

Article

Professional Knowledge and Self-Efficacy Expectations of Pre-Service Teachers Regarding Scientific Reasoning and Diagnostics

Dagmar Hilfert-Rüppell ^{1,*} , Monique Meier ², Daniel Horn ²  and Kerstin Höner ¹

¹ Institut for Science Education, Technische Universität Braunschweig, 38106 Braunschweig, Germany; k.hoener@tu-braunschweig.de

² Biology Education, Universität Kassel, 34132 Kassel, Germany; mmeier@uni-kassel.de (M.M.); daniel.horn@uni-kassel.de (D.H.)

* Correspondence: d.hilfert-ruempel@tu-braunschweig.de

Abstract: Understanding and knowledge of scientific reasoning skills is a key ability of pre-service teachers. In a written survey (open response format), biology and chemistry pre-service teachers ($n = 51$) from two German universities claimed central decisions or actions school students have to perform in scientific reasoning in the open inquiry instruction of an experiment. The participants' answers were assessed in a quality content analysis using a rubric system generated from a theoretical background. Instruments in a closed response format were used to measure attitudes towards the importance of diagnostics in teacher training and the domain-specific expectations of self-efficacy. The pre-service teacher lacked pedagogical (didactics) content knowledge about potential student difficulties and also exhibited a low level of content methodological (procedural) knowledge. There was no correlation between the knowledge of student difficulties and the approach to experimenting with expectations of self-efficacy for diagnosing student abilities regarding scientific reasoning. Self-efficacy expectations concerning their own abilities to successfully cope with general and experimental diagnostic activities were significantly lower than the attitude towards the importance of diagnostics in teacher training. The results are discussed with regard to practical implications as they imply that scientific reasoning should be promoted in university courses, emphasising the importance of understanding the science-specific procedures (knowing how) and epistemic constructs in scientific reasoning (knowing why).

Keywords: professional knowledge; scientific reasoning skills; self-efficacy; students' difficulties; diagnostic competencies



Citation: Hilfert-Rüppell, D.; Meier, M.; Horn, D.; Höner, K. Professional Knowledge and Self-Efficacy Expectations of Pre-Service Teachers Regarding Scientific Reasoning and Diagnostics. *Educ. Sci.* **2021**, *11*, 629. <https://doi.org/10.3390/educsci11100629>

Academic Editors: Moritz Krell, Andreas Vorholzer and Andreas Nehring

Received: 8 August 2021

Accepted: 6 October 2021

Published: 11 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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 (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Inquiry-based teaching is seen as contributing to content, procedural, and epistemic learning goals of science education [1]. Therefore, a basic understanding of the systematic approach to conducting science investigations is required in the competencies of scientific inquiry (e.g., in the US [2] and in the UK [3]). In Germany, these are reported for biology, chemistry, and physics with similarly formulated educational standards [4–6], allowing these competencies to be promoted in a networked manner both vertically within a subject and horizontally across subjects [7]. This applies also to the field of school education and in the training of pre-service teachers. Both areas are in turn directly related, i.e., the knowledge and skills (for scientific inquiry) of the teacher shape the teaching with learning opportunities (for inquiry-based science education) and thus the potential learning success on the part of the students [8]. As a result, at the international and national levels, standards for the teaching profession are also being formulated and science teaching competencies described for teacher education [9–11]. Common to them are the requirements for future teachers to build up science education about and through scientific inquiry as their own

competence, as well as to learn how to teach scientific inquiry in order to be able to promote corresponding competencies in students. For example, (ongoing) teachers have to be competent themselves in designing empirical approaches to test hypotheses, and need knowledge as well as skills in hypothesis-led experimentation [12]. Studies on the development of pre-service science teachers' scientific reasoning competencies show that explicit reflections about scientific reasoning (i.e., learning about science; [13]) contributes more to the development of scientific reasoning competencies than only doing science without reflecting about it [14,15]. In the course of this, knowledge about and understanding of scientific inquiry and scientific reasoning is relevant [16], and can also be used as constructs for the analysis and assessment of learning activities and learning outcomes [17].

1.1. Scientific Reasoning

In many studies, the terms scientific reasoning and scientific thinking are used interchangeably because the boundary between reasoning and thinking is blurred (e.g., [18–20]). Scientific reasoning thus forms the basis for critical thinking and is only one, albeit very significant, aspect in the process of thinking about (scientific) facts (e.g., [21–24]). Scientific reasoning can therefore be interpreted as a subset of critical thinking skills (cognitive and metacognitive processes and dispositions) that are essential for scientific procedures in the problem-solving process, the evidence of information in scientific disciplines, and the epistemological incorporation of scientific methods and paradigms [25]. There are a range of views on the structure of scientific reasoning and on the number of its components. Two groupings can be distinguished: while one emphasises scientific reasoning as a broad and complex component representing a particular skill, understanding, or competence [26], the other grouping opposes the advocacy of multidimensional theories [27].

Since the 1970s, according to Dunbar and Fugelsang [28], scientific reasoning has also been viewed as a way of solving problems, with great efforts being made to identify strategies that scientists use to solve problems. Competencies required for scientific reasoning are viewed as a complex construct, encompassing both the skills required for scientific problem solving and the ability to reflect on problem solving at a meta-level [29,30]. The scientific discovery process is best conceptualised as involving both reasoning and problem-solving skills, with the ultimate goal of generating, testing, and then evaluating a hypothesis about a causal or categorical relationship based on the results. Both of these skills—strategy development and reasoning processes—require knowledge to identify key features of a problem at hand [31]. Problem solving requires three types of knowledge in a complex construct (knowing that, knowing how, knowing why, cf., [29,32,33]). Explicit procedural knowledge, in turn, addresses the execution level and thus the “knowing how” and practical implementation of actions to solve problems [32,33]. In his structural model of scientific reasoning, Mayer [34] identified, in addition to personal variables, such as (prior) knowledge and cognition, four process-related skills/subcompetencies:

1. Formulating scientific questions;
2. Generating hypotheses;
3. Planning scientific investigations;
4. Interpreting data, which are differentiated via different aspects of competence, e.g., [35].

In the execution of these process steps in problem-oriented and inquiry-based teaching, content knowledge is to be applied and methodological knowledge is to be developed and applied in equal measure [19]. Experimentation is considered to play a central role in the process of scientific inquiry as a content and method in science education [36]. “Scientific reasoning is defined as the inquiry processes [. . .] [and] the reasoning skills involved in experimentation, evidence evaluation, and inference making addressed to scientific understanding” [37] (p. 106). In the present study, scientific reasoning is predominantly defined from a procedural perspective and thus addressed in the most open-ended problem-solving process possible using the method of experimentation.

In order to perform diagnostics and scaffolding in the (experimental) inquiry process, knowledge about scientific reasoning, as well as about the implementation and associated

difficulties on the part of the learners of scientific reasoning, plays a central role for (pre-service) teachers. Given this significance, the key difficulties faced by students will be outlined below in a summary of the literature.

1.2. Literature Summary on Student Difficulties in the Experimental Problem Solving Process

Experimentation demands and promotes a wide range of cognitive, psychomotor, and social skills in students [35]. In this context, the SDDS model (Scientific Discovery as Dual Search; [22]) as well as the structural model for scientific reasoning [34] with their anchoring in problem-solving research is relevant for both the conceptualisation and the assessment of competencies in the science subjects of biology [37–41], chemistry [42–44], and physics [45–47], respectively, and is central in natural science studies [7,48]. In promoting the competencies underlying each model, the inquiry-based learning approach is shown to be superior to direct instruction on these [49]. In inquiry-based learning, the degree of student activity or, respectively, the open-endedness in the experimentation can be designed differently (levels of inquiry [50]) and thus influence learning success. In the literature, on the one hand, the high value of independence in experimentation compared to teacher demonstration experiments is emphasised (e.g., [51,52]) and, on the other hand, the guided inquiry approach with targeted support of learners in the phases for experimental problem solving is attributed the highest effectiveness [53]. Vorholzer and von Aufschnaiter [54] identify three main dimensions in which the implementation of guidance can vary: (a) the degree of autonomy, (b) the degree of conceptual information, and (c) the cognitive domain of guidance. Independent experimentation can be understood as a relatively complex cognitive problem-solving process that is particularly challenging for students (cf., [55,56]) and consequently requires scaffolding in the different dimensions, process steps, and/or competencies. Accordingly, pre-service science teachers must be provided with learning opportunities to acquire these competencies, to foster their own reasoning in science alongside how to teach students how to reason. They should also be enabled to plan and implement targeted lessons that enable their students to acquire the scientific reasoning competencies for experimentation and experimental problem solving required by the educational standards of many countries [4–6,57]. Neither mere boiler-plate imitation of experiments nor participation in teacher demonstration experiments by students leads to the desired mastery of the scientific reasoning process [58]; nor are approaches opened too early effective [59,60]. The latter can create numerous difficulties/barriers for students to overcome during experimentation. According to the structural areas/phases of hypothesis, planning/execution, and testing, as well as conclusions and evaluation of results in the models of Klahr [22] and Mayer [34], the misconceptions and student difficulties described here descriptively and proven empirically are summarised in Table 1.

Table 1. Overview of process-related misconceptions and difficulties of students in experimentation published in the literature (references in brackets on the right side) structured and summarised according to the three phases for scientific reasoning of the SDDS model [22].

Phase:	Search Hypothesis	[39,61–67]
	No hypothesis is formulated.	
	No alternative hypotheses are formulated.	
	There is no idea about the purpose of hypothesising.	
	Hypotheses are not related to the research question.	
	Hypotheses are formulated and/or changed in the scientific inquiry process.	
	In the hypothesis, the variables are not defined or are defined incorrectly.	
	Hypotheses are formulated without justification (conjectures).	
	Assumptions are justified by themselves.	
	Assumptions are justified with reference to everyday life.	

Table 1. Cont.

Phase:	Planning and Testing	[39,56,63,66–76]
	Experiment is not suitable for testing the hypothesis. No plan and/or unstructured trial and error (no plan, change all). Missing and/or incorrect operationalisation of variables. Unsystematic handling of variables, i.e., several variables are confounded with each other or the same variable is varied unsystematically. Planning of an experiment that does not lead to the desired results. Lack of a measurement concept/measurement repetitions. Lack of a control approach and/or attention to control variables. Selection of materials is unsystematic, incomplete and/or done by trial and error. Difficulties in handling (simple) materials (e.g., pipette). Wrong observation focus. Interference factors are perceived but not eliminated.	
Phase:	Evaluate Evidence	[63,66,67,72,75,77–80]
	Hypothesis is confirmed without considering the results (confirmatory bias). Hypothesis is adjusted to the results. Data are not (re)related to hypothesis. Not all results are considered; (unexpected) results are (partly) ignored. Unexpected results are attributed to errors in the experimental procedure. Missing or incomplete (error) reflection of the results. Wrong conclusion from coherent experiments. Due to unsystematic variable variation, no conclusions can actually be drawn, or illogical conclusions are drawn.	

1.2.1. Formulating/Search Hypothesis

When generating hypotheses, students often find it difficult to make educated guesses. They do not consider preconditions such as justification or verifiability—in principle or with the given experimental materials. The procedure that several hypotheses are to be set up is mostly unknown to them or only the one they think is correct is considered (e.g., [64,81]).

1.2.2. Design and Execution of the Experiment

It is often observed that the students do not plan out their approach to the experiment, but instead carry it out immediately and instinctively with no pre-determined method in mind. This can result in the steps in the experiment not being purposeful and the procedure being changed several times [72]. Hammann and colleagues [40] describe this unstructured trial-and-error approach as “no plan, change all”. The most important scientific reasoning skill describes the control of variables strategy [70]. The difficulties of students here lie in identifying the dependent and independent variable ([82,83], among others), and in the fact that the variable control is often not considered and confounding variables are not excluded [63]. Moreover, a control approach is usually missing, and measurement repetitions are rarely performed [39,69,82,83]. Sometimes students try to create an effect rather than conduct a goal-directed experiment [75,78]. Inefficient experimentation is also evident, e.g., the same experiment is repeated multiple times [81]. Kraeva [44] was able to identify six different approaches of students in conducting chemistry experiments through video analysis, classified in terms of the attributes “plan” versus “try” as well as “maintain”, “develop”, and “discard”. “Revision” as a planning-reviewing strategy was described as relatively successful, whereas ‘imitation’ (exploratory-alternative-less) was described as a relatively less successful strategy. This corresponds with the courses of action taken by students when experimenting to clarify a biological phenomenon in the sense of a process-oriented or explorative type [39].

1.2.3. Evaluation of the Evidence

The generated data are partly disregarded, while conclusions are drawn illogically and for the most part not related back to the hypothesis [67]. Hammann ([84], p. 200) refers to the hasty termination of the search for possible hypotheses and thus to the wrong conclusion as “positive capture”. He calls the “most robust finding in the literature” the “confirmation bias” [62,79], i.e., the tendency to confirm hypotheses while ignoring contradictory data [84]. Dunbar [85] distinguishes between two approaches to data evaluation: the “find-evidence-goal” approach, i.e., looking for results that confirm the hypothesis, and the “find-hypothesis-goal” approach, i.e., looking for new hypotheses after non-confirming results that then reflect the results. Evaluation of the evidence in the form of referring back to the hypothesis rarely occurs, as does discussion of error [39].

The probability of success for a specific experimentation process depends on the one hand, on the requirements resulting from the individual phases of the experimentation process [86], and, on the other hand, on personal characteristics, such as interest in the subject or general cognitive performance of the students [55]. For diagnosis, in this case the assessment of students’ abilities and performance in experimental problem solving, the discrepancy between students’ conceptions and students’ actions on scientific concepts for experimentation is used [87,88]. Draude [89] was able to demonstrate deficits regarding the diagnosis of student difficulties in experimentation for physics teachers, whereby the necessary prerequisites are hardly developed and promoted in the teacher training programs [75,90]. In a longitudinal two-year study with biology pre-service teachers, insights into the structure and change of their diagnostic competence and possible influencing factors were obtained [91]. Diagnoses of student difficulties and performance in experimental problem solving can only succeed if pre-service teachers themselves have the appropriate (professional) knowledge of the subject matter to be diagnosed. Otherwise, diagnostic processes will be impaired by the existing knowledge gaps [92]. Before a promotion in this area can and should be targeted and discussed as a training element in teacher education, we pursue with this study the concern to describe pre-service science teachers’ scientific reasoning competencies in order to derive the relevance of possible curricular implementations in subject, subject didactics, and educational science for diagnostics in experimentation. As described, the latter is of importance in all study elements, but is located differently and requires a theory-related consideration in the following.

1.3. Relevance of Diagnostic Competencies for (Pre-Service) Teachers

Diagnostic competence of teachers describes both the ability to successfully cope with the diagnostic tasks arising in the teaching profession and the quality of the diagnostic performance [93]. Making efficient instructionally relevant decisions is impossible without being able to identify, understand, and even predict instructionally relevant situations and events [94]. Thus, the investment of a (subject-related) diagnostic competence of (pre-service) teachers seems to be an indispensable prerequisite for the teaching profession [95]. In Shulman’s three main domains of professional knowledge (content knowledge (CK), pedagogical content knowledge (PCK), pedagogical knowledge (PK) [96], which is the most widely used classification in the literature, knowledge about assessment and diagnosis is classified in the domain of general didactical knowledge (pedagogical knowledge, PK) [97]. Kramer et al. [95] describe either PK (more generic: e.g., teaching disorder [98]) or the subject-specific facets CK or PCK (e.g., diagnosing biology instruction, [99]) as relevant to the application of diagnostic activities and diagnostic accuracy, depending on the diagnostic focus. Results of path analyses utilising Rasch measures showed that both PCK and PK were statistically significantly in relation to pre-service teachers’ diagnostic activities. Additionally, biology teachers’ PCK was positively related to diagnostic accuracy [95].

