

Analysis of content knowledge and pedagogical content knowledge of pre-service-chemistry teachers regarding the fundamental concept of 'structure-property-relations'

Analyse des fachlichen und fachdidaktischen Wissens von
Chemie-Lehramtsstudierenden
zum Basiskonzept "Struktur-Eigenschafts-Beziehungen"

Dissertation

zur Erlangung des akademischen Grades eines
Doktors der Naturwissenschaften
– Dr. rer. nat. –

vorgelegt von
Stefanie Herzog

geboren in Güstrow

Mathematisch-Naturwissenschaftliche Fakultät
der Christian-Albrechts-Universität zu Kiel

Kiel 2019

Gutachter:

Prof. Dr. Ilka Parchmann

Associate Prof. Dr. Vanessa Kind

Datum der Disputation:

04. Dezember 2018

Abstract

Professional knowledge of teachers has received considerable attention in recent years. For mathematics and the sciences, in particular, specific models representing teachers' knowledge have been set up theoretically and verified or altered empirically.

The field for pre-service teachers' professional knowledge is less well established, especially when considering the development of their chemistry-related knowledge during their education programs. As the whole of chemistry still seemed too broad for a thorough analysis regarding pre-service teachers' professional knowledge, a fundamental concept, covering comparable issues in a large variety of topics, was chosen as the focus. In order to contribute information to this field, a research project was initiated aiming to develop adequate, i.e. reliable and valid, test instruments for measuring pre-service teachers' professional knowledge regarding structure-property-relations. The fundamental concept as well as reasons for its selection are provided together with a presentation of literature on student difficulties focusing both on secondary and tertiary level.

Item development for content knowledge and pedagogical content knowledge as well as the use of these resulting instruments in a cross-sectional study are described. Apart from assessing pre-service teachers' subject-specific knowledge, the items are analyzed regarding possible factors influencing their difficulty. While a qualitative analysis of answers to content knowledge items reveals that pre-service teachers exhibit problems comparable to those described in literature, the test instruments show acceptable statistical values. Several factors can be identified that explain variance in test outcomes. For content knowledge, the following factors could be identified: gender, successfully passing exams in organic chemistry, grade in school leaving certificate, last school grade in chemistry, and having studied chemistry in upper secondary. For pedagogical content knowledge, those factors are: gender, successfully passing exams in teaching chemistry, successfully passing exams in organic chemistry, the participants' teaching program, the progress in their studies, as well as the grade in school leaving certificate. The combinations of factors explain 34.6% and 23.0% of variance in test results on content knowledge and pedagogical content knowledge, respectively. A dimensional analysis of knowledge presents content knowledge and pedagogical content knowledge to be separable but related dimensions, with a latent correlation between both dimensions of $\psi = .72$.

In a subsequent step, the measures on structure-property-relations are combined with those on chemical representations, which have been developed and tested by another researcher. A resulting dimensional analysis shows all four dimensions to be separable and the dimensions showing latent correlations between the two content knowledge dimensions of $\psi = .87$ and between the two pedagogical content knowledge dimensions of $\psi = .67$. A following addition of a test on pedagogical knowledge yields a five-dimensional model of pre-service teachers' professional knowledge as assessed through these measures. The correlation pattern is according to expectations, the highest between the chemistry-specific knowledge dimensions and considerably lower between any of the chemistry-specific dimensions and pedagogy.

In the end, results are discussed, deriving at both suggestions for further research studies and a short practitioner's perspective on the topic.

Zusammenfassung

Professionswissen von Lehrkräften hat in den letzten Jahren beachtliche Aufmerksamkeit erzielt. Insbesondere für die Fächer Mathematik und die Naturwissenschaften sind spezifische Modelle zum Wissen von Lehrkräften theoretisch fundiert aufgestellt und empirisch verifiziert oder überarbeitet worden.

Das Feld der Forschung zum Professionswissen von Lehramtsstudierenden ist weniger stark ausgeprägt, insbesondere unter Berücksichtigung einer Entwicklung des chemiespezifischen Wissens über die Dauer des Lehramtsstudiums. Da der gesamte Bereich der Chemie für eine detaillierte Untersuchung des Professionswissens von Lehramtsstudierenden zu umfangreich erscheint, ist ein Basiskonzept, welches vergleichbare Aspekte bei einer Vielzahl an Themen betrachtet, als Untersuchungsschwerpunkt gewählt worden. Um in diesem Bereich Ergebnisse beisteuern zu können, ist ein Forschungsprojekt initiiert worden, in welchem adäquate, also reliable und valide Testinstrumente entwickelt werden sollten, die das Professionswissen von Lehramtsstudierenden im Hinblick auf das Struktur-Eigenschafts-Konzept messen sollten. Das Basiskonzept als auch die Gründe für dessen Auswahl werden dargestellt, gemeinsam mit einem Literaturüberblick über Schwierigkeiten von Lernenden im sekundären und tertiären Bereich.

