing Achievement Levels. *J. Educ. Res.* **2009**, *102*, 453–462. [\[CrossRef\]](#)
45. Perry, N.E.; Hutchinson, L.; Thauberger, C. Talking about teaching self-regulated learning: Scaffolding student teachers' development and use of practices that promote self-regulated learning. *Int. J. Educ. Res.* **2008**, *47*, 97–108. [\[CrossRef\]](#)
46. Coladarci, T. Accuracy of teacher judgment of students' responses to standardized test items. *J. Educ. Psychol.* **1986**, *78*, 141–146. [\[CrossRef\]](#)
47. Hoge, R.D.; Coladarci, T. Teacher-based judgments of academic achievement: A review of literature. *Rev. Educ. Res.* **1989**, *59*, 297–313. [\[CrossRef\]](#)
48. Partenio, I.; Taylor, R.L. The relationship of teacher ratings and IQ: A question of bias? *Sch. Psychol. Rev.* **1985**, *14*, 79–83.



49. Demaray, M.K.; Elliott, S.N. Teachers' judgements of students' academic functioning: A comparison of actual and predicted performances. *Sch. Psychol. Q.* **1998**, *13*, 8–24. [[CrossRef](#)]
50. Feinberg, A.B.; Shapiro, E.S. Accuracy of teacher judgments in predicting oral reading fluency. *Sch. Psychol. Q.* **2003**, *18*, 52–65. [[CrossRef](#)]
51. Karing, C.; Matthäi, J.; Artelt, C. Genauigkeit von Lehrerurteilen über die Lesekompetenz ihrer Schülerinnen und Schüler in der Sekundarstufe I—Eine Frage der Spezifität? *Z. Pädagogische Psychol.* **2011**, *25*, 159–172. [[CrossRef](#)]
52. Südkamp, A.; Kaiser, J.; Möller, J. Accuracy of teachers' judgements of students' academic achievement: A Meta-analysis. *J. Educ. Psychol.* **2012**, *104*, 743–762. [[CrossRef](#)]
53. Begeny, J.C.; Eckert, T.L.; Montarello, S.A.; Storie, M.S. Teachers' perceptions of students' reading abilities: An examination of the relationship between teachers' judgments and students' performance across a continuum of rating methods. *Sch. Psychol. Q.* **2008**, *23*, 43–55. [[CrossRef](#)]
54. Bates, C.; Nettelbeck, T. Primary school teachers' judgments of reading achievement. *Educ. Psychol.* **2001**, *21*, 177–187. [[CrossRef](#)]
55. Capizzi, A.M.; Fuchs, L. Effects of curriculum-based measurement with and without diagnostic feedback on teacher planning. *Remedial Spec. Educ.* **2005**, *26*, 159–174. [[CrossRef](#)]
56. Krauss, S.; Brunner, M.; Kunter, M.; Baumert, J.; Blum, W.; Neubrand, M.; Jordan, A. Pedagogical content knowledge and content knowledge of secondary mathematics teachers. *J. Educ. Psychol.* **2008**, *100*, 716–725. [[CrossRef](#)]
57. Jäger, R.S. Diagnostischer Prozess. In *Handbuch der Psychologischen Diagnostik*; Petermann, F., Eid, M., Eds.; Hogrefe: Göttingen, Germany, 2006; pp. 89–96, ISBN 9783801719111.
58. Ohle, A.; McElvany, N.; Horz, H.; Ullrich, M. Text-picture integration—Teachers' attitudes, motivation and self-related cognitions in diagnostics. *J. Educ. Res. Online* **2015**, *7*, 11–33.
59. Tolsdorf, Y.; Markic, S. Development of an instrument and evaluation pattern for the analysis of chemistry student teachers' diagnostic competence. *Eur. J. Phys. Chem. Educ.* **2017**, accepted.
60. Weisberg, H.E. *The Total Survey Error Approach: A Guide to the New Science of Survey Research*; The University of Chicago Press: Chicago, IL, USA, 2005; ISBN 9780226891286.
61. Mayring, P. *Qualitative Content Analysis: Theoretical Foundation, Basis Procedures and Software Solutions*; Beltz: Klagensfurt, Germany, 2014.
62. Tolsdorf, Y.; Markic, S. Exploring student teachers' knowledge concerning diagnostics in science lessons. In *Science Education Research: Engaging Learners for a Sustainable Future, Proceedings of the ESERA 2015 Conference, Helsinki, Finland, 31 August–4 September 2015*; Lavonen, J., Juuti, K., Lampiselkä, J., Uitto, A., Hahl, K., Eds.; University of Helsinki: Helsinki, Finland, 2016.
63. Swanborn, P.G. A common base for quality control criteria in quantitative and qualitative research. *Qual. Quant.* **1996**, *30*, 19–35. [[CrossRef](#)]
64. Corno, L.; Snow, R.E. Adapting teaching to individual differences among learners. In *Third Handbook of Research on Teaching*; Wittrock, M.C., Ed.; Macmillan: New York, NY, USA, 1986; pp. 605–629, ISBN 0029003105.



© 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).