Divergent assumptions exist about what comprises teachers’ diagnostic competence, stemming from the fact that different aspects such as subject matter, method, and target are modeled. The conceptualisation of diagnostic activities in which knowledge is applied in order to solve specific problems can be seen as equivalent to scientific reasoning

skills [100]. Crucial for a sustainable diagnostic cycle is the transformation of competence into performance mediated by situational skills of perception (P), interpretation (I), and (action) decision (D) in the sense of Blömeke, Gustafsson, and Shavelson's [101] model, in which teachers' competence is viewed as a continuum with multiple transitions (P-I-D model of competence transformation). In this context, diagnostic competence presupposes the correct perception of relevant classroom features (noticing) and their evaluation with reference to theoretically grounded, pedagogical action knowledge (reasoning) [102]. Draude [89] distinguishes between predictive and action diagnostic competence of physics teachers. While the predictive competence measured the extent to which teachers could predict students' experimental difficulties in a particular physics experiment, the action-accompanying diagnostic competence measured the extent to which teachers diagnosed difficulties during the students' experimentation process. He also found deficits in both areas, so that a promotion of (pre-service) teachers' skills in this regard seems to be indicated [89]. In the present study, therefore, the predictive diagnostic competence of biology and chemistry pre-service teachers with regard to student difficulties, misconceptions, and the necessary central decisions students have to make during open-ended experimentation was assessed by means of a text-based description of a teaching scenario for a student experiment. In terms of examining diagnostic competence, self-report is common in research, so there is a need for tools to survey diagnostic and reflective skills in a natural setting [103,104].

1.4. Self-Efficacy Expectations

Self-efficacy expectations are considered another major aspect of teachers' professional competence; following Baumert and Kunter's [105] model of professional competence, self-efficacy expectations are considered relevant in addition to knowledge and attitudes. Self-assessments related to motivation, personal engagement, or self-efficacy also appear to be of value in better understanding the interplay between motivational and affective states and diagnostic activities. Self-efficacy, first introduced by Bandura [106] as an aspect of social cognitive learning theory, is described as the strength of one's belief in one's ability to perform a particular task or achieve a particular outcome. Thus, assessing self-efficacy is less about what skills and abilities individuals possess and more about what they believe they can do with the skills and abilities they possess [107]. In this regard, competent performance is guided in part by higher-level self-regulatory abilities [108]. These include general abilities to diagnose task demands, construct and evaluate alternative courses of action, set perspective-close goals to guide one's efforts, and create self-incentives to maintain engagement in stressful activities and manage stress and distracting thoughts [109]. Self-efficacy correlates with academic performance [110,111], task persistence, motivation [112], and resilience in academic contexts [113]. Self-efficacy varies depending on the situation and therefore needs to be considered or captured in a domain- and context-specific manner [114]. Students' self-efficacy in science education has been studied in the science subjects of mathematics [115–118], physics [117], and chemistry [118,119]. With regard to problem solving in mathematics, it was shown that even when students have the ability to solve problems, those who have a strong self-efficacy expectancy are more effective problem solvers [116,117].

Regarding the self-efficacy expectancy of (pre-service) teachers in science, there are some less empirical findings [120–122]. In a study by Yürük [120], pre-service teachers who had taken more science courses in college, felt better prepared to teach science content and had higher levels of self-efficacy. Riese and Reinhold [123] addressed the relationship between physics teachers' CK and PCK (compare [96], see above) and their general and classroom self-efficacy. They found a significant positive correlation between teaching-related self-efficacy and CK. Kurbanoglu and Akim [121], meanwhile, show that low self-efficacy expectancy regarding the subject of chemistry predicts chemistry laboratory anxiety and has a negative effect on freshmen's attitudes toward chemistry. In biology, there have been very few studies on self-efficacy (compare [114,122,124]). The findings of

Mahler, Großschedl, and Harms [124] indicate that teacher education in college, attending professional development courses, and self-study provide learning opportunities to promote self-efficacy and enthusiasm for teaching. In addition, the authors found that self-efficacy and subject-specific enthusiasm were positively related to PCK.

While self-efficacy or self-efficacy expectations are increasingly a focus of inquiry in teacher education, very few studies can be identified that address the application of self-efficacy theory to diagnostic skills or self-reported perceptions of self-efficacy as a predictor of actual diagnostic skills in pre-service teachers. Motivation, attitude, and knowledge were found to be significant positive predictors of diagnostic skills with respect to learning behavior [125], whereas reflection on experience and self-efficacy were not found to be relevant. In a study of German secondary mathematics teachers at two measurement points, a causal effect of teacher self-efficacy expectations on subsequent instructional quality (self-reported teachers' self-efficacy and instructional quality) was partially found [126]. Given the primarily heterogeneous and partly contradictory findings, the need for research on self-efficacy or self-efficacy expectations in the context of diagnostic skills becomes clear.

1.5. Claim and Research Questions

In addition to content knowledge about the addressed context in an experiment and about experimentation (CK), the teacher needs pedagogical content knowledge about typical difficulties of students in the experimental implementation of the context as well as about possibilities for action in the experimental instructional setting (PCK) [127,128]. In this study, these dimensions of professional knowledge are addressed as important components in the formation of diagnostic competence. Consequently, the promotion of diagnostic competence requires that these areas of knowledge are either developed or are already present in the pre-service teachers. Especially in the first third of university studies, it is important to clarify what previous knowledge students bring with them or do not bring with them in the development of diagnostic competence for the assessment of students' skills to experiment or are not developed in the basic subject didactic training. On this basis, university teaching-learning programmes can be optimized and tailored to the target group. Since research studies in science education with a focus on the evaluation of diagnostic competence in combination with subject-specific pedagogical knowledge are rare so far [129,130], an explorative approach was chosen for the present study. In a chemistry course and a biology course, respectively, at two institutions at the University of Braunschweig and the University of Kassel the first step was to qualitatively investigate the professional knowledge of methodological difficulties and central decisions/actions of students in science experimentation among pre-service teachers in the first third of their university teaching studies. For this purpose, the following explorative-qualitative research questions were considered:

RQ1: Which difficulties in experimentation pre-service teachers with the subject biology and/or chemistry are able to describe on the basis of their pedagogical content knowledge (PCK) in the first third of their university studies?

RQ2: To what extent can pre-service teachers with the subject biology and/or chemistry describe central methodological contents and actions for experimentation in the first third of their university studies and what (methodological) content knowledge is predominant here (CK)?

A number of individual and contextual factors may influence the willingness and ability of (pre-service) teachers to implement diagnostic activities (while experimenting) in the teaching profession (cf., e.g., [125,131]). Attitudes toward the relevance and importance of diagnostic content in teacher education and, more broadly, the teaching profession are difficult to predict due to overlaps in the addressed knowledge domains. As a content element of pedagogical/didactic and/or educational study elements, both higher and lower attitude expressions would be expected according to the findings of Cramer [132]. Similarly, for the area of self-efficacy expectations with the specification of diagnostic competence in subject-related settings of experimentation, there is a lack of empirical

findings that can be used to make an educated guess about the expression in the sample studied here. Consequently, the following additional descriptive-quantitative questions were examined in order to draw statements about the expression of diagnosis-related attitudes and self-efficacy expectations among pre-service teachers:

RQ3: What attitudes towards the relevance of diagnostics in teacher training and what self-efficacy expectations for diagnostics in experimental settings show pre-service teachers in the first third of their university studies and how are they related in terms of expression?

RQ4: Is there a relation between pre-service teachers' attitudes towards the relevance of diagnostics in teacher training resp. self-efficacy expectations for diagnostic activities in experimental settings:

- a. . . . and their pedagogical content knowledge about difficulties of students in experimentation?
- b. . . . and their (methodological) content knowledge about central components and decisions in scientific experimentation.

2. Materials and Methods

2.1. Procedure

The written survey on which this paper is based was conducted at the beginning of two regular obligatory courses identified in the module plan from the 2nd semester of the (bachelor/teacher) degree programme in chemistry at the University of Braunschweig and biology at the University of Kassel. The procedure and information provided, as well as the time frame for completion (approximately 30 min) were identical in the cohorts. An online questionnaire was used to collect (a) demographic and academic information, (b) pedagogical content-related diagnostic knowledge about student difficulties and central actions or decisions of students in experimentation, and (c) the relevance of diagnostics in teacher training as well as domain-specific self-efficacy expectations. When recording the self-efficacy expectations, the participants were free to specify them in relation to the subject biology or chemistry. They had to make a selection beforehand and were assigned to the chemistry or biology sub-sample according to this selection (see Section 2.2).

Participation in the survey was anonymous and voluntary. For the release of the socio-demographic, quantitative, and qualitative data of the closed-ended and open-ended questions, the participants provided a declaration of consent for anonymous analysis and publication.

2.2. Participants

The sample consisted of 51 pre-service teachers of biology and/or chemistry. Of these, 66.7% were female and the average age was 22 ± 2.6 years. Of the participants, 34 were studying to become teachers at grammar schools and 17 were studying to become teachers at secondary schools. The relevant sociodemographic data for both the total sample and the sub-samples are presented in Table 2. In relation to the sub-sample, the proportions of pre-service teachers in their second or fourth semester were 47% in biology and 62% in chemistry. While a further 43% of pre-service teachers in biology were in their 6th semester, the remaining proportion in chemistry was distributed among the higher semesters. The average number of semesters is similar in both sub-samples, as is the distribution of gender and the type of school targeted in the study programme (see Table 2).

Table 2. Demographic characteristics of the participants.

	<i>n</i>	Age in Years	Sex		School Type		Semester in Biology	Semester in Chemistry
		<i>M</i> (<i>SD</i>)	Female <i>n</i> (%)	Male <i>n</i> (%)	GYM <i>n</i> (%)	HR <i>n</i> (%)	<i>M</i> (<i>SD</i>)	<i>M</i> (<i>SD</i>)
Total sample	51	22.0 (2.6)	34 (67)	17 (33)	34 (67)	17 (33)		
Sub-sample biology	30	21.9 (2.5)	22 (73)	8 (27)	19 (63)	11 (37)	4.7 (1.8)	
Sub-sample chemistry	21	22.2 (2.8)	12 (57)	9 (43)	15 (71)	6 (29)		4.4 (2.1)

Annotation: *n* = number; *M* = mean; *SD* = standard deviation; GYM = “Gymnasium” (grammar schools); HR = “Haupt-/Realschule” (secondary schools).

Approximately three-quarters of the participants had taken one or two courses in subject didactics at the time of the survey (biology 67%, chemistry 77%); all others had already taken three or more courses in subject didactics. At the University Braunschweig, the largest proportion of pre-service teachers recruited were those from the chemistry sub-sample. Only two participants of this sub-sample completed the survey in the context of a chemistry didactic course at the University Kassel. Further participants from this course could not be included in the analysis due to too high a semester number and missing data. Accordingly, a description of the university teaching and pre-conditions for the chemistry sub-sample focuses on the curricular structures at the University Braunschweig. Here, pre-service teachers take subject courses in the first semesters, in which investigations tend to follow detailed instructions in laboratory practicals. Content knowledge and skills in natural sciences working methods and techniques and handling laboratory materials are to be developed. In the 4th semester, pre-service teachers usually attend courses with didactic content for the first time, in which, among other things, the hypothetical-deductive processes for scientific inquiry and experimental problem solving are addressed. The data collection took place at the beginning of the seminar “Simple scientific experiments”.

The biology sub-sample mainly consists of pre-service teachers with biology as a subject from the University of Kassel. Here, too, there is an exception to two participants who come from the University of Braunschweig and who chose biology rather than chemistry in the survey section on domain-specific self-efficacy expectations. However, since more than 90% participated in the survey at the beginning of a course in the didactics of biology at the University Kassel, the curricular structures available here are used to describe the university teaching- and pre-conditions for the biology sub-sample. The first two semesters of the biology teaching programme at the University of Kassel are also dominated by subject-related biology courses (incl. laboratory). In addition, the basic module “Introduction to Biology Didactics” with lecture and exercise should be completed in this stage. In this module, two individual sessions provide basic background and information on the scientific inquiry process and the associated subject-methodological (procedural) knowledge. In each subsequent semester there are further subject didactic courses with different emphases. The survey took place at the beginning of the course “Scientific inquiry methods and lab techniques in biology teaching”, which is central to the content area of scientific inquiry; it can therefore not be assumed that the participants have received in-depth training in this area.

2.3. Instruments

Some authors critique a lack of validity evidence for instruments to assess scientific reasoning competencies (e.g., [133,134]) and point out that multiple choice assessment can hardly be seen as situations closely representing real life (e.g., [135]). In some studies with students response processes have also been examined qualitatively (through thinking aloud, eye-tracking studies, video recordings and written recordings), and these studies confirmed that respondents use procedural and epistemic knowledge (e.g., [136,137]).

Moreover, as science is constituted among others by specialized language [138], a central part of this study is a qualitative instrument to determine pre-service teachers' scientific reasoning knowledge concerning student' difficulties.

2.3.1. Student Difficulties/Misconceptions and Actions in Experimentation

The central concern of the open-ended questions posed here in two parts is to record the pre-service teachers' Pedagogical (didactic) Content Knowledge (PCK) about typical students' difficulties/misconceptions as well as their Content (methodological) Knowledge (CK) about the processing of an experimental task for an authentic teaching scenario. The latter is presented to the subjects in the form of a short progression plan from the perspective of the teacher and the experimental task given to the students via a task sheet (see Supplementary Materials S1). The phenomenon to be investigated here for dissolving sugar is visualized to the (fictitious) students as well as to the pre-service teachers by a video with a dialogue between two friends at and about a cup of tea ('tea conversation') and used to derive the question "How does the time needed to dissolve sugar depend on different influencing factors?" This task stem is followed by two open-ended questions or subtasks. In part 1, the central difficulties of the students in accomplishing the set experimental task or the planning and execution of an experiment on the dissolving time of sugar are examined with regard to the content knowledge and the experimental implementation. Part 2 explicitly aims to attain methodological (procedural) knowledge for the design of experiments in the sense of the scientific inquiry process and to also indirectly overcome possible obstacles to master this process for students. In this part, four central decisions or necessary actions of the students to accomplish the experimental task are to be articulated by the pre-service teachers in writing.

2.3.2. Relevance of Diagnostics in the Teacher Training Program

The measurement instrument for assessing attitudes toward the importance of diagnostics in teacher education includes five items, the language and content of which were adapted to fit the focus of the present study, based on Lorenz [139]. Based on the characteristic values obtained in this sample for Cronbach's alpha [140] (pp. 281–302) with 5 items, $\alpha = 0.64$, and a low discriminatory power for the item: it is important for teachers to be able to correctly assess students' performance in experimentation ($r_{it} = 0.165$), a reduction of the scale to four items is made ($\alpha = 0.70$). A complete overview of all items can be found in Appendix A.

2.3.3. Domain-Specific Self-Efficacy Expectations

Based on the concept of self-efficacy as devised by Schwarzer and Jerusalem [141], an instrument was developed for the domain-specific assessment of subjective certainty in coping with diagnosis-related teaching activities in general, as well as in instructional experimental settings. In addition to an explicit focus on this selected activity domain, the item formulation describes the perceived ability to be assessed here as "proficiency". Moreover, the ability of an action is specified in its effectiveness via the inclusion of challenges or obstacles or barriers to action in the item formulation [109]. The items used in the present survey [142] were administered and piloted to 98 pre-service teachers majoring in biology over three semester cohorts. Based on the piloting data, the dimensionality of the scale was tested using exploratory factor analysis (principal component analysis with varimax rotation) and internal consistency was tested using Cronbach's alpha. Factor analysis revealed two factors with an eigenvalue > 1 that explained 59% of the variance. On the first factor, four items loaded highly (factor load (a_{jq}) > 0.649). One additional item loaded similarly on both factors; based on content, this item was assigned to the second factor. Thus, factor one (experiment-related diagnostic activities) is represented by four items whose content is explicitly diagnostic and/or experimental. On the second factor, another three items loaded similarly highly ($a_{jq} > 0.709$). Together with the content-related item, this factor describes diagnostic activities implicitly and without the inclusion of

experimentation/processes (general diagnostic activities). The reliability of the two scales formed according to the two-factor model was in the good range ($\alpha \geq 0.70$). The items and (factor analytic) findings for piloting the instrument on domain-specific self-efficacy expectations can be found in Appendix B.