Die Aufgabenentwicklung für das Fachwissen und fachdidaktische Wissen sowie der Einsatz der resultierenden Instrumente in einer Querschnittserhebung werden beschrieben. Neben einer Untersuchung des fachspezifischen Wissens von Lehramtsstudierenden werden auch die Items im Hinblick auf mögliche schwierigkeitsgenerierende Faktoren analysiert. Während eine qualitative Analyse der Antworten zu den Fachwissensitems zeigt, dass Lehramtsstudierende vergleichbare Schwierigkeiten haben wie in der Literatur beschrieben, weisen die Testinstrumente selbst zufriedenstellende statistische Kennwerte auf. Verschiedene Faktoren können identifiziert werden, die Varianz im Testergebnis erklären. Für das Fachwissen ergeben sich Geschlecht, das Bestehen von Klausuren in der organischen Chemie, die Abiturnote, die letzte Chemienote in der Schule sowie die Belegung von Chemie in der Oberstufe als Faktoren. Beim fachdidaktischen Wissen handelt es sich um die Faktoren Geschlecht, das Bestehen von Klausuren in der organischen Chemie und Fachdidaktik, das Studiensystem (Bachelor/Master vs. Staatsexamen), der Zeitpunkt im Studium sowie die Abiturnote. Die Kombination der Faktoren klären beim Fachwissen 34.6% und beim fachdidaktischen Wissen 23.0% der Varianz in den Testergebnissen auf. Eine Dimensionsanalyse des Wissens zeigt, dass Fachwissen und fachdidaktisches Wissen trennbare Konstrukte sind, die mit einer latenten Korrelation von $\psi = .72$ zusammenhängen.

In einer anschließenden Analyse werden die Instrumente zum Struktur-Eigenschafts-Konzept mit denen zu chemischen Repräsentationen kombiniert, welche von einer anderen Doktorandin entwickelt und überprüft worden sind. Die sich ergebende Dimensionsanalyse zeigt, dass vier Konstrukte trennbar sind und dass die zwei Konstrukte zum Fachwissen eine latente Korrelation von $\psi = .87$ und die beiden Konstrukte zum fachdidaktischen Wissen eine latente Korrelation von $\psi = .67$ aufweisen. Eine weitere Ergänzung um ein Instrument zum pädagogischen Wissen ergibt ein fünfdimensionales Modell zum Professionswissen von Lehramtsstudierenden. Das Korrelationsmuster ist erwartungskonform: die höchsten Korrelationen treten zwischen den chemiespezifischen Wissensdimensionen auf, während zwischen den chemiespezifischen Wissensdimensionen und dem pädagogischen Wissen jeweils deutlich geringere Korrelationen auftreten.

Am Ende werden die Ergebnisse diskutiert und sowohl Vorschläge für zukünftige Forschung als auch eine kurze Perspektive aus unterrichtspraktischer Sicht abgeleitet.

Table of Contents

1 INTRODUCTION.....	1
2 THEORETICAL FRAMEWORK.....	8
2.1 Research context.....	8
2.1.1 Secondary education in Germany	8
2.1.2 Tertiary education regarding a teaching career in Germany	11
2.1.2.1 General aspects	11
2.1.2.2 Aspects related to chemistry	14
2.2 Teachers' professional knowledge	17
2.2.1 Structure of professional knowledge for teaching	17
2.2.2 Empirical competence models for students	24
2.2.3 Competence models for (pre-service) teachers	28
2.3 Structure-Property-Relations (SPR)	35
2.3.1 Content knowledge (CK)	37
2.3.1.1 Topic-transcending consistent perspective	38
2.3.1.2 Levels of thinking	39
2.3.1.3 Chemistry-related reasoning	40
2.3.1.4 Transferability to tertiary level	43
2.3.2 Pedagogical Content Knowledge (PCK)	45
2.3.2.1 Short overview of PCK models	45
2.3.2.2 Curricular development of SPR	51
2.3.2.3 Problems regarding student understanding reported in connection with SPR	54
3 RESEARCH PROJECT AND RESEARCH OBJECTIVES	60
4 ITEM DEVELOPMENT	64
4.1 Content Knowledge.....	64
4.1.1 Sources for and content of items.....	64