In the survey presented in this article, a further subject-related specification took place in the item formulation and content orientation. In accordance with the sub-samples, the subjects were provided with either biology or chemistry items to assess their self-efficacy expectations. Using a 4-point Likert scale ranging from “does not apply at all” (1) to “fully applies” (4), the pre-service teachers’ self-reports were recorded in relation to the respective items. The empirically derived and validated scales from the pilot were adopted; the biology sub-sample showed in the very good range ($\alpha \geq 0.83$) and the chemistry sub-sample showed acceptable Cronbach’s alpha values ($\alpha \geq 0.61$).

2.4. Data Analysis

2.4.1. Qualitative Analysis Methods

The written student responses on knowledge about student difficulties and actions/procedures during experimentation were analysed qualitatively according to the procedure of a summarising content analysis with deductive-inductive category formation [143] using the programmes MAXQDA 2020 and Excel 2016. For the first part with written pre-service responses on student difficulties and misconceptions, the formation of content categories was initially carried out deductively with a theoretical foundation based on research findings, which primarily explicate the prior knowledge of the researchers involved in the study [39,47], as well as on results from a wide range of literature sources (see Table 1). Where possible, the references were classified in the process-related sub-steps of experimentation in Mayer’s [34] structural model of scientific reasoning, whereby the formulation of the question was already specified in the material. In the second part (procedure for experimenting), the central decisions listed were sorted into subcategories inductively formed on the material on the basis of similarities in content. These were then classified by subsumption into supercategories based on the hypothetical-deductive procedure in the experiment [34]. This systematic inclusion of deductively and inductively formed categories served to identify and explore further meaning components in the process of creating the category system up to the category definition through anchor examples [144]. The assignment of the student responses to the categories for the 1st and 2nd part was carried out independently by two trained raters. Interrater reliability was estimated using Cohen’s Kappa with the programme IBM SPSS Statistics (version 27) and is ‘substantial’ in the 1st part (Cohen’s $\kappa = 0.80$; $p \leq 0.001$) and ‘almost perfect’ in the 2nd part (Cohen’s $\kappa = 0.903$; $p \leq 0.001$) [145]. Finally, the relative frequency for each code in each category was calculated for each group of biology or chemistry students.

2.4.2. Quantitative Methods of Analysis

In order to analyse the attitudes towards the importance of diagnostics in teacher education and domain-specific self-efficacy expectations towards diagnostic skills/actions in teaching-learning settings for science experimentation, descriptive procedures, representations, and associated characteristic values were used and analysed with the programme IBM SPSS Statistics (version 27). In the first step, deductively derived as well as newly constructed items for the assessment of self-efficacy expectations were used in a pilot sample in order to test them accordingly factor-analytically and in their reliability (see Section 2.3.3). A further reliability test was also carried out for the two empirically derived scales on domain-specific self-efficacy expectations as well as for the newly constructed scale on the significance of diagnostics in teacher training in the sample on which this study is based. Taber [146] was used to assess the Cronbach’s for internal consistency. In the next step, the sample was analysed in the subjects included here and the constructs examined in each case (Mann–Whitney U test; Wilcoxon test) in order to uncover any differences in attitudes and self-assessments that could influence possible correlations

with the results on the open-ended task and allow conclusions to be drawn about subject specifics. According to Cohan [147], possible effects associated with this are rated as insignificant for $r < 0.10$, weak for $r = 0.10\text{--}0.30$; medium for $r = 0.30\text{--}0.50$ and strong for $r > 0.50$. In accordance with the goal of describing a comprehensive picture of the extent of subject-specific methodological (procedural) knowledge of experimentation in combination with diagnosis-related competence assessments in this area, frequencies and mean values in the respective characteristics are reported and, in the last step, the correlations between the quantitative and qualitative data are exploratively tested via correlations (Spearman rank correlation). Due to the small samples that predominate, especially in the subject sub-samples, and a violation of the criterion for normal distribution in selected scales, non-parametric procedures were used throughout.

3. Results

In total, 49 fully completed questionnaires by the pre-service teachers for the first sub-task (student's difficulties) and 50 for the second subtask (procedure for experimenting/key decisions) were included in the predictive diagnostics.

(RQ1) To analyze participants' responses to the first subtask, process-related difficulties and misconceptions among students described in the theoretical literature (see Table 1) were applied to the material. The pre-service teachers' responses were paraphrased prior to analysis, which involved rewriting the responses' core content in a concise descriptive form [143]. This then allowed them to be reliably assigned to the categories. For example, the statement "... The students do not know what things are important/what one needs to pay attention to—no connection to the research question—no hypotheses, ..." was reformulated into the core components "creating a link to the research question" and "no hypothesis generated". The statement "... should be clarified and potentially what a conjecture/thesis entails" was reduced to "technical term conjecture/thesis not known". A total of 293 statements could be coded from the 49 answer sheets examined. Of these, 166 statements referred to process-related difficulties and/or misconceptions by students, which corresponds to roughly 57% (Table 3). In addition, the respondents frequently mentioned difficulties arising from the "instructional setting", specifically from the open-ended task structure and materials pool, which overwhelmed students and required them to make decisions. Statements referring to students' decisions about how to divide up tasks and/or disagreements within the group were assigned to the category "social format". Statements referring to teachers' instructional planning were assigned to the category "teacher". A total of 94 statements (32%) were assigned to these non-process-related categories. Difficulties related to subject-specific content knowledge or technical terms were mentioned in 33 statements (11%).

In the following, the difficulties that students have to deal with during each sub-steps within the hypothetico-deductive scientific inquiry process according to the pre-service teachers' statements are explained (see Table 3).

In terms of the **phenomenon**, the pre-service teachers mentioned a lack of or differences in prior knowledge. (The students will not yet be familiar with the phenomenon of diffusion; [...] was not worked through as a group, meaning that the students will not be able to think deeply about the phenomenon. It also seems that no ideas were taken up or discussed. Hence, there is no bridge to their prior knowledge [...]).

Potential difficulties in dealing with the provided **research question** to be investigated in the fictitious experimental instruction setting (see Supplementary Materials S1) were also identified. These difficulties concerned the students failing to understand or refer back to the research question ("The students might not be able to make connections between the conducted experiment and the research question").

Table 3. Overview of the students' process-related misconceptions and difficulties faced during experimentation identified by the pre-service teachers (incl. paraphrases) according to the three phases for scientific reasoning of the SDDS model [22].

Difficulties Students Face (Deductive, See Table 1)	Paraphrase from Material	<i>h</i>
Phase: Search Hypothesis		
Phenomenon		
Different levels of prior knowledge	Students have sharply divergent understandings (or sometimes no understanding) of how sugar dissolves in tea.	13
Research question		
Integrating the research question into the experimentation process	Perhaps a too complicated research question that might lead to misunderstandings during implementation. As a result, there is no bridge to students' prior knowledge and the students cannot incorporate their contributions into the research question.	8
Hypothesis Generation		
No hypothesis generated	No hypotheses.	1
No understanding of hypothesis generation	No background subject-related knowledge is present, nor is fundamental knowledge of scientific knowledge acquisition through the generation of hypotheses.	2
Suppositions not linked to research question	[...] generate a hypothesis that makes reference to the research question and subsequently guides the experimentation phase.	1
Phase: Planning and Testing		
Planning		
Selection of materials unsystematic, incomplete and/or via trial-and-error	[...] that they throw together materials at random. [...] there are still problems with the selection of materials, since typically only the necessary materials are made available and the students try to use everything, even when it's not necessary.	19
Trying things out in an unstructured way (no plan, change all variables)	Confusion with respect to materials selection, since more materials are available -> perhaps the students want to switch to using other materials during the experiment. [...] change their minds while conducting the experiment if they get the feeling they have selected the "wrong" factor.	3
Planning an experiment that (does not) achieve its objective	No foundation for planning an experiment.	3
Dealing with variables in an unsystematic way (multiple variables are confounded or a single variable is varied unsystematically)	Students conduct the experiment with two independent variables simultaneously, which does not lead to an unambiguous result.	19
Aware of confounders, but do not eliminate them	When conducting the experiment, it could be difficult to stick to one influencing factor and not hold constant conditions constant. That they do not recognize or eliminate confounders when constructing a self-developed experiment.	2
Lack of control group approach	Furthermore, students often forget to create blinded or comparison samples.	2
Variable selection		
Deciding which variables to include	[...] which influencing factors play a role in the solubility of sugar in water. [...] which experiment should investigate a certain influencing factor.	42
Conducting the experiment		
Variables operationalized not at all or incorrectly	Imprecision in keeping time, since the time point at which all of the sugar has dissolved is often difficult to determine.	14
Imprecise measurement/Measurement error	[...] that they measure the water by eyeballing it or count without looking at a clock.	
Order of experimental steps	[...] students probably do not yet know exactly how an experiment with a phenomenon, planning, observation and interpretation is structured. I consider difficulties in conducting the experiment graver here [...].	12
Conducting experiment incorrectly	Experiment not conducted in a structured way.	6
Documentation	Measurements not recorded.	3

Table 3. Cont.

Difficulties Students Face (Deductive, See Table 1)	Paraphrase from Material	<i>h</i>
Phase: Evaluate Evidence		
Unexpected data are attributed to errors in conducting the experiment	They might change their results when they have the feeling that something is not right or they have done something incorrectly.	1
Insufficient reflection on the experimental results	Errors are not taken into consideration.	1
Replicability/validity	[...] replicability or precision of measurement will pose problems for the students.	2
Wrong conclusion drawn from consistent experiments	The students might not be able to correctly interpret their observations.	1
Conclusions are not possible because variables were not varied systematically, and/or fallacious conclusions are drawn	[...] not investigate multiple influencing factors at the same time, since they then cannot say which factor actually influences the sugar's dissolution. Analysis does not take place, causal conclusions are not possible.	10
Data are not related (back) to hypothesis/research question	[...] hypotheses are actually confirmed or falsified based on the insights gained while conducting the experiment.	1

Annotation: *h* = absolute frequency of mentions from a total of *n* = 49 pre-service science teachers.

Only a few statements mentioned difficulties regarding **hypothesis generation** compared to the later steps of the process. Three participant statements referred to the aspects of having no hypothesis, hypothesis formulation, and generating multiple hypotheses, while one statement referred to the link between the supposition and research question. No pre-service teachers mentioned formulating justifications for hypotheses as a difficulty faced by students. One pre-service teacher wrote: "To me, the connection between technical terms and phenomena observed in everyday life does not seem to be pronounced enough in sixth grade in order to bridge the gap from multiple hypotheses [...] to independently conducting an experiment to test these hypotheses". This statement addresses the association between the hypothesis and the need to plan an experiment that successfully tests the hypothesis.

With regard to **planning**, no respondents mentioned potential problems connecting the hypothesis and experiment, i.e., planning an experiment that is able to actually test the proposed hypothesis. In contrast, difficulties with planning a meaningful experiment and selecting appropriate materials (e.g., utilization of the materials pool) were mentioned very frequently ("It is possible that the students might become overwhelmed by the various materials on offer"). Application of the control-of-variables strategy was also described by many respondents ("Moreover, various factors might be unwittingly coupled with one another, meaning that only a single factor is not investigated"). Three mentioned difficulties concerned trying things out in an unstructured way ("no plan-change all variables") ("that the students change their minds while conducting the experiment if they get the feeling they have selected the 'wrong' factor").

Variable selection was included as a unique overarching category, since it accounted for 25% of process-related statements and was mentioned by 28 participants ("It could be the case that no influencing factors are selected, but only changes that exert no influence"; Ultimately, the students need to know and/or be able to identify the influencing factors in order to independently develop an experimental approach"; "A difficulty for students would be recognizing the difference between loose sugar and sugar cubes").

With respect to **conducting the experiment**, general problems were mentioned: "Difficulties arise in conducting the experiment". The spectrum of statements referring to students' difficulties with respect to failing to or incorrectly operationalizing variables, from imprecise measurements to measurement errors, ranged from taking imprecise measurements with the stopwatch to the use of different amounts and/or volumes of the

chemicals and measurement differences between students (“When conducting the experiment, replicability or precision of measurements will pose problems for the students.”; “constantly switch the person doing the measurement, which can lead to imprecise measurements”). Students’ difficulties regarding documentation also fall under this category (“that the results need to be documented; the learners might not do this and forget the times they measured, for example”). With respect to evaluating evidence, only one respondent mentioned that students might attribute unexpected results to an error made in conducting the experiment (“They might change their results when they have the feeling that something is not right or they have done something incorrectly.”). Insufficient reflection by the students on the experimental results was mentioned once (“Errors are not considered—No analysis takes place (causal conclusions are not possible).”). The most frequently mentioned difficulty was that no conclusions are possible because the variables were varied unsystematically (“Students conduct the experiment with two influencing factors simultaneously, which does not lead to an unambiguous result”; “Focusing on just one aspect is probably difficult for some groups of students, so they try to investigate multiple factors simultaneously and only realize at the end that they cannot draw any conclusions from the investigation”).

Problems with **evaluating evidence** due to lack of a comparison or blinded sample were also mentioned (“Furthermore, students often forget to create blinded or comparison samples and thus cannot draw any concrete conclusions from their results”).

Overall, these problems were attributed to students being overwhelmed by the “lack of guidance” in the self-directed procedure, which did not involve following prescriptive, externally prescribed experimental steps (“Moreover, it is not precisely described how the experiment should proceed”; “that the students can become overwhelmed by this freedom and autonomy”). These statements were assigned to a non-process-related overarching category, the “instructional setting”, which the pre-service teachers referred to 71 times, with particular focus on obstacles stemming from the open-ended nature of the task. Additionally included in this category was students’ lack of experience in dealing with experimental materials, which was mentioned 10 times ([...] difficulty correctly using the given materials”; [...] that the students are not familiar with all the materials and how they are used”). Overall, 94 statements were coded into one of three such non-process-related categories, the “instructional setting”, “social format” and “teacher”. Twelve statements concerning students’ decisions about how to divide up tasks and/or disagreements within the group were assigned to the category “social format” (“Students cannot come to an agreement within the group”). Eleven statements addressing the teacher’s instructional planning were assigned to the “teacher” category (“[...] handouts to assist students or opportunities to repeat explanations are not included in the experiment, which could lead to excessive questions”; “Since no thermometer is available, the students cannot make any statements about the water temperature”). Beyond these overarching categories and subcategories, difficulties related to subject-specific content knowledge and technical terms were also reported. Lack of familiarity with technical terms was mentioned 14 times, with the term “influencing factor” coming up seven times, “conjecture” three times and other technical terms four times (“have problems finding out what influencing factors are exactly”; “even the term ‘influencing factor’ should be clarified and possibly also what a conjecture/thesis entails”) (“Effect on dissolution speed is difficult in sixth grade”). That subject-specific content knowledge might be lacking, insufficient or fragmentary was mentioned 19 times (“With respect to subject-specific content knowledge, the difficulty might arise that some students do not know that sugar dissolves more quickly and easily at higher temperatures”; “Background subject-related content knowledge is not yet present”).

To summarize, the respondents considered here ($n = 49$) mentioned four of the nine overarching categories on average ($M = 4.00$, $SD = 1.26$; $\min = 2$; $\max = 7$) in their written responses asking about potential difficulties faced by students, including roughly two of the five process-related overarching categories (phenomenon, hypothesis generation/research question, planning, variable selection, conducting the experiment, and evaluating evidence)

($M = 2.33$, $SD = 1.18$; $\min = 0$; $\max = 4$). Three pre-service teachers did not mention any process-related difficulties. Statements falling under the overarching category of hypothesis generation were by far the least frequent (2% of the total number of statements in the process-related categories). Considered together with the phenomenon and the research question, this rose to 10% of all statements in the process-related categories, which is similar to the share of statements referring to evaluating evidence (15%). Planning was mentioned most frequently by the respondents (29% of all statements in the process-related categories), with the lion's share referring to selecting appropriate materials and the control-of-variables strategy, each of which made up 11% of all statements in the process-related categories.

Differences in the quality of participants' statements were related to their subjects of study, with the pre-service chemistry teachers using technical terms like "familiarity with the RGT rule" (authors' note: Reaction velocity-Temperature-Regulation), "materials surface", "solubility product", and "law of mass action" more frequently than the pre-service biology teachers, who used more general formulations like "dissolution time depends on the temperature", "recognizing the difference between loose sugar and a sugar cube", "factors like the solubility or saturation of a liquid and corresponding effect on the dissolution speed". The only further differences between the biology ($Mdn = 6.00$) and chemistry pre-service teachers ($Mdn = 4.50$) uncovered concerned the total number of difficulties identified in all nine overarching categories (Mann-Whitney U-test: $z = -2.377$, $p = 0.017$; $r = 0.34$). Examining the frequencies of the process-related and non-process-related categories revealed that this difference reflected a larger share of non-process-related categories in the biology pre-service teachers' statements. Consequently, no significant differences in the number of process-related difficulties mentioned were found. A total of 24% of the biology pre-service teachers ($n = 29$) referred to the groups' social fabric as it related to completing the tasks, compared to just 10% of the chemistry pre-service teachers ($n = 20$). A similar pattern was found for the "teacher" category, which was mentioned by 24% of biology pre-service teachers but only 5% of chemistry pre-service teachers.

(RQ2) The university pre-service teacher statements about four key decisions students need to make when solving problems experimentally could be assigned to seven superordinate categories and 26 subcategories (Table 4). A total of 325 statements were analyzed, of which 298 were assigned to the following process-related superordinate categories, which referred to phases of the experimentation process [22,34]: decisions about the "phenomenon", the "research question and/or hypothesis", "working with and identifying variables", "planning", "conducting inclusive documentation", "analysis and interpretation". All statements by participants were analysed, regardless of the number of decisions a respondent mentioned, which ranged from 2 to 14 in total (both process-related and non-process-related).

On average, the pre-service teachers mentioned decisions about a bit more than three (3.34) process phases/elements (=superordinate categories). In addition, five pre-service teachers made six mentions of decisions by teachers, such as ensuring appropriate group composition or an appropriate task. None of the participants mentioned all of the process phases/elements considered here in their responses. Turning to the subcategories, the pre-service teachers most frequently mentioned decisions related to the influencing factors (around 20% of the 298 total mentions of process-related decisions) and the experimental materials (12%), in line with the difficulties students were considered to face in this area (see Table 3). Among decisions about "variables" and "planning", the aspects of avoiding confounders, mentioned twice by two different pre-service teachers, and including a control group, mentioned three times by two different pre-service teachers, were grossly underrepresented. This was also the case for estimating the experimental validity within the superordinate category of "analysis and interpretation", which was mentioned by three pre-service teachers one time each.

Table 4. Overview of categories (superordinate and subcategories) in the statements about key decisions students face when solving problems experimentally, as well as corresponding anchoring examples and number of mentions.

Superordinate Category Decisions About	Subcategory	Anchoring Example	<i>h</i>	<i>f</i>
... the phenomenon	Retrieving (prior) knowledge about the phenomenon	The students need to visualize the phenomenon and think about how they can link it to their prior knowledge.	4	
			4	8%
... the research question and/or hypotheses	Develop a research question/hypothesis	To start with, the students should think about a research question (or a conjecture) that they can subsequently answer with the experiment.	17	
	Working with the research question/assigned task	They need to come to agreement amongst themselves on how they can best express the characteristic to be observed and in what way they will investigate it.	2	
	Understanding the research question	Students need to understand the research question.	1	
			20	36%
... working with and identifying variables	Selecting one influencing factor	The students need to decide which influencing factor they want to test.	59	
	Controlling for other factors	Investigate one factor in the experiment! [...] pay attention to other factors that remain constant.	12	
	Decisions about the measured variable	It needs to be determined when the time will be stopped (when has the sugar dissolved?).	8	
	Avoiding confounders	[...] As part of this, the students must minimize the presence of confounding factors.	3	
	Decisions about the dependent variable (degree of breakdown, amount)	They need to agree on whether to use ground sugar or sugar cubes.	18	
			100	86%
... planning	Selection of appropriate materials	Decision about selecting materials from the list of materials.	36	
	Planning the experiment	[...] decide how they will proceed.	17	
	Determining the order of work steps	The learners need to familiarize themselves with which work steps they will conduct in which order.	4	
	Planning a control group experiment	At least one comparative experiment must be conducted.	4	
	Replication/Reliability	Conduct each experiment at least twice.	4	
			65	84%
... conducting the experiment (including documentation)	Conducting the experiment	Conduct the experiment in accordance with the experimental plan.	24	
	Documentation	Selecting documentation of the experiment.	21	
	Observation	The students now observe [...].	15	
			60	70%
... analysis and interpretation	Conclusion	They need to interpret their results (conclusion).	17	
	Evaluating evidence/causal relationships	Based on what evidence they can see how the influencing factor they are investigating affects the sugar's dissolution.	8	
	Reflecting on errors	[...] pay attention to errors that might have crept in and potentially conduct the experiment again.	7	
	Referring back to the hypothesis/research question	Test the hypothesis	6	
	Paying attention to validity	Analysis (Was the hypothesis refuted or not?) Does the experiment I have conducted actually answer my question?	9 2	
			49	50%

Table 4. Cont.

Superordinate Category Decisions About	Subcategory	Anchoring Example	<i>h</i>	<i>f</i>
Decisions unrelated to the scientific method	Dividing up tasks within the group	They need to divide up the various tasks within the experiment [. . .]	11	
	Completing the group work together	The group should work together so that everyone is aware of what has been done.	9	
	Working precisely and carefully	The students need to work very precisely.	3	
		Time management: “The time allotted should also be considered.”	2	
	Other	Cleaning up after the experiment	1	
	Alignment with teacher (minimizing errors)	1		
			27	38%

Annotation: *h* = absolute frequency of mentions from a total of *n* = 50 pre-service science teachers; *f* = relative frequency of mentions in the subordinate category (%); each respondent could identify multiple influencing factors.

Overall, the pre-service teachers frequently adopted a results-focused perspective: “[...] After conducting the experiment, the students should collect the results, discuss them as a group [...]”; “How can I answer my question for others. How can I be sure of my result, does the experiment I have conducted really answer my question”. Around half of the pre-service teachers adopted an understanding-oriented perspective and explicitly mentioned the need for students to “understand the research question/assigned task” and the need to make decisions about “selecting appropriate materials” or the “order of work steps”. In this context, references were made to the open-ended structure of the task, the third most commonly mentioned difficulty for students in Part 1 (see Table 3), (“The learners need to familiarize themselves with which work steps they will conduct in which order”; “... it could be difficult if they do not have sufficient practice in conducting experiments and the students forget steps like documenting the results”; How do I conduct the experiment, which work steps logically follow one another”) as well as decisions about selecting materials from the materials pool (“They need to ascertain what materials are necessary to match the selected influencing factor”; “They need to decide what materials they want to use and which ones they don’t need or don’t want to use”). Three pre-service teachers focused exclusively on deciding on an influencing factor, such as water temperature or water volume, or deciding whether to stir or shake the samples during the experiment.

When comparing the two subsamples, it was found that the biology pre-service teachers (Mdn = 4.00) made reference to more process steps (=superordinate categories) when discussing decisions that need to be made in carrying out the example experiment than the chemistry pre-service teachers (Mdn = 3.00; Mann-Whitney U-Test: $z = -2.639$, $p = 0.008$; $r = 0.37$). A similar pattern was found for the total number of process-related decisions across all subcategories among the biology (Mdn = 6.00) and chemistry pre-service teachers (Mdn = 5.00). The biology students reported more unique process-related decisions here as well; the difference just failed to reach significance and represented a small effect ($r = 0.27$; Mann-Whitney U-Test: $z = -1.931$, $p = 0.056$).

(RQ3) The university pre-service teachers’ attitudes toward including diagnostic elements and promoting diagnostic skills in university teacher education were positive and quite pronounced. The full-sample mean was substantially above the scale midpoint, at $M = 3.13$ ($SD = 0.395$), (see Appendix B). No significant differences in the strength of these attitudes were found between the two subsamples ($Mdn_{\text{Chemistry}} = 3.00$; $Mdn_{\text{Biology}} = 3.25$; $z = -1.414$, $p = 0.157$). Likewise, there were no significant differences between the two subsamples of pre-service teachers in perceived domain-specific self-efficacy with respect to general ($Mdn_{\text{Chemistry}} = 2.50$; $Mdn_{\text{Biology}} = 2.50$) as well as experimentation-related diagnostic activities ($Mdn_{\text{Chemistry}} = 2.50$; $Mdn_{\text{Biology}} = 2.25$). However, there was a significant difference in the full sample between the two perceived self-efficacy scales and attitudes towards the importance of diagnostics in university teacher education. The

pre-service teachers' subjective perception of their own skills in successfully carrying out general and experimentation-related diagnostic activities ($Mdn_{\text{Self-efficacy factor1/2}} = 2.50$; $Mdn_{\text{Relative importance of diagnostics}} = 3.00$) was lower than their perception of the importance of this teaching and learning topic for university teacher education (Self-efficacy factor 1: $z = 5.612$, $p < 0.001$; Self-efficacy factor 2: $z = 5.345$, $p < 0.001$; $n = 50$). The effect sizes can be considered large, $r > 0.75$. However, no correlation between these two motivational constructs was found. The pre-service teachers saw potential for improvement in the diagnostic competences referred to in the perceived self-efficacy scale, as they tended to "somewhat disagree" or "somewhat agree" to these items (e.g., $M_{\text{Self-efficacy factor1-Bio}} = 2.38$, $SD = 0.61$; Table 5).

Table 5. Item scores for the domain-specific perceived self-efficacy scales.

Item	Biology ($n = 30$)		Chemistry ($n = 20$)	
	M (SD)	r_{it}	M (SD)	r_{it}
<i>Experimentation-related diagnostic activities</i>				
Even when I am experiencing stress, I can still diagnose students' errors in experimentation in biology/chemistry class well.	2.20 (0.71)	0.708	2.40 (0.60)	0.524
I am certain that I am able to recognize children's specific difficulties in experimentation in biology/chemistry even when under a large amount of time pressure.	2.40 (0.77)	0.737	2.75 (0.44)	0.447
In biology/chemistry, I am able to accurately assess my students' level of learning prerequisites, even when I have little time available.	2.43 (0.63)	0.603	2.65 (0.50)	0.380
In biology/chemistry, I am able to accurately assess my students' experimentation skills, even when I have little time available.	2.37 (0.72)	0.803	2.45 (0.51)	0.430
	$M_{\text{scale}} = 2.35$ $SD_{\text{scale}} = 0.60$ Cronbachs $\alpha = 0.86$		$M_{\text{scale}} = 2.56$ $SD_{\text{scale}} = 0.36$ Cronbachs $\alpha = 0.66$	
<i>General diagnostic activities</i>				
I am able to successfully integrate diagnostic activities to accompany learning in my biology/chemistry instruction, even when I am under time pressure.	2.13 (0.78)	0.628	2.20 (0.62)	0.217
Despite a high level of heterogeneity, I am able to create tasks in biology/chemistry that allow me to appropriately check both weaker and stronger students' knowledge levels.	2.57 (0.77)	0.665	2.50 (0.69)	0.546
I am able to successfully take into account students' learning processes when formulating individual learning goals in biology/chemistry, even when these differ markedly.	2.47 (0.78)	0.743	2.35 (0.67)	0.440
In biology/chemistry, I am able to accurately assess my students' thought and work processes, even when I have little time available.	2.37 (0.67)	0.576	2.55 (0.51)	0.393
	$M_{\text{scale}} = 2.38$ $SD_{\text{scale}} = 0.61$ Cronbachs $\alpha = 0.83$		$M_{\text{scale}} = 2.40$ $SD_{\text{scale}} = 0.42$ Cronbachs $\alpha = 0.61$	

Annotation: n = sample size, M = mean, SD = standard deviation, r_{it} = discriminatory power.

(RQ4) Quantifying the qualitative findings into frequency scores for the difficulties students face and the different process-related decisions/actions during experimentation made it possible to test the associations between these and self-efficacy expectations and attitudes. (a) With respect to the first (qualitative) subtask, the ^anumber of difficulties mentioned in all superordinate categories, number of difficulties in the process-related superordinate categories, and number of process-related superordinate categories addressed were included in the analysis. There are no significant correlations between these knowledge-based expressions and attitudes toward the importance of diagnostics (e.g., ^aSpearman's $\rho = 0.005$, $p = 0.974$). The same is found in comparison to the domain-specific self-efficacy expectations. In this sample, the expression of self-efficacy expectations with respect to general as well as experimentation-related diagnostic activities is not related to the pre-service teachers' knowledge of student difficulties in experimentation. (b) For

the second (qualitative) subtask, the ^bnumber of process-related superordinate categories addressed in the mentioned decisions and the number of decisions mentioned in the process-related superordinate categories were included. These results also did not correlate with perceived domain-specific self-efficacy or ^battitudes towards the importance of diagnostics in university teacher education in the present sample (e.g., ^bSpearman's $\rho = 0.038$, $p = 0.790$).

4. Discussion

What began in the 1990s with a call for “Science for All” [148] has led to the setting of obligatory educational goals in countries’ and states’ curricula and standards regarding scientific inquiry and scientific reasoning as a component of scientific literacy (e.g., [4–6,149]). In this respect, school curricula in the natural sciences (e.g., NGSS [149], University location federal state 1 [150,151], University location federal state 2 [152,153]) require students to be able to conduct and reflect on scientific inquiry processes. Subject-specific content on promoting scientific reasoning in the classroom are also anchored in curricular standards for teacher education (e.g., [9,11] and thus must be taught in university teacher education. Appropriate teaching-learning concepts for scientific inquiry in university education can give future science teachers a better understanding of the difficulties and the (mis)understandings, alternative ideas or misconceptions students experience during experimentation, in order to be able to include those diagnosis-related aspects. Accordingly, there are lines of research focusing on (pre-service) teachers, e.g., measuring and assessing pre-service teachers scientific reasoning competencies in higher education [133,154,155], verification of validity [156], evaluation of translated versions [157], but these are rather limited and predominantly of a quantitative nature (e.g., [158]). Many studies also focus predominantly on developing and testing concepts and materials to promote subject-specific, pedagogical content knowledge and pedagogical knowledge related to scientific reasoning (e.g., video vignettes as a support for scientific reasoning, including video vignettes as a tool to promote students’ learning: e.g., [48,98,99,159]; seminar concepts: e.g., [160]). However, the first step is to identify pre-service teachers’ knowledge state in order to appropriately adapt university courses and to develop and/or employ alternative teaching approaches. Consequently, the present study did not focus on developing teaching-learning concepts for the subject didactic training of pre-service teachers; instead, the primary focus lies on potential prerequisites for learning such diagnostic activities, not only in the knowledge areas mentioned (CK; PCK), but also in terms of attitudes and self-efficacy regarding diagnostics and scientific inquiry (with a focus on experimentation). To summarize, it became clear that the pre-service teachers in the present study were only able to identify and cite a varying yet small number of potential student difficulties in the experimental problem-solving process from the comprehensive catalogue, presumably based on their own experience in school and basic training in subject didactics. The same was true of procedural knowledge in carrying out an investigation of an illustrative experimental phenomenon (“knowing how” [29,30]). In contrast, the teacher education students had a strong sense of the importance of diagnosing students’ experimentation skills and the hurdles students face, as well as a moderate level of self-efficacy in carrying out such diagnostic activities. The research questions underlying these findings can be answered and discussed in detail as follows:

(RQ1) The pre-service teachers cited difficulties students face in all steps of the experimentation process (see Table 1), but with very different frequencies (see Table 3). They predominantly described difficulties in the planning phase, including variable selection, followed by the implementation phase. Contrary to the usually higher self-assessment of scientific reasoning abilities [133], only a small number of unique difficulties were described with respect to evaluating evidence and most of all, formulating a research question and hypothesis, and the number of participants mentioning these areas was very low as well. With respect to the research question, this is possibly due to the structure of the task and teaching scenario applied in the study, in which students were provided with an

overarching research question (see Supplementary Material S1). Moreover, it can be concluded that pre-service teachers not only lack pedagogical (didactics) content knowledge about potential student difficulties in these phases, but also exhibit a lower level of content methodological (procedural) knowledge, e.g., with respect to formulating hypotheses. Students in the early semesters of university teacher education are still largely unfamiliar with the procedure and content of hypothesis formulation [155], which is at least partially due to the lack of or minimal opportunities to learn how to formulate research questions and hypotheses in school-based science instruction [42].