4.1.2	Design of items.....	67
4.1.2.1	Complexity	67
4.1.2.2	Number of perspectives	68
4.1.2.3	Educational allocation	69
4.1.2.4	Item format	70
4.2	Pedagogical Content Knowledge	73
4.2.1	Sources for and content of items.....	73
4.2.2	Design of items.....	75
4.3	Communicative item validation	79
5	STUDY I: ANALYSIS OF TEST INSTRUMENTS.....	83
5.1	Research objective	83
5.2	Test administration	85
5.3	Data preparation & statistical measures.....	86
5.4	CK and PCK instruments.....	88
5.5	Results.....	91
5.5.1	Demographic data of the participants in study I.....	91
5.5.2	CK – individual item analysis.....	92
5.5.3	CK – resulting instrument.....	107
5.5.4	CK – factors influencing item difficulty.....	108
5.5.5	PCK – individual item analysis.....	113
5.5.6	PCK – resulting instruments.....	118
5.5.7	Performance on tests.....	119
5.6	Summary and discussion of findings	122
6	STUDY II: STRUCTURE OF PROFESSIONAL KNOWLEDGE	129
6.1	Research objective	129
6.2	Instruments.....	131
6.3	Data preparation & statistical measures.....	136
6.4	Results.....	138

6.4.1 Description of sample.....	138
6.4.2 Descriptive data analysis	139
6.4.3 Factors explaining variance in test outcome.....	142
6.4.4 Dimensional analysis of all chemistry items from the <i>KiL</i> -project: Chemical Representations (CRep) and SPR.....	151
6.5 Summary and discussion of findings	160
7 FINAL DISCUSSION AND IMPLICATIONS	169
8 ADDENDUM – A PRACTITIONER'S PERSPECTIVE	174
9 REFERENCES.....	182
10 APPENDICES	a
10.1 Overview of items.....	a
10.2 SPR items in study II.....	a
10.3 Overview of items in Wright Map (SPR only).....	b
10.4 Overview of items in Wright Map (CRep and SPR).....	c

Table of Figures; labels are shortened

Figure 1: Intended knowledge and possible experiences received through instruction at school	3
Figure 2: Intended subject-specific knowledge areas received through instruction at university for pre-service teacher's role as a learner	3
Figure 3: Demand for an integration of university knowledge and reflected school knowledge as well as experiences for pre-service teacher's role as a future teacher	4
Figure 4: Advanced organizers quickly allowing the reader to deduce which knowledge domain, educational level and role of the pre-service teacher a particular chapter will focus on	7
Figure 5: German educational system	8
Figure 6: Three-part division in teacher education	13
Figure 7: Maximum chemistry content ECTS-points in academic track (Gymnasium) teaching program per state	14
Figure 8: <i>COACTIV</i> generic structure model of professional competence	20
Figure 9: PISA framework for scientific literacy	25
Figure 10: Facets in competence model according to SCHECKER & PARCHMANN	26
Figure 11: Competence model for all sciences according to ESNaS	28
Figure 12: Content sub-dimension in <i>COACTIV</i>	29
Figure 13: Competency model of mathematical PCK according to TEDS-M study	31
Figure 14: Competency model according to RIESE	32
Figure 15: Competence model for chemistry teachers' CK according to <i>ProwiN</i>	33
Figure 16: Competence model for chemistry teachers' PCK according to <i>ProwiN</i>	33
Figure 17: Correlations between latent dimensions in <i>COACTIV</i>	34
Figure 18: Example of submicro-meso-macro-thinking	40
Figure 19: Consensus Model of Teacher Professional Knowledge Bases including PCK	48
Figure 20: Excerpt of curricular sequence for students' understanding of a discipline-specific cross-cutting concept	51
Figure 21: Learning progression on structure and function, I	52
Figure 22: Curricular development of the SPR-concept	53
Figure 23: Overview of students' problems	58
Figure 24: Learning progression on structure and function, II	61
Figure 25: CK item foci based on possible curricular development and described difficulties	65
Figure 26: Resulting model characterizing CK items	72
Figure 27: Resulting model including relevant aspects characterizing PCK items	78
Figure 28: Overview of testing session	85
Figure 29: Difficulty of CK items from test booklet A	92
Figure 30: Difficulty of CK items from test booklet B	93
Figure 31: Difficulty of CK items from test booklet C	93