Overall, the pre-service teachers in the present study frequently attribute students' difficulties to the instructional setting, e.g., the open-ended nature of the task, working with the experimental materials, coming to an agreement and dividing up roles within the group during experimentation are most frequently mentioned. The participants' knowledge of PCK with regard to potential student difficulties and misconceptions hardly extends to aspects of subject-specific methodological concepts, such as planning an experiment that actually tests the hypothesis(es), the question of confounding variables or dealing with unexpected data. Overall, the respondents used rather general terminology, with scientific terminology [161] used only to a very small extent. Comparing the sub-samples, however, the pre-service chemistry teachers used technical terms more frequently, which may be due to the specific experiment selected in the example teaching unit (sugar's dissolution time) at the university location Braunschweig at the time of the survey. In contrast, it can be seen that the pre-service biology teachers cite more difficulties students face than the pre-service chemistry teachers. However, this difference is only due to a higher proportion of non-process-related categories mentioned by the pre-service biology teachers. As with the use of scientific terminology, this may be due to the number of subject didactics courses already completed by each group at the universities in which the study was conducted. Even though the study was conducted before any courses on scientific reasoning at both locations, the pre-service biology students in Kassel had already briefly dealt with this content area in their introductory lecture and tutorial.

(RQ2) With regard to key decisions (procedure for experimenting), the university pre-service science teachers tended to return to aspects they had also mentioned with respect to difficulties and misconceptions. The majority of respondents referred to decisions about working with and identifying variables in the experiment, planning, and actually conducting the experiment (including documentation). Interestingly, however, more than a third of the pre-service teachers also mentioned decisions about the research question or hypothesis, i.e., experimental steps that were underrepresented in the responses to the first subtask. It is possible that the pre-service teachers made greater reference here to hurdles had experienced themselves in their practical laboratory training at university. For example, one participant writes: "Now the students should not be thrown into the experiment like that. Forming hypotheses is often still a sticking point in tasks, even for university students, albeit at a higher level. [...] However, gaining knowledge should always be the goal of the experimental phase. Finally, in order to consolidate knowledge, the hypotheses must be verified or falsified based on the knowledge gained during the experiment." Results- and understanding-oriented statements are also made here, presumably due to the fictitious, open-ended experimental situation. With appropriate adaptive support (scaffolding), pre-service science teachers could be encouraged to further develop their process-oriented and results-oriented thinking patterns in this regard [162]. Considering the statements from both subtasks together, the pre-service biology teachers describe significantly more decisions in the process-related superordinate categories and also tend to describe more process-related decisions in the subcategories than the pre-service chemistry teachers. Nevertheless, subject-specific methodological (procedural) knowledge regarding scientific reasoning is rather low in the overall sample, similar to other studies in this area (e.g., [160]).

(RG3) The pre-service teachers rated diagnostics or diagnostic activities concerning experimentation as highly important for university teacher education, corresponding to the high perceived importance of diagnostics in teacher training more generally [163].

However, in-service teachers have a divergent view. Only around one-third of the in-service teachers surveyed by Lorenz [139] believed that diagnostic training is required to correctly assess students' competencies. It is possible that in-service teachers see their university studies in retrospect as focused on theoretical content and content related to their school subject, while diagnostics can only be performed when actually teaching in schools and thus can be learned in practice. Despite an increase in diagnostic knowledge during their studies compared to the beginning of their studies, pre-service teachers still rate it as below average compared to their knowledge of their subject and pedagogical knowledge [91,164]. In addition, diagnosing students' performance and difficulties in experimentation is more challenging than in other areas of science education, as it can only be assessed to a rather limited extent through written tasks. However, knowledge of the experimental competencies of the students being assessed—in particular, explicit knowledge of difficulties and misconceptions, which enables science teachers to assess student achievement against curricular expectations—is fundamental in science subjects [163]. This knowledge goes along with higher self-efficacy [124], which is in turn positively related to the successful transfer of content from teacher training (e.g., [165–167]) and ultimately to student achievement [168].

(RQ4a/b) This study's results concerning self-efficacy show that pre-service science teachers' subjective perceptions of their ability to successfully carry out general and experimentation-related diagnostic activities are significantly lower than their attitudes regarding the importance of diagnostics in teacher education. No correlations between knowledge of students' difficulties and key decisions in experimentation procedures and self-efficacy in diagnosing students' abilities in scientific reasoning acquisition were found in the sample studied here. In a study of 495 pre-service biology teachers [161], advanced students' (Master's or State Examination degree ≥ 7) self-efficacy to plan and conduct biology lessons correlated with the knowledge of what data to collect when assessing experimentation skills. No such correlation was found for pre-service teacher students at the undergraduate level. The present study examined pre-service teachers at an early stage of university teacher education, in the first third of their studies, which may be associated with a low level of knowledge in assessing self-efficacy and its associated dimensions. It is also possible that our findings were due to the surveyed students' low assessment of their own scientific inquiry competence. Conducting such surveys at the beginning of students' studies is therefore highly relevant in order to be able to promote such competences in a targeted way based on the obtained results. Khan and Krell [169] investigated the scientific reasoning competencies of pre-service Canadian science teachers and how they improved as a result of an intensive methods course including a 15-week internship. This intervention significantly improved the students' competencies in planning experiments and testing models, demonstrating that these are trainable through instruction on the scientific method including an internship with the opportunity to model, engage in, and reflect upon inquiry instruction in the science classroom. Further intervention studies demonstrate an increase in scientific reasoning competences as a result of combining different scaffolding formats [12,48,170]. In this context, it also seems interesting that, in addition to training (through scaffolds), greater background knowledge of scientific reasoning may influence pre-service teachers' competencies. Participants with prior university degrees (in other subjects) performed better in a multiple-choice questionnaire surveying scientific reasoning competencies than participants with no prior university degrees, perhaps because the former group could draw on greater background knowledge [169].

Limitations

Although the sample size was small, especially when it came to the biology and chemistry subsamples, and the findings require replication among physics education students and students at other universities, the study was able to provide insight on the views of pre-service biology and chemistry teachers in the first third of university teacher education on students' difficulties, misconceptions and key decisions in experimentation. This sample

was specifically selected because the pre-service science teachers were still at a relatively early phase of their teacher training, i.e., they were attending their first course on subject didactic training, and prior internships in their subject of study (biology/chemistry) and introductory lectures were not expected to have had much influence. The present study's sample therefore made it possible to assess the status quo of pre-service science teachers' professional knowledge and self-efficacy regarding scientific reasoning and diagnostics, in order to derive possible support measures.

At both university locations where the surveys took place, teaching education in the natural sciences focuses on scientific inquiry in research and teaching, and this context is likely to be significantly affected the students' specific content and pedagogical content knowledge. The second subject the pre-service teachers were studying might have also exerted an influence. For example, a measurement instrument for upper secondary students uncovered a distinction between "experimental ways of thinking and working" and content knowledge [170]. Pre-service teachers studying two science subjects have more opportunities to learn how to conduct scientific work, i.e., how to apply and reflect on scientific reasoning skills in different contexts and with different strategies [68], which leads to higher levels of competence [14,171].

The present survey of perceived difficulties and key decisions students face in experimentation had respondents refer to an example experimentation context (see Supplementary Material S1). While it would be possible to evaluate and carry out scientific reasoning skills in a context-free manner, it cannot be generally assumed that the same pre-service science teachers will provide the same responses on a task with the same format referring to a different context. Even though declarative knowledge of the context has only a minor influence on the successful completion of scientific reasoning test items, students with similar levels of procedural knowledge perform differently in such tasks (for an example regarding hypothesis-testing skills: [171]), which may be due to their different levels of declarative knowledge. Similarly, a different strategy can be used with each phenomenon or scientific problem to be discovered and researched, even in the same domain [68].

Another methodological limitation that deserves mention here concerns the scales measuring the students' attitudes and self-efficacy, which should be interpreted in comparison to other findings on these constructs. Neither follows the recommendation to apply scales with a large number of points [108], as they only contain four levels, potentially limiting the range of respondents' assessments. However, other studies have assessed self-efficacy with equally short Likert scales (e.g., [172]), while some studies have applied more comprehensive scales (e.g., [173]). In the future, it should be examined whether the number of scale points influences assessments of abilities and attitudes in a given area.

5. Conclusions

This study surveyed knowledge of students' difficulties, misconceptions, and key decisions in experimental problem-solving in the context of an authentic lesson plan, attitudes concerning the importance of diagnostics in teacher training, and self-efficacy related to diagnostics in experimental settings via an online questionnaire with a sample of pre-service biology and chemistry teachers. The developed instruments/tasks make it possible to survey pre-service teachers' knowledge and attitudes towards diagnosing experimentation skills within university teacher education in the subjects of biology and chemistry. They might be also transferable to physics education, and could serve as a foundation for conceptualizing university teaching-learning settings concerning scientific reasoning as well as for future intervention studies within subject didactics training regarding diagnostic knowledge of experimentation and attitudes. The results indicate that knowledge about students' misunderstandings, difficulties and problems during experimentation must be imparted to pre-service science teachers during university education. It is essential to understand scientific procedures (knowing how) and epistemic constructs (knowing why) regarding scientific reasoning in experimental problem-solving. Imparting this procedural and epistemic knowledge to teacher education at an early phase of their

studies via appropriate adaptive support (scaffolding) could help pre-service science teachers develop process-oriented and results-oriented thinking patterns as well as diagnostic skills in this regard [162]. To put this in perspective, knowledge of students' difficulties and frequent sticking points can help teachers design lessons in a student-oriented way (cf., [174]). Studies on subject-specific teaching quality in biology with respect to diagnostic competencies have shown that PCK and PK are statistically significantly related to pre-service teachers' diagnostic activities, and biology teachers' PCK is positively related to diagnostic accuracy [95]. Teachers are expected to provide students with learning opportunities that help them develop 21st-century skills, including core competencies in subject areas such as science. Zimmermann [19] concludes from her review of the literature on scientific reasoning that it is possible to teach both key features of science, i.e., the subject-specific content of scientific disciplines (e.g., biology, physics) and skills in experimentation and evaluating evidence. Previous studies on scientific reasoning show that progress has been made in research how it can be done to help students become scientifically literate adults by applying their scientific reasoning skills (cf., [19]). Therefore, it is necessary that (pre-service) science teachers are taught concepts and theories important for experimental inquiry processes, i.e., the key decisions that must be made based on concrete examples in open-ended experimentation process in order to conduct a successful experiment. In addition to knowledge about individual student learning characteristics that may be relevant to the learning process of scientific reasoning [175], procedural knowledge of scientific reasoning enables teachers to diagnose students' difficulties and misconceptions in the experimentation process. Further research using the instruments employed in this study on a larger sample of pre-service science teachers would contribute to the development of models of diagnostic competence acquisition (e.g., [176,177]) as well as professional vision (cf., [178]) for experimental problem-solving in competence-oriented science instruction and identify similarities and differences across subjects.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/educsci11100629/s1>, Task-Sheet S1: Tasks to analyse a classroom setting for experimentation—Difficulties and Approach of students.

Author Contributions: Conceptualization, D.H.-R. and M.M.; formal analysis, D.H.-R. and M.M.; investigation, D.H.-R., M.M., D.H.; writing—original draft D.H.-R.; writing—review and editing, D.H.-R., M.M., D.H., K.H.; project administration, D.H.-R., M.M.; funding acquisition, K.H., D.H.-R., M.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Bundesministerium für Bildung und Forschung (Federal Ministry of Education and Research), at the Technische Universität Braunschweig in the project Diagonal-NaWi, grant number: 01JA1909 and at the Universität Kassel in the project PRONET², grant number: 01JA1805. Both projects are part of the “Qualitätsinitiative Lehrerbildung”, a joint initiative of the Federal Government and the Länder which aims to improve the quality of teacher training. The authors are responsible for the content of this publication. Funding of APC: The authors acknowledge support by the Open Access Publication Funds of Technische Universität Braunschweig.

Institutional Review Board Statement: All participants were students at two German universities. They took part voluntarily and signed an informed consent form. Pseudonymization of participants was guaranteed during the study and the implementation took place online-based in a stress-free environment at home. Due to all these measures in the implementation of the study, an audit by an ethics committee was waived.

Informed Consent Statement: Written informed consent was obtained from all participants involved in the study.

Data Availability Statement: Information and queries on the data used can be obtained from the authors of this article.

Acknowledgments: We would like to thank Femke Sander for her support in rating as well all pre-service teachers who participated in the study. We also thank Di Fuccia (University of Kassel) for his support in recruiting pre-service teachers with chemistry as a subject. We thank the academic

editors, and two anonymous reviewers whose recommendations substantially improved the quality of the article.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Item parameters for the scale relevance of diagnostics in the teacher training program ($n = 50$).

Item Wording	<i>M</i> (<i>SD</i>)	r_{it}
A teacher can correctly assess students' performance in experimentation even without diagnostic components in teacher education.	2.92 (0.57)	0.55
It is important for teachers to be able to correctly assess the characteristics of the students that are relevant to learning and performance in experimentation.	3.42 (0.54)	0.51
Diagnostic skills of pre-service teachers in the assessment of experimental processes should be mandatorily promoted in teacher education.	3.30 (0.58)	0.59
Even without conducting diagnostic units in the course of study, a teacher can correctly assess the characteristics of students that are relevant to learning and performance in experimentation.	2.88 (0.52)	0.28
$M_{scale} = 3.13$; $SD_{scale} = 0.39$; Cronbach's $\alpha = 0.70$		

Appendix B

Table A2. Item analysis on domain-specific self-efficacy expectations with a pilot sample of $n = 98$.

Item Wording	Factor Loading (a_{jq})		h_2
	I	II	
Experiment-related diagnostic activities			
Even when I am under stress, I am still able to diagnose students' errors when experimenting in biology lessons.	0.818		0.697
I am confident in recognising children's specific difficulties in experimentation in biology despite great time pressure. *	0.763		0.663
In biology, I am able to identify the learning requirements of my students, even when I have little time. *	0.762		0.583
In biology, I am able to confidently understand my students' experimental skills, even when I have little time.	0.650	0.341	0.539
$n = 97$; $M_{scale} = 2.60$; $SD_{scale} = 0.44$; Cronbach's $\alpha = 0.79$			
General diagnostic activities			
In biology, I am able to integrate a diagnostic activity that accompanies learning in my teaching, even when I am under time pressure. *	0.549	0.418	0.476
Despite a high degree of heterogeneity, I am able to formulate tasks in biology with which I can appropriately test the level of knowledge of both weaker and stronger students. *		0.763	0.603
In biology, I am able to take into account the learning processes of the students when formulating individual learning objectives, even if these are very different. *		0.732	0.554
In the subject of biology, I manage to grasp the thought and work processes for gaining knowledge of my students, even when I have little time. *	0.319	0.710	0.606
$n = 95$; $M_{scale} = 2.46$; $SD_{scale} = 0.44$; Cronbach's $\alpha = 0.70$			
% of variance	46.03		12.99

Annotation: Principal component analysis with Varimax rotation; * adapted from [142].

References

- Vorholzer, A.; von Aufschnaiter, C.; Boone, W.J. Fostering Upper Secondary Students' Ability to Engage in Practices of Scientific Investigation: A Comparative Analysis of an Explicit and an Implicit Instructional Approach. *Res. Sci. Educ.* **2020**, *50*, 333–359. [CrossRef]
- NSTA/ASTE Standards for Science Teacher Preparation. 2020. Available online: <https://static.nsta.org/pdfs/2020NSTAStandards.pdf> (accessed on 2 August 2021).