Figure 32: Drawing and explanation of person P0778.....	102
Figure 33: Variety of items addressing hydrogen bonds.....	110
Figure 34: Difficulty of PCK items in test booklet B.....	113
Figure 35: Difficulty of PCK items in test booklet C.....	113
Figure 36: Overview of overall test performance.....	120
Figure 37: Test performance according to year of studies (test booklet B).....	120
Figure 38: Wright Map of CK and PCK.....	141
Figure 39: Overview of different variables that might have effect on test performance.....	143
Figure 40: Mean CK test outcome per chemistry level at upper secondary.....	144
Figure 41: Mean PCK test outcome per chemistry level at upper secondary.....	144
Figure 42: Wright Map of all four chemistry dimensions.....	154
Figure 43: Test results regarding the students' year of their studies.....	156
Figure 44: Knowledge development for each dimension from beginning to end of studies.....	157
Figure 45: Structure of professional knowledge of pre-service chemistry teachers.....	158
Figure 46: Structures of teachers' and pre-service teachers' professional knowledge.....	168
Figure 47: Knowledge acquisition at university.....	171
Figure 48: Developmental model for the SPR concept.....	177
Figure 49: Setup for testing steam permeability of a textile.....	179

Table of Tables; labels are shortened

Table 1: Overview of competency areas for chemistry.....	11
Table 2: Suggestion regarding content that should be covered by students aiming to teach chemistry.....	15
Table 3: Suggestions for contents regarding chemistry-education courses.....	16
Table 4: Overview of some exemplary models of teachers' professional knowledge	19
Table 5: Facets of SCHECKER & PARCHMANN model.....	26
Table 6: PCK sub-dimension according to <i>COACTIV</i> model.....	30
Table 7: Descriptions of knowledge domains according to BALL ET AL.....	30
Table 8: Facets of the PCK dimension according to RIESE.....	32
Table 9: Short overview of fundamental concepts for lower secondary education	35
Table 10: Selection of examples of how the SPR concept provides consistent chemical perspective.....	38
Table 11: A closer look at possible sub-dimensions of PCK.....	47
Table 12: Exemplary content of SPR-items.....	67
Table 13: Classification of an exemplary item according to the varying factors	71
Table 14: Comparison of kinds of knowledge to complexity levels for PCK.....	77
Table 15: Assignment of items to test booklets	89
Table 16: Item formats per test booklet (TB).....	90
Table 17: Answer patterns and coding scheme for item E1.....	95
Table 18: Answer patterns and coding scheme for item E5.....	97
Table 19: Answer patterns and coding scheme for item E17, verbal answers	100
Table 20: Answer patterns and coding scheme for item E17, drawing of chemical structures..	101
Table 21: Answer patterns and coding scheme for item R11_1, definition of hydrogen bond ...	103
Table 22: Answer patterns and coding scheme for item R11_1, drawing of hydrogen bond	105
Table 23: Exemplary drawings of a hydrogen bond.....	106
Table 24: Discriminatory power per CK item, booklet B	107
Table 25: Statistical values of the three CK scales	108
Table 26: Characterization of items regarding hydrogen bonds.....	112
Table 27: Answer patterns and coding scheme for item I6	116
Table 28: Coding of answers for item S4.....	116
Table 29: Distribution of answers for item S4.....	117
Table 30: Coding of answers for item EM10	118
Table 31: Distribution of answers for item EM10	118
Table 32: Statistical values of the two PCK scales.....	119
Table 33: Additional measures in study II and expected correlations with the SPR knowledge dimensions.....	135

Table 34: Comparison of sample of study I and II.....	139
Table 35: Model fit analysis, SPR dimensions.....	140
Table 36: Overview of means for both SPR dimensions.....	142
Table 37: Results for correlations of person parameters and certain variables.....	147
Table 38: Results for <i>t</i> -tests of certain variables on person parameters.....	147
Table 39: Results for ANOVA of certain variables on person parameters.....	148
Table 40: Distribution of chemistry level in upper secondary education by gender and corresponding chi-square test.....	149
Table 41: Regression analysis regarding success in SPR-CK test.....	150
Table 42: Regression analysis regarding success in SPR-PCK test.....	151
Table 43: Model fit analysis, CRep and SPR dimensions.....	152
Table 44: Overview of means for all four dimensions.....	155
Table 45: Correlations of the person parameters of the four content dimensions with other constructs.....	159
Table 46: Comparison of an inquiry-based approach with different aims and quite similar paths.....	176

Acknowledgments

I would like to extend my thanks to

- the research groups of Peter Childs and Odilla Finlayson, as well as Oliver Tepner and Vanessa Kind, for a fruitful discussion of the items regarding understandability, difficulty for students, as well as opportunities for improving individual items
- Stefan Schwarzer, for providing the picture of the nano gold solutions
- the chemistry education department at IPN for ideas regarding central aspects of the SPR concept and valuable feedback concerning item formulation and assessment of student answers
- the Leibniz Association for the funding of the research project KiL (Project number SAW-2011-IPN-2)

1 INTRODUCTION

One leading theme in educational research has been presented by SHULMAN (1986):

“Our central question concerns the transition from expert student to novice teacher. How does the successful college student transform his or her expertise in the subject matter into a form that high school students can comprehend? When this novice teacher confronts flawed or muddled textbook chapters or befuddled students, how does he or she employ content expertise to generate new explanations, representations, or clarifications? What are the sources of analogies, metaphors, examples, demonstrations, and rephrasings?” (p.8)

In order to eventually find out how university knowledge – making the pre-service teacher an *expert student* in Shulman’s statement – is used and altered when transferred from the learning at university to teaching at school, one must first assess how much pre-service teachers¹ actually know about the content and the aspects relevant for teaching. That is what this dissertation project attempts: developing instruments that will allow for a sensible measurement of pre-service teachers’ knowledge.

In Germany and other countries, chemistry knowledge in secondary education is structured according to basic or fundamental concepts. A large amount of international research studies have revealed, however, that secondary students have numerous problems with a systematic understanding of fundamental chemical concepts (see among others BUTTS & SMITH, 1997; GILBERT & TREAGUST, 2009; KIND, 2004). If one considers teachers as the central mediator between secondary student and chemistry knowledge, the question arises how teachers themselves understand and then make use of these fundamental concepts. Looking at teacher professionalization in general and pre-service teacher education in particular, one

¹ The term pre-service teacher will be used to indicate a person enrolled in a university (or equivalent higher education institution) program for becoming a teacher regardless of his/her specific stage in the degree program; the person may have undergone some practical training in class but is not yet fully certified to teach without any supervision. The terms *student teacher*, *student in teacher training* or *teacher training student*, which are also often used to describe members of this group, have deliberately not been chosen for this thesis because parts in the theoretical framework will focus on secondary students and confusion between the two ‘student’ populations was to be avoided. Another term referring to pre-service teacher is *teacher trainee*; this term, however, lacks the connotation of the person being theoretically educated at a tertiary institution and has therefore not been chosen.

might wonder whether these observed problems regarding comprehending fundamental concepts can already be found in pre-service teacher education.

German students arrive at university with quite heterogeneous prior knowledge regarding chemistry. This is a consequence of the education system in Germany and is true for both length and depth of their secondary chemistry education. Regardless of the heterogeneity, chemistry at secondary level is sometimes taught through inquiry-based or context-based learning approaches (NENTWIG, DEMUTH, PARCHMANN, GRÄSEL & RALLE, 2007; SCHMIDKUNZ & LINDEMANN, 1992). Through these approaches, students are usually presented with macroscopic properties of a substance or an industrial product relevant in everyday life as a starting point of their own investigations. Eventually, they are supposed to derive an underlying understanding of submicroscopic aspects of these products (cf. among others BENNETT & LUBBEN, 2006; PARCHMANN, GRAESEL, ET AL., 2006). While these approaches, therefore, present prime examples of one side of the structure-property-relationship that is vital for understanding chemistry, i.e. macroscopic properties can be explained through submicroscopic structures, this understanding is intended regardless of the teaching approach. The other direction of this relationship is just as relevant, especially for producing products or substances with a specific function: obtaining desirable macroscopic properties by constructing a certain submicroscopic structure. While the former is usually stressed more heavily in school, the latter is focused on more closely in university laboratory work. These two sides of the macroscopic – submicroscopic relationship show why an understanding of this concept is essential for comprehending chemistry. This is emphasized even further as *structure-property-relations* presents one of four² basic or fundamental chemical concepts taught in German secondary schools (cf. NENTWIG ET AL., 2007). Each fundamental concept comprises a wide variety of individual topics, so that the concept of structure-property-relations (SPR) contains, for example, topics such as the observed color of different substances, the acid strength of various acids, the nucleophilic or electrophilic character of different organic substances as well as the degrees of resistance to heat of different plastics. It becomes obvious even from this rather short list that the intention of secondary chemistry education is to facilitate a learner's acquisition of