3. Department for Education. Teachers' Standards. Guidance for School Leaders, School Staff and Governing Bodies. 2011. Available online: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/665520/Teachers_Standards.pdf (accessed on 2 August 2021).
4. Beschlüsse der Kultusministerkonferenz. Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss. Beschluss vom 16.12.2004. Available online: https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2004/2004_12_16-Bildungsstandards-Biologie.pdf (accessed on 2 August 2021).
5. Beschlüsse der Kultusministerkonferenz. Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss. Beschluss vom 16.12.2004. Available online: https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2004/2004_12_16-Bildungsstandards-Chemie.pdf (accessed on 2 August 2021).
6. Beschlüsse der Kultusministerkonferenz. Bildungsstandards im Fach Physik für den Mittleren Schulabschluss. Beschluss vom 16.12.2004. Available online: https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2004/2004_12_16-Bildungsstandards-Physik-Mittleren-SA.pdf (accessed on 2 August 2021).
7. Wellnitz, N.; Hecht, M.; Heitmann, P.; Kauertz, A.; Mayer, J.; Sumfleth, E.; Walpuski, M. Modellierung des Kompetenzteilbereichs Naturwissenschaftliche Untersuchungen. *Z. Erzieh.* **2017**, *20*, 556–584. [CrossRef]
8. Darling-Hammond, L. Teacher Quality and Student Achievement. *Educ. Policy Anal. Arch.* **2000**, *8*, 1. [CrossRef]
9. Morrell, P.D.; Park Rogers, M.A.; Pyle, E.J.; Roehrig, G.; Veal, W.R. Preparing Teachers of Science for 2020 and Beyond: Highlighting Changes to the NSTA/ASTE Standards for Science Teacher Preparation. *J. Sci. Teach. Educ.* **2020**, *31*, 1–7. [CrossRef]
10. Committee on the Development of an Addendum to the National Science Education Standards on Scientific Inquiry; Board on Science Education; Division of Behavioral and Social Sciences and Education; National Research Council. *Inquiry and the National Science Education Standards: A Guide for Teaching and Learning*; Olson, S., Loucks-Horsley, S., Eds.; National Academies Press: Washington, DC, USA, 2000; p. 9596, ISBN 9780309064767.
11. Kultusministerkonferenz. Ländergemeinsame Inhaltliche Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung. Beschluss der Kultusministerkonferenz vom 16.10.2008 i. d. F. vom 16.05.2019. Available online: https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2008/2008_10_16-Fachprofile-Lehrerbildung.pdf (accessed on 2 August 2021).
12. Bruckermann, T.; Aschermann, E.; Bresges, A.; Schlüter, K. Metacognitive and Multimedia Support of Experiments in Inquiry Learning for Science Teacher Preparation. *Int. J. Sci. Educ.* **2017**, *39*, 701–722. [CrossRef]
13. Hodson, D. Learning Science, Learning about Science, Doing Science: Different Goals Demand Different Learning Methods. *Int. J. Sci. Educ.* **2014**, *36*, 2534–2553. [CrossRef]
14. Hartmann, S.; Upmeier zu Belzen, A.; Krüger, D.; Pant, H.A. Scientific Reasoning in Higher Education: Constructing and Evaluating the Criterion-Related Validity of an Assessment of Preservice Science Teachers' Competencies. *Z. Psychol.* **2015**, *223*, 47–53. [CrossRef]
15. Mathesius, S.; Hartmann, S.; Upmeier zu Belzen, A.; Krüger, D. Scientific Reasoning as an Aspect of Preservice Biology Teacher Education. Assessing competencies using a paper-pencil test. In *The Future of Biology Education Research*; Tal, T., Yarden, A., Eds.; The Technion, Israel Institute of Technology/The Weizmann Institute of Science: Haifa, Israel, 2016; pp. 93–110.
16. Bernholt, S.; Neumann, K.; Nentwig, P. *Making It Tangible: Learning Outcomes in Science Education*; Waxmann: Münster, Germany, 2012; pp. 13–28.
17. Opitz, A.; Heene, M.; Fischer, F. Measuring Scientific Reasoning—A Review of Test Instruments. *Educ. Res. Eval.* **2017**, *23*, 78–101. [CrossRef]
18. Koslowski, B. *Theory and Evidence: The Development of Scientific Reasoning*; MIT Press: Cambridge, MA, USA, 1996.
19. Zimmerman, C. The Development of Scientific Reasoning Skills. *Dev. Rev.* **2000**, *20*, 99–149. [CrossRef]
20. Paul, R.; Elder, L. *The Thinker's Guide to Scientific Thinking: Based on Critical Thinking Concepts & Principles*; Thinker's Guide Library: Tomales, CA, USA, 2012.
21. Kuhn, D.; Amsel, E.; O'Loughlin, M. The Development of Scientific Thinking Skills. In *Developmental Psychology Series*; Academic Press: San Diego, CA, USA, 1988.
22. Klahr, D. Exploring Science. In *The Cognition and Development of Discovery Processes*; MIT Press: Cambridge, MA, USA, 2000.
23. Schafersman, S.D. An Introduction to Science: Scientific Thinking and the Scientific Method. Available online: <http://www.geosunysb.edu/esp/files/scientific-method.html> (accessed on 3 August 2021).
24. Janoušková, S.; Pyskatá Rathouská, L.; Žák, V.; Urválková, E.S. The Scientific Thinking and Reasoning Framework and Its Applicability to Manufacturing and Services Firms in Natural Sciences. *Res. Sci. Technol. Educ.* **2021**, 1–22. [CrossRef]
25. Dowd, J.E.; Thompson, R.J.; Schiff, L.A.; Reynolds, J.A. Understanding the Complex Relationship between Critical Thinking and Science Reasoning among Undergraduate Thesis Writers. *Life Sci. Educ.* **2018**, *17*, ar4. [CrossRef]
26. Kuhn, D. *The Skills of Argument*; Cambridge University Press: Cambridge, UK, 1991.
27. Fischer, F.; Kollar, I.; Ufer, S.; Sodian, B.; Hussmann, H.; Pekrun, R.; Neuhaus, B.; Dorner, B.; Pankofer, S.; Fischer, M.; et al. Scientific Reasoning and Argumentation: Advancing an Interdisciplinary Research Agenda in Education. *Frontline Learn. Res.* **2014**, *2*, 28–45. [CrossRef]
28. Dunbar, K.; Fugelsang, J. Scientific Thinking and Reasoning. In *The Cambridge Handbook of Thinking and Reasoning*; Holyoak, K.J., Morrison, R.G., Eds.; Cambridge University Press: Cambridge, UK, 2005; pp. 705–725.

29. Krell, M.; Redman, C.; Mathesius, S.; Krüger, D.; van Driel, J. Assessing Pre-Service Science Teachers' Scientific Reasoning Competencies. *Res. Sci. Educ.* **2018**, *50*, 2305–2329. [[CrossRef](#)]
30. Lawson, A.E. The Nature and Development of Scientific Reasoning: A Synthetic View. *Int. J. Sci. Math. Educ.* **2004**, *2*, 307–338. [[CrossRef](#)]
31. Morris, B.J.; Croker, S.; Masnick, A.; Zimmerm, C. The Emergence of Scientific Reasoning. In *Current Topics in Children's Learning and Cognition*; Kloos, H., Ed.; IntechOpen: London, UK, 2012; pp. 61–82.
32. Kind, P.; Osborne, J. Styles of Scientific Reasoning: A Cultural Rationale for Science Education? *Sci. Ed.* **2017**, *101*, 8–31. [[CrossRef](#)]
33. Arnold, J.C.; Mühlhng, A.; Kremer, K. Exploring Core Ideas of Procedural Understanding in Scientific Inquiry Using Educational Data Mining. *Res. Sci. Technol. Educ.* **2021**, 1–21. [[CrossRef](#)]
34. Mayer, J. Erkenntnisgewinnung als Wissenschaftliches Problemlösen. In *Theorien in der Biologiedidaktischen Forschung: Ein Handbuch für Lehramtsstudenten und Doktoranden*; Krüger, D., Vogt, H., Eds.; Springer: Berlin, Germany; New York, NY, USA, 2007; pp. 177–186.
35. Baur, A.; Emden, M. How to Open Inquiry Teaching? An Alternative Teaching Scaffold to Foster Students' Inquiry Skills. *Chem. Teach. Int.* **2021**, *3*, 20190013. [[CrossRef](#)]
36. Rieß, W.; Wirtz, M.A.; Barzel, B.; Schulz, A. Integration der theoretischen und empirischen Befunde zum Experimentieren im mathematisch-naturwissenschaftlichen Unterricht. In *Experimentieren im Mathematisch-Naturwissenschaftlichen Unterricht: Schüler*; Waxmann Verlag: Münster, Germany, 2012; pp. 353–364.
37. Andersen, C.; Garcia-Mila, M. Scientific Reasoning During Inquiry. In *Science Education*; Taber, K.S., Akpan, B., Eds.; SensePublishers: Rotterdam, The Netherlands, 2017; pp. 105–117.
38. Arnold, J.C.; Boone, W.J.; Kremer, K.; Mayer, J. Assessment of Competencies in Scientific Inquiry through the Application of Rasch Measurement Techniques. *Educ. Sci.* **2018**, *8*, 184. [[CrossRef](#)]
39. Meier, M. *Entwicklung und Prüfung eines Instrumentes zur Diagnose der Experimentierkompetenz von Schülerinnen und Schülern*; Logos Verlag: Berlin, Germany, 2016.
40. Hammann, M.; Phan, T.T.H.; Bayrhuber, H. Experimentieren als Problemlösen: Lässt sich das SDDS-Modell Nutzen, um Unterschiedliche Dimensionen beim Experimentieren zu messen? In *Kompetenzdiagnostik*; Prenzel, M., Gogolin, I., Krüger, H.-H., Eds.; Springer-Verlag GmbH: Wiesbaden, Germany, 2007; pp. 33–50.
41. Nowak, K.H.; Nehring, A.; Tiemann, R.; Upmeier zu Belzen, A. Assessing Students' Abilities in Processes of Scientific Inquiry in Biology Using a Paper-and-Pencil Test. *J. Biol. Educ.* **2013**, *47*, 182–188. [[CrossRef](#)]
42. Nehring, A.; Nowak, K.H.; zu Belzen, A.U.; Tiemann, R. Predicting Students' Skills in the Context of Scientific Inquiry with Cognitive, Motivational, and Sociodemographic Variables. *Int. J. Sci. Educ.* **2015**, *37*, 1343–1363. [[CrossRef](#)]
43. Nehring, A.; Stiller, J.; Nowak, K.H.; Upmeier zu Belzen, A.; Tiemann, R. Naturwissenschaftliche Denk- und Arbeitsweisen im Chemieunterricht-eine Modellbasierte Videostudie zu Lerngelegenheiten für den Kompetenzbereich der Erkenntnisgewinnung. *ZfDN* **2016**, *22*, 77–96. [[CrossRef](#)]
44. Kraeva, L. *Problemlösestrategien von Schülerinnen und Schülern Diagnostizieren*; Logos Verlag: Berlin, Germany, 2020.
45. Neumann, I. *Beyond Physics Content Knowledge: Modeling Competence Regarding Nature of Scientific Inquiry and Nature of Scientific Knowledge*; Logos Verlag: Berlin, Germany, 2011.
46. Maiseyenko, V.; Schecker, H.; Nawrath, D. Kompetenzorientierung des Naturwissenschaftlichen Unterrichts-Symbiotische Kooperation bei der Entwicklung eines Modells Experimenteller Kompetenz. *PhyDid A Phys. Didakt. Sch. Hochsch.* **2013**, *1*, 1–17.
47. Nababan, N.P.; Nasution, D.; Jayanti, R.D. The Effect of Scientific Inquiry Learning Model and Scientific Argumentation on The Students' Science Process Skill. *J. Phys. Conf. Ser.* **2019**, *1155*, 012064. [[CrossRef](#)]
48. Hilfert-Rüppell, D.; Eghtessad, A.; Höner, K. Interaktive Videovignetten aus naturwissenschaftlichem Unterricht—Förderung der Diagnosekompetenz von Lehramtsstudierenden hinsichtlich der Experimentierfähigkeit von Schülerinnen und Schülern. *Medien Pädagogik* **2018**, *31*, 125–142. [[CrossRef](#)]
49. Alfieri, L.; Brooks, P.J.; Aldrich, N.J.; Tenenbaum, H.R. Does Discovery-Based Instruction Enhance Learning? *J. Educ. Psychol.* **2011**, *103*, 1–18. [[CrossRef](#)]
50. Abrams, E.; Southerland, S.A.; Evans, C. Inquiry in the Classroom: Identifying Necessary Components of a Useful Definition. In *Inquiry in the Science Classroom: Challenges and Opportunities*; Abrams, E., Southerland, S., Silva, P., Eds.; Information Age Publishing: Charlotte, NC, USA, 2008; pp. 11–42.
51. Barzel, B.; Reinhoffer, B.; Schrenk, B. Das Experimentieren im Unterricht. In *Experimentieren im Mathematisch-Naturwissenschaftlichen Unterricht: Schüler Lernen Wissenschaftlich Denken und Arbeiten*; Rieß, W., Wirtz, M.A., Barzel, B., Schulz, A., Eds.; Waxmann Verlag: Münster, Germany, 2012; pp. 103–128.
52. Ural, E. The Effect of Guided-Inquiry Laboratory Experiments on Science Education Students' Chemistry Laboratory Attitudes, Anxiety and Achievement. *J. Educ. Train. Stud.* **2016**, *4*, 217–227. [[CrossRef](#)]
53. Hmelo-Silver, C.E.; Duncan, R.G.; Chinn, C.A. Scaffolding and Achievement in Problem-Based and Inquiry Learning: A Response to Kirschner, Sweller, and Clark (2006). *Educ. Psychol.* **2007**, *42*, 99–107. [[CrossRef](#)]
54. Vorholzer, A.; von Aufschnaiter, C. Guidance in Inquiry-Based Instruction-An Attempt to Disentangle a Manifold Construct. *Int. J. Sci. Educ.* **2019**, *41*, 1562–1577. [[CrossRef](#)]
55. Höner, K.; Eghtessad, A.; Hilfert-Rüppell, D.; Kraeva, L. Naturwissenschaftliches Potenzial?—Diagnose von Schülerfähigkeiten zum experimentellen Problemlösen. *J. Für Begabtenförderung* **2017**, *2*, 8–23.

56. Hammann, M.; Phan, T.T.H.; Ehmer, M.; Grimm, T. Assessing Pupils' Skills in Experimentation. *J. Biol. Exp.* **2008**, *42*, 66–72. [[CrossRef](#)]
57. National Academy of Engineering and National Research Council. *STEM Integration in K-12 Education: Status, Prospects, and an Agenda for Research*; National Academies Press: Washington, DC, USA, 2014.
58. Bell, R.L.; Blair, L.M.; Crawford, B.A.; Lederman, N.G. Just Do It? Impact of a Science Apprenticeship Program on High School Students' Understandings of the Nature of Science and Scientific Inquiry. *J. Res. Sci. Teach.* **2003**, *40*, 487–509. [[CrossRef](#)]
59. Klahr, D.; Nigam, M. The Equivalence of Learning Paths in Early Science Instruction: Effect of Direct Instruction and Discovery Learning. *Psychol. Sci.* **2004**, *15*, 661–667. [[CrossRef](#)]
60. Kirschner, P.A.; Sweller, J.; Clark, R.E. Why Minimal Guidance During Instruction Does Not Work: An Analysis of the Failure of Constructivist, Discovery, Problem-Based, Experiential, and Inquiry-Based Teaching. *Educ. Psychol.* **2006**, *41*, 75–86. [[CrossRef](#)]
61. Dunbar, K.; Klahr, D. Developmental Differences in Scientific Discovery Processes. In *Complex Information Processing: The Impact of Herbert, A. Simon*; Lawrence Erlbaum Associates, Inc.: New Jersey, NJ, USA, 1989; pp. 109–143.
62. Chinn, C.A.; Brewer, W.F. The Role of Anomalous Data in Knowledge Acquisition: A Theoretical Framework and Implications for Science Instruction. *Rev. Educ. Res.* **1993**, *63*, 1–49. [[CrossRef](#)]
63. Hammann, M.; Phan, T.T.H.; Ehmer, M.; Bayrhuber, H. Fehlerfrei Experimentieren. *Math. Nat. Unterr.* **2006**, *59*, 292–299.
64. Klahr, D.; Fay, A.L.; Dunbar, K. Heuristics for Scientific Experimentation: A Developmental Study. *Cogn. Psychol.* **1993**, *25*, 111–146. [[CrossRef](#)]
65. Lubben, F.; Millar, R. Children's Ideas about the Reliability of Experimental Data. *Int. J. Sci. Educ.* **1996**, *18*, 955–968. [[CrossRef](#)]
66. Baur, A. Fehler, Fehlkonzepte und Spezifische Vorgehensweisen von Schülerinnen und Schülern beim Experimentieren: Ergebnisse einer Videogestützten Beobachtung. *ZfDN* **2018**, *24*, 115–129. [[CrossRef](#)]
67. Hilfert-Rüppell, D.; Höner, K. *Diagnosing Student's Scientific Reasoning Skills Based on Authentic Videos of Inquiry Experiments (Working Title)*; TU Braunschweig: Braunschweig, Germany, 2021.
68. Glaser, R.; Schauble, L.; Raghavan, K.; Zeitz, C. Scientific Reasoning Across Different Domains. In *Computer-Based Learning Environments and Problem Solving*; De Corte, E., Linn, M.C., Mandl, H., Verschaffel, L., Eds.; Springer: Berlin/Heidelberg, Germany, 1992; pp. 345–371.
69. Germann, P.J.; Aram, R.; Burke, G. Identifying Patterns and Relationships among the Responses of Seventh-Grade Students to the Science Process Skill of Designing Experiments. *J. Res. Sci. Teach.* **1996**, *33*, 79–99. [[CrossRef](#)]
70. Chen, Z.; Klahr, D. All Other Things Being Equal: Acquisition and Transfer of the Control of Variables Strategy. *Child Dev.* **1999**, *70*, 1098–1120. [[CrossRef](#)] [[PubMed](#)]
71. Kirchner, S.; Priemer, B. Probleme von Schülern Mit Offenen Experimentieraufgaben in Physik. In *Naturwissenschaftlicher Unterricht im internationalen Vergleich*; LIT Verlag: Münster, Germany, 2007; pp. 346–348.
72. Wahser, I.; Sumfleth, E. Training Experimenteller Arbeitsweisen Zur Unterstützung Kooperativer Kleingruppenarbeit Im Fach Chemie. *ZfDN* **2008**, *14*, 219–241.
73. Meier, M.; Mayer, J. Experimentierkompetenz praktisch erfassen—Entwicklung und Validierung eines anwendungsbezogenen Aufgabendesigns. In *Lehr- und Lernforschung in der Biologiedidaktik*; Harms, U., Bogner, F.X., Eds.; StudienVerlag: Innsbruck, Austria, 2012; Volume 5, pp. 81–98.
74. Meier, M.; Mayer, J. Selbständiges Experimentieren: Entwicklung und Einsatz eines anwendungsbezogenen Aufgabendesigns. *MNU* **2014**, *67*, 4–10.
75. Kechel, J.-H.; Wodzinski, R. Schülerschwierigkeiten beim Experimentieren zum Hooke'schen Gesetz. In *Authentizität und Lernen-das Fach in der Fachdidaktik*; Universität Regensburg: Regensburg, Germany, 2016; pp. 170–172.
76. Nerdel, C. Kompetenzorientiert und aufgabenbasiert für Schule und Hochschule. In *Grundlagen der Naturwissenschaftsdidaktik*, 1st ed.; Springer: Berlin/Heidelberg, Germany, 2017.
77. Carey, S.; Evans, R.; Honda, M.; Jay, E.; Unger, C. 'An Experiment Is When You Try It and See If It Works': A Study of Grade 7 Students' Understanding of the Construction of Scientific Knowledge. *Int. J. Sci. Educ.* **1989**, *11*, 514–529. [[CrossRef](#)]
78. Schauble, L.; Klopfer, L.E.; Raghavan, K. Students' Transition from an Engineering Model to a Science Model of Experimentation. *J. Res. Sci. Teach.* **1991**, *28*, 859–882. [[CrossRef](#)]
79. Chinn, C.A.; Brewer, W.F. An Empirical Test of a Taxonomy of Responses to Anomalous Data in Science. *J. Res. Sci. Teach.* **1998**, *35*, 623–654. [[CrossRef](#)]
80. Ludwig, T.; Priemer, B. Begründungen und Überzeugungen beim Beibehalten und Verwerfen von eigenen Hypothesen in Real- und Simulationsexperimenten. In *Konzepte Fachdidaktischer Strukturierung für den Unterricht*; Bernholt, S., Ed.; LIT Verlag: Münster, Germany, 2012; pp. 313–315.
81. De Jong, T.; Van Joolingen, W.R. Scientific Discovery Learning with Computer Simulations of Conceptual Domains. *Rev. Educ. Res.* **1998**, *68*, 179–201. [[CrossRef](#)]
82. Arnold, J.; Kremer, K.; Mayer, J. Understandig Students' Experiments—What Kind of Support Do They Need in Inquiry Tasks? *Int. J. Sci. Educ.* **2014**, *36*, 2719–2749. [[CrossRef](#)]
83. Arnold, J.; Kremer, K.; Mayer, J. Schüler als Forscher. Experimentieren kompetenzorientiert Unterrichten und Beurteilen. *MNU* **2014**, *67*, 83–91.
84. Hammann, M. Kompetenzentwicklungsmodelle: Merkmale und ihre Bedeutung - dargestellt anhand von Kompetenzen beim Experimentieren. *MNU* **2004**, *57*, 196–203.

85. Dunbar, K. Concept Discovery in a Scientific Domain. *Cogn. Sci.* **1993**, *17*, 397–434. [[CrossRef](#)]
86. Härtig, H.; Neumann, K.; Erb, R. Experimentieren als Interaktion von Situation und Person: Ergebnisse einer Expertenbefragung. *ZfjDN* **2017**, *23*, 71–80. [[CrossRef](#)]
87. Cappell, J. *Fachspezifische Diagnosekompetenz Angehender Physiklehrkräfte in der Ersten Ausbildungsphase*; Studien zum Physik- und Chemielernen; Logos Verlag: Berlin, Germany, 2013.
88. Bögeholz, S.; Joachim, C.; Hasse, S.; Hammann, M. Kompetenzen von (angehenden) Biologielehrkräften Zur Beurteilung von Experimentierkompetenzen. *Unterrichtswissenschaft* **2016**, *44*, 40–54.
89. Draude, M. *Die Kompetenz von Physiklehrkräften, Schwierigkeiten von Schülerinnen und Schülern beim Eigenständigen Experimentieren zu Diagnostizieren*; Logos Verlag Berlin: Berlin, Germany, 2016.
90. Kechel, J.-H. *Schülerschwierigkeiten beim Eigenständigen Experimentieren: Eine Qualitative Studie am Beispiel einer Experimentieraufgabe zum Hooke'schen Gesetz*; Logos Verlag Berlin: Berlin, Germany, 2016.
91. Dübbelde, G. *Diagnostische Kompetenzen Angehender Biologie-Lehrkräfte im Bereich der Naturwissenschaftlichen Erkenntnisgewinnung*; Universität Kassel: Kassel, Germany, 2013. Available online: <https://kobra.uni-kassel.de/handle/123456789/2013122044701> (accessed on 3 August 2021).
92. Beretz, A.-K.; Lengnink, K.; Aufschnaiter, C. Diagnostische Kompetenz gezielt Fördern-Videoeinsatz im Lehramtsstudium Mathematik und Physik. In *Diagnose und Förderung Heterogener Lerngruppen: Theorien, Konzepte und Beispiele aus der MINT-Lehrerbildung*; Selter, C., Michaelis, J., Lengnink, K., Knipping, C., Hößle, C., Hußmann, S., Eds.; Waxmann: Münster, Germany; New York, NY, USA, 2017; pp. 149–168.
93. Schrader, J. Theorie und Praxis der Erwachsenenbildung. In *Struktur und Wandel der Weiterbildung*; Bertelsmann: Bielefeld, Germany, 2011.
94. Chernikova, O.; Heitzmann, N.; Fink, M.C.; Timothy, V.; Seidel, T.; Fischer, F. Facilitating Diagnostic Competences in Higher Education-A Meta-Analysis in Medical and Teacher Education. *Educ. Psychol. Rev.* **2020**, *32*, 157–196. [[CrossRef](#)]
95. Kramer, M.; Förtsch, C.; Boone, W.J.; Seidel, T.; Neuhaus, B.J. Investigating Pre-Service Biology Teachers' Diagnostic Competences: Relationships between Professional Knowledge, Diagnostic Activities, and Diagnostic Accuracy. *Educ. Sci.* **2021**, *11*, 89. [[CrossRef](#)]
96. Shulman, L.S. Knowledge and Teaching. Foundations of the New Reform. *Harv. Educ. Rev.* **1987**, *57*, 1–22. [[CrossRef](#)]
97. Kunter, M.; Klusmann, U.; Baumert, J. Professionelle Kompetenz von Mathematiklehrkräften: Das COACTIV-Modell. In *Lehrprofessionalität. Bedingungen, Genese, Wirkungen und ihre Messung*; Zlatkin-Troitschanskaia, O., Beck, K., Sembill, D., Nickolaus, R., Mulder, R., Eds.; Beltz: Weinheim, Germany, 2009; pp. 153–165.
98. Barth, V.L.; Piwowar, V.; Kumschick, I.R.; Ophardt, D.; Thiel, F. The Impact of Direct Instruction in a Problem-Based Learning Setting. Effects of a Video-Based Training Program to Foster Preservice Teachers' Professional Vision of Critical Incidents in the Classroom. *Int. J. Educ. Res.* **2019**, *95*, 1–12. [[CrossRef](#)]
99. Kramer, M.; Förtsch, C.; Stürmer, J.; Förtsch, S.; Seidel, T.; Neuhaus, B.J. Measuring Biology Teachers' Professional Vision: Development and Validation of a Video-Based Assessment Tool. *Cogent. Educ.* **2020**, *7*, 1823155. [[CrossRef](#)]
100. Kramer, M.; Förtsch, C.; Neuhaus, B.J. Can Pre-Service Biology Teachers' Professional Knowledge and Diagnostic Activities Be Fostered by Self-Directed Knowledge Acquisition via Texts? *Educ. Sci.* **2021**, *11*, 244. [[CrossRef](#)]
101. Blömeke, S.; Gustafsson, J.-E.; Shavelson, R.J. Beyond Dichotomies: Competence Viewed as a Continuum. *Z. Psychol.* **2015**, *223*, 3–13. [[CrossRef](#)]
102. Seidel, T.; Stürmer, J. Modeling and Measuring the Structure of Professional Vision in Preservice Teachers. *Am. Educ. Res. J.* **2014**, *51*, 739–771. [[CrossRef](#)]
103. Blomberg, G.; Sherin, M.G.; Renkl, A.; Glogger, I.; Seidel, T. Understanding Video as a Tool for Teacher Education: Investigating Instructional Strategies to Promote Reflection. *Instr. Sci.* **2014**, *42*, 443–463. [[CrossRef](#)]
104. Meissel, K.; Meyer, F.; Yao, E.S.; Rubie-Davies, C.M. Subjectivity of Teacher Judgments: Exploring Student Characteristics That Influence Teacher Judgments of Student Ability. *Teach. Teach. Educ.* **2017**, *65*, 48–60. [[CrossRef](#)]
105. Baumert, J.; Kunter, M. Das Kompetenzmodell von COACTIV. In *Professionelle Kompetenz von Lehrkräften: Ergebnisse des Forschungsprogramms COACTIV*; Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S., Neubrand, M., Eds.; Waxmann: Münster, Germany, 2011; pp. 29–53.
106. Bandura, A. *Social Foundations of Thought and Action: A Social Cognitive Theory*; Prentice-Hall, Inc.: Englewood Cliffs, NJ, USA, 1986.
107. Bong, M.; Skaalvik, E.M. Academic Self-Concept and Self-Efficacy: How Different Are They Really? *Educ. Psychol. Rev.* **2003**, *15*, 1–40. [[CrossRef](#)]
108. Bandura, A.; Freeman, W.H.; Lightsey, R. Self-Efficacy: The Exercise of Control. *J. Cogn. Psychother.* **1999**, *13*, 158–166. [[CrossRef](#)]
109. Bandura, A. Guide for Constructing Self-Efficacy Scales. In *Self-Efficacy Beliefs of Adolescents*; Pajares, F., Urdan, T., Eds.; Information Age Publishing: Charlotte, NC, USA, 2006; Volume 5, pp. 307–337.
110. Pajares, F. Self-Efficacy Beliefs in Academic Settings. *Rev. Educ. Res.* **1996**, *66*, 543–578. [[CrossRef](#)]
111. Andrew, S. Self-Efficacy as a Predictor of Academic Performance in Science. *J. Adv. Nurs.* **1998**, *27*, 596–603. [[CrossRef](#)] [[PubMed](#)]
112. Schunk, D.H.; Pajares, F. Self-Efficacy Theory. In *Handbook of Motivation at School*; Wenzel, K.R., Wigfield, A., Eds.; Routledge/Taylor & Francis Group: New York, NY, USA, 2009; pp. 35–53.
113. Komarraju, M.; Nadler, D. Self-Efficacy and Academic Achievement: Why Do Implicit Beliefs, Goals, and Effort Regulation Matter? *Learn. Individ. Differ.* **2013**, *25*, 67–72. [[CrossRef](#)]

114. Handtke, K.; Bögeholz, S. Self-Efficacy Beliefs of Interdisciplinary Science Teaching (Self-ST) Instrument: Drafting a Theory-Based Measurement. *Educ. Sci.* **2019**, *9*, 247. [[CrossRef](#)]
115. Hackett, G.; Betz, N.E. An Exploration of the Mathematics Self-Efficacy/Mathematics Performance Correspondence. *J. Res. Math. Educ.* **1989**, *20*, 261–273. [[CrossRef](#)]
116. Pajares, F.; Kranzler, J. Self-Efficacy Beliefs and General Mental Ability in Mathematical Problem-Solving. *Contemp. Educ. Psychol.* **1995**, *20*, 426–443. [[CrossRef](#)]
117. Pajares, F.; Graham, L. Self-Efficacy, Motivation Constructs, and Mathematics Performance of Entering Middle School Students. *Contemp. Educ. Psychol.* **1999**, *24*, 124–139. [[CrossRef](#)]
118. Fencil, H.; Scheel, K. Engaging Students: An Examination of the Effects of Teaching Strategies on Self-Efficacy and Course Climate in a Nonmajors Physics Course. *J. Coll. Sci. Teach.* **2005**, *35*, 20.
119. Dalgety, J.; Coll, R.K. Exploring First-Year Science Students' Chemistry Self-Efficacy. *Int. J. Sci. Math. Educ.* **2006**, *4*, 97–116. [[CrossRef](#)]
120. Yürük, N. The Predicators of Pre-Service Elementary Teachers' Anxiety about Teaching Science. *J. Balt. Sci. Educ.* **2011**, *10*, 17–26.
121. Kurbanoglu, N.I.; Akim, A. The Relationships between University Students' Chemistry Laboratory Anxiety, Attitudes, and Self-Efficacy Beliefs. *Aust. J. Teach. Educ.* **2010**, *35*, 48–59. [[CrossRef](#)]
122. Trujillo, G.; Tanner, K.D. Considering the Role of Affect in Learning: Monitoring Students' Self-Efficacy, Sense of Belonging, and Science Identity. *Life Sci. Educ.* **2014**, *13*, 6–15. [[CrossRef](#)] [[PubMed](#)]
123. Riese, J.; Reinhold, P. Empirische Erkenntnisse zur Struktur Professioneller Handlungskompetenz von angehenden Physik-lehrkräften. *ZfDN* **2010**, *16*, 167–187.
124. Mahler, D.; Großschedl, J.; Harms, U. Opportunities to Learn for Teachers' Self-Efficacy and Enthusiasm. *Educ. Res. Int.* **2017**, *2017*, 1–17. [[CrossRef](#)]
125. Klug, J.; Bruder, S.; Schmitz, B. Which Variables Predict Teachers Diagnostic Competence When Diagnosing Students' Learning Behavior at Different Stages of a Teacher's Career? *Teach. Teach.* **2016**, *22*, 461–484. [[CrossRef](#)]
126. Holzberger, D.; Philipp, A.; Kunter, M. How Teachers' Self-Efficacy Is Related to Instructional Quality: A Longitudinal Analysis. *J. Educ. Psychol.* **2013**, *105*, 774–786. [[CrossRef](#)]
127. Blömeke, S. Qualitativ-Quantitativ, Induktiv-Deduktiv, Prozess-Produkt, National-International. In *Forschung zur Lehrerbildung*; Lüders, M., Ed.; Waxmann Verlag: Münster, Germany, 2007; pp. 13–36.
128. Barke, H.-D.; Hazari, A.; Yitbarek, S. *Misconceptions in Chemistry: Addressing Perceptions in Chemical Education*; Springer: Berlin/Heidelberg, Germany, 2009.
129. Tolsdorf, Y.; Markic, S. Exploring Chemistry Student Teachers' Diagnostic Competence—A Qualitative Cross-Level Study. *Educ. Sci.* **2017**, *7*, 86. [[CrossRef](#)]
130. Tolsdorf, Y.; Markic, S. Development of an Instrument and Evaluation Pattern for the Analysis of Chemistry Student Teachers' Diagnostic Competence. *Int. J. Phys. Chem. Educ.* **2017**, *9*, 1–10.
131. Capizzi, A.M.; Fuchs, L.S. Effects of Curriculum-Based Measurement with and without Diagnostic Feedback on Teacher Planning. *Remedial Spec. Educ.* **2005**, *26*, 159–174. [[CrossRef](#)]
132. Cramer, C. Beurteilung Des Bildungswissenschaftlichen Studiums Durch Lehramtsstudierende in Der Ersten Ausbildungsphase Im Längsschnitt. *Z. Pädagogik* **2013**, *59*, 66–82.
133. Bıcak, B.E.; Borchert, C.E.; Höner, K. Measuring and Fostering Preservice Chemistry Teachers' Scientific Reasoning Competency. *Educ. Sci.* **2021**, *11*, 496. [[CrossRef](#)]
134. Osborne, J. The 21st Century Challenge for Science Education: Assessing Scientific Reasoning. *Think. Ski. Creat.* **2013**, *10*, 265–279. [[CrossRef](#)]
135. Shavelson, R.J. Assessing Business-Planning Competence using the Collegiate Learning Assessment as a Prototype. *Empir. Res. Vocat. Educ. Train.* **2012**, *4*, 77–90. [[CrossRef](#)]
136. Roberts, R.; Gott, R.; Glaesser, J. Students' approaches to open-ended science investigation: The importance of substantive and procedural understanding. *Res. Pap. Educ.* **2010**, *25*, 377–407. [[CrossRef](#)]
137. Mathesius, S.; Upmeier zu Belzen, A.; Krüger, D. Eyetracking als Methode zur Untersuchung von Lösungsprozessen bei Multiple-Choice-Aufgaben zum wissenschaftlichen Denken. In *Lehr- und Lernforschung in der Biologiedidaktik*; Hammann, M., Lindner, M., Eds.; Studienverlag: Innsbruck, Germany, 2018; pp. 225–244.
138. Osbeck, L.M. Scientific Reasoning as Sense-Making: Implications for Qualitative Inquiry. *Qual. Psychol.* **2014**, *1*, 34–46. [[CrossRef](#)]
139. Lorenz, C. Schriften aus der Fakultät Humanwissenschaften der Otto-Friedrich-Universität Bamberg. In *Diagnostische Kompetenz von Grundschullehrkräften: Strukturelle Aspekte und Bedingungen*; University of Bamberg Press: Bamberg, Germany, 2011.
140. Cronbach, L.J.; Meehl, P.E. Construct Validity in Psychological Tests. *Psychol. Bull.* **1955**, *52*, 281–302. [[CrossRef](#)]
141. Schwarzer, R.; Jerusalem, M. Das Konzept Der Selbstwirksamkeit. In *Selbstwirksamkeit und Motivationsprozesse in Bildungsinstitutionen*; Jerusalem, M., Hopf, D., Eds.; Beltz Verlag: Weinheim, Germany; Basel, Switzerland, 2002; pp. 28–53.
142. Sprenger, M.; Wartha, S.; Lipowsky, F. *Skalenhandbuch der Fortbildungsstudie QUASUM—Wirkungen einer Qualifizierungsmaßnahme zum Thema Rechenstörungen auf das Diagnostische Wissen und die Selbstwirksamkeitserwartungen von Mathematiklehrpersonen*; Pädagogische Hochschule Karlsruhe: Karlsruhe, Germany, 2015; Unpublished work.
143. Mayring, P. *Qualitative Inhaltsanalyse: Grundlagen und Techniken*; Beltz: Weinheim, Germany; Basel, Switzerland, 2015.

144. Stamann, C.; Janssen, M.; Schreier, M. Qualitative Inhaltsanalyse-Versuch einer Begriffsbestimmung und Systematisierung. *Forum Qual. Soc. Res.* **2016**, *17*, 16. [CrossRef]
145. Landis, J.R.; Koch, G.G. The Measurement of Observer Agreement for Categorical Data. *Biometrics* **1977**, *33*, 159–174. [CrossRef] [PubMed]
146. Taber, K. The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Res. Sci. Educ.* **2017**, *48*, 1273–1296. [CrossRef]
147. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*; Lawrence Erlbaum Associates: Hillsdale, NJ, USA, 1988.
148. Fensham, P. Science for All. *Educ. Leadersh.* **1986**, *44*, 18–23.
149. Next Generation Science Standards (NGSS). Topic Arrangements of the Next Generation Science Standards. Available online: <https://www.nextgenscience.org/sites/default/files/AllTopic.pdf> (accessed on 4 August 2021).
150. Hessisches Kultusministerium. Kerncurriculum Gymnasiale Oberstufe. Biologie. Available online: <https://kultusministerium.hessen.de/sites/default/files/media/kcgo-bio.pdf> (accessed on 5 August 2021).
151. Hessisches Kultusministerium. Kerncurriculum Gymnasiale Oberstufe. Chemie. Available online: <https://kultusministerium.hessen.de/sites/default/files/media/kcgo-ch.pdf> (accessed on 5 August 2021).
152. Niedersächsisches Kultusministerium. Kerncurriculum für das Gymnasium-Gymnasiale Oberstufe, die Gesamtschule-Gymnasiale Oberstufe, das Berufliche Gymnasium das Abendgymnasium, das Kolleg. Biologie. Available online: https://cuvo.nibis.de/cuvo.php?p=search&k0_0=Dokumentenart&v0_0=Kerncurriculum&k0_1=Schulbereich&v0_1=Sek+II&k0_2=Fach&v0_2=Biologie& (accessed on 5 August 2021).
153. Niedersächsisches Kultusministerium. Kerncurriculum für das Gymnasium-Gymnasiale Oberstufe, die Gesamtschule-Gymnasiale Oberstufe, das Berufliche Gymnasium das Abendgymnasium, das Kolleg. Chemie. Available online: https://cuvo.nibis.de/cuvo.php?p=search&k0_0=Dokumentenart&v0_0=Kerncurriculum&k0_1=Schulbereich&v0_1=Sek+II&k0_2=Fach&v0_2=Chemie& (accessed on 5 August 2021).
154. Ziepprecht, K.; Meier, M. Umsetzung und Weiterentwicklung von Modellen zur curricularen Vernetzung in hochschuldidaktischen Lernumgebungen in PRONET und PRONET². In *Vielfältige Wege biologiedidaktischer Forschung. Vom Lernort Natur zur Naturwissenschaftlichen Erkenntnisgewinnung in die Lehrerprofessionalisierung*; Meier, M., Wulff, C., Ziepprecht, K., Eds.; Waxmann: Münster, Germany, 2021; pp. 203–218.
155. Becerra-Labra, C.; Gras-Martí, A.; Martínez Torregrosa, J. Teaching Physics with a Fundamental-Problem-Based Approach: Effects on Conceptual Learning, Attitudes and Interests of University Students. *Rev. Bras. Ensino Fis.* **2006**, *29*, 95–103. [CrossRef]
156. Krüger, D.; Hartmann, S.; Nordmeier, V.; Upmeyer zu Belzen, A. Measuring Scientific Reasoning Competencies: Multiple Aspects of Validity. In *Student Learning in German Higher Education*; Zlatkin-Troitschanskaia, O., Pant, H.A., Toepper, M., Lautenbach, C., Eds.; Springer Fachmedien Wiesbaden: Wiesbaden, Germany, 2020; pp. 261–280.
157. Krell, M.; Mathesius, S.; van Driel, J.; Vergara, C.; Krüger, D. Assessing Scientific Reasoning Competencies of Pre-Service Science Teachers: Translating a German Multiple-Choice Instrument into English and Spanish. *Int. J. Sci. Educ.* **2020**, *42*, 2819–2841. [CrossRef]
158. Mahler, D.; Bock, D.; Bruckermann, T. Preservice Biology Teachers Scientific Reasoning Skills and Beliefs about Nature of Science: How do They Develop and is There a Mutual Relationship during the Development? *Educ. Sci.* **2021**, *9*, 558. [CrossRef]
159. Kersting, N. Using Video Clips of Mathematics Classroom Instruction as Item Prompts to Measure Teachers' Knowledge of Teaching Mathematics. *Educ. Psychol. Meas.* **2008**, *68*, 845–861. [CrossRef]
160. Kunz, H.; Wolowski, J. Wissenschaftliches Denken und Arbeiten im Kompetenzorientierten Biologieunterricht-Aufbau von Fachmethodischem Wissen in der Qualifizierung Angehender Lehrkräfte. In *Vielfältige Wege biologiedidaktischer Forschung. Vom Lernort Natur zur Naturwissenschaftlichen Erkenntnisgewinnung in die Lehrerprofessionalisierung*; Meier, M., Wulff, C., Ziepprecht, K., Eds.; Waxmann: Münster, Germany, 2021; pp. 189–202.
161. Joachim, C.; Hammann, M.; Carstensen, C.H.; Bögeholz, S. Modeling and Measuring Pre-Service Teachers' Assessment Literacy Regarding Experimentation Competences in Biology. *Educ. Sci.* **2020**, *10*, 140. [CrossRef]
162. Van der Valk, T.; de Jong, O. Scaffolding Science Teachers in Open-Inquiry Teaching. *Int. J. Sci. Educ.* **2009**, *31*, 829–850. [CrossRef]
163. Hilfert-Rüppell, D.; Looß, M.; Klingenberg, K.; Eghtessad, A.; Höner, K.; Müller, R.; Strahl, A.; Pietzner, V. Scientific Reasoning of Prospective Science Teachers in Designing a Biological Experiment. *Lehrbild. Prüfstand* **2013**, *6*, 135–154.
164. Rauin, U.; Meier, U. Subjektive Einschätzungen des Kompetenzerwerbs in der Lehramtsausbildung. In *Forschung zur Lehrerbildung*; Lüders, M., Ed.; Waxmann: Münster, Germany, 2007; pp. 102–131.
165. Pugh, K.J.; Bergin, D.A. Motivational Influences on Transfer. *Educ. Psychol.* **2006**, *41*, 147–160. [CrossRef]
166. Simosi, M. The Moderating Role of Self-Efficacy in the Organizational Culture–Training Transfer Relationship. *Int. J. Train. Dev.* **2012**, *16*, 92–106. [CrossRef]
167. Schütze, B.; Rakoczy, K.; Klieme, E.; Besser, M.; Leiss, D. Training Effects on Teachers' Feedback Practice. The Mediating Function of Feedback Knowledge and the Moderating Role of Self-Efficacy. *ZDM* **2017**, *49*, 475–489. [CrossRef]
168. Fauth, B.; Decristan, J.; Decker, A.-T.; Büttner, G.; Hardy, I.; Klieme, E.; Kunter, M. The Effects of Teacher Competence on Student Outcomes in Elementary Science Education: The Mediating Role of Teaching Quality. *Teach. Teach. Educ.* **2019**, *86*, 102882. [CrossRef]
169. Khan, S.; Krell, M. Scientific Reasoning Competencies: A Case of Preservice Teacher Education. *Can. J. Sci. Math. Techn. Educ.* **2019**, *19*, 446–464. [CrossRef]

170. Vorholzer, A.; von Aufschnaiter, C.; Kirschner, S. Entwicklung und Erprobung eines Tests zur Erfassung des Verständnisses Experimenteller Denk- und Arbeitsweisen. *ZfDN* **2016**, *22*, 25–41. [[CrossRef](#)]
171. Lawson, A.E.; Clark, B.; Cramer-Meldrum, E.; Falconer, K.A.; Sequist, J.M.; Kwon, Y.-J. Development of Scientific Reasoning in College Biology: Do Two Levels of General Hypothesis-Testing Skills Exist? *J. Res. Sci. Teach.* **2000**, *37*, 81–101. [[CrossRef](#)]
172. Bosse, S.; Spörer, N. Erfassung der Einstellung und der Selbstwirksamkeit von Lehramtsstudierenden zum inklusiven Unterricht. *Empir. Sonderpädagogik* **2014**, *4*, 279–299.
173. Jerusalem, M.; Satow, L. Schulbezogene Selbstwirksamkeitserwartung. In *Skalen zur Erfassung von Lehrer- und Schülermerkmalen*; Jerusalem, M., Schwarzer, R., Eds.; Freie Universität Berlin: Berlin, Germany, 1999; p. 15.
174. Caspary, R.; Blanck, B.; Spychiger, M.; Bosch, B.; Steinbrink, D.; Beutelspacher, A.; Kahl, R.; Osten, M.; Schumacher, R.; Spitzer, M. Der Produktive Umgang mit Fehlern. In *Nur wer Fehler Macht, Kommt Weiter. Wege zu einer Neuen Lernkultur*; Herder: Freiburg, Germany, 2008; pp. 49–72.
175. Schlatter, E.; Lazonder, A.W.; Molenaar, I.; Janssen, N. Individual Differences in Children's Scientific Reasoning. *Educ. Sci.* **2021**, *11*, 471. [[CrossRef](#)]
176. Herppich, S.; Praetorius, A.-K.; Förster, N.; Glogger-Frey, I.; Karst, K.; Leutner, D.; Behrmann, L.; Böhmer, M.; Ufer, S.; Klug, J.; et al. Teachers' Assessment Competence: Integrating Knowledge-, Process-, and Product-Oriented Approaches into a Competence-Oriented Conceptual Model. *Teach. Teach. Educ.* **2018**, *76*, 181–193. [[CrossRef](#)]
177. Behrmann, L.; van Ophuysen, S. Das Vier-Komponenten-Modell der Diagnosequalität. In *Diagnostische Kompetenz von Lehrkräften: Theoretische und Methodische Weiterentwicklungen*; Pädagogische Psychologie und Entwicklungspsychologie; Südkamp, A., Praetorius, A.-K., Eds.; Waxmann: Münster, Germany; New York, NY, USA, 2017; pp. 38–41.
178. Vogt, F.; Schmiemann, P. Assessing Biology Pre-Service Teachers' Professional Vision of Teaching Scientific Inquiry. *Educ. Sci.* **2020**, *10*, 332. [[CrossRef](#)]