² Sometimes the reader may find that five or six fundamental concepts are mentioned in literature for German secondary school curricula; this is due to the fact that one of these concepts *chemical reactions* is split in upper secondary into *donor-acceptor-concept* and *kinetics and chemical equilibrium-concept* (and sometimes the latter being further subdivided into the two).

chemistry content (i.e. subject-matter knowledge). As a possible by-product, students might gain experiences regarding their own or their classmates' difficulties, regarding representations and models used by the teacher, regarding experiments that could illustrate certain things and regarding a general structure of going from simple to more complex explanations. Sometimes, students integrate intended knowledge and these unintended experiences. Figure 1 shows these products of secondary chemistry education, with which pre-service teachers start their university program.

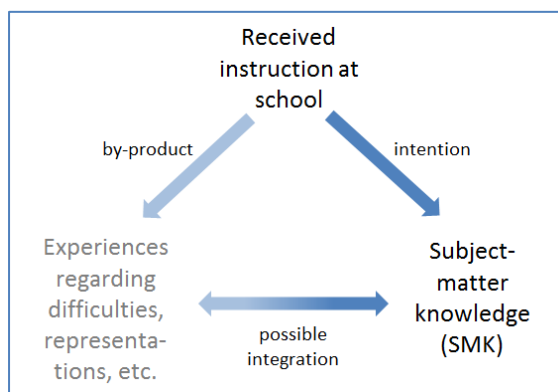


Figure 1: Intended knowledge and possible experiences received through instruction at school

In the current situation of German pre-service chemistry teacher education, university students have to assume two roles simultaneously: One being that of a *learner* in which they are exposed to a variety of chemistry content subdivided into the discipline's traditional areas of general, inorganic, organic and physical chemistry. According to models on teachers' professional knowledge (cf. BAUMERT & KUNTER, 2006; BLÖMEKE ET AL., 2014; SHULMAN, 1986), content knowledge needs to be flanked by pedagogical knowledge and pedagogical content knowledge in order to generate successful teachers. This notion is also reflected in a sense in teacher education programs in Germany since pre-service teachers have to pass courses in these three areas right from the beginning of their university career. As these three strands are usually taught by different staff or even different departments, an explicit interconnectedness between them is often marginal and thus a possible integration of the subject-specific knowledge areas, as presented in Figure 2, is up to the learner. Regarding the integration of the knowledge

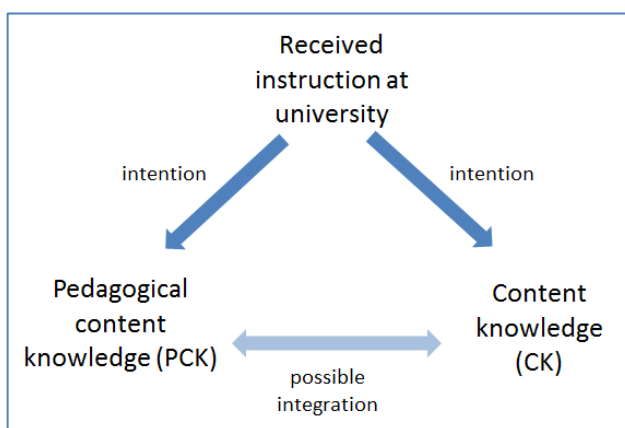


Figure 2: Intended subject-specific knowledge areas received through instruction at university for pre-service teacher's role as a learner

areas, the learner's situation at tertiary level is, therefore, very similar to that of the learner at secondary level.

The other role of a pre-service teacher is that of a *future teacher* in which he/she – already while at university – has to (anticipate to) stand in front of a class and apply his/her acquired knowledge. In order to deliver quality teaching at school in the future, the pre-service teacher has to integrate i) the pedagogical content knowledge learnt at university with his/her knowledge and experience gained at school in a reflection process and ii) this knowledge about teaching with the content knowledge studied at university. Figure 3 shows this demand. When integrating the different knowledge areas, pre-service teachers have to specifically combine two categorization schemes for teaching chemistry content knowledge: the categorization scheme of fundamental concepts learnt at school and that of the traditional sub-disciplines learnt at university. The question arises how well (or even if) pre-service teachers manage to integrate chemistry knowledge acquired from those two separate categorization schemes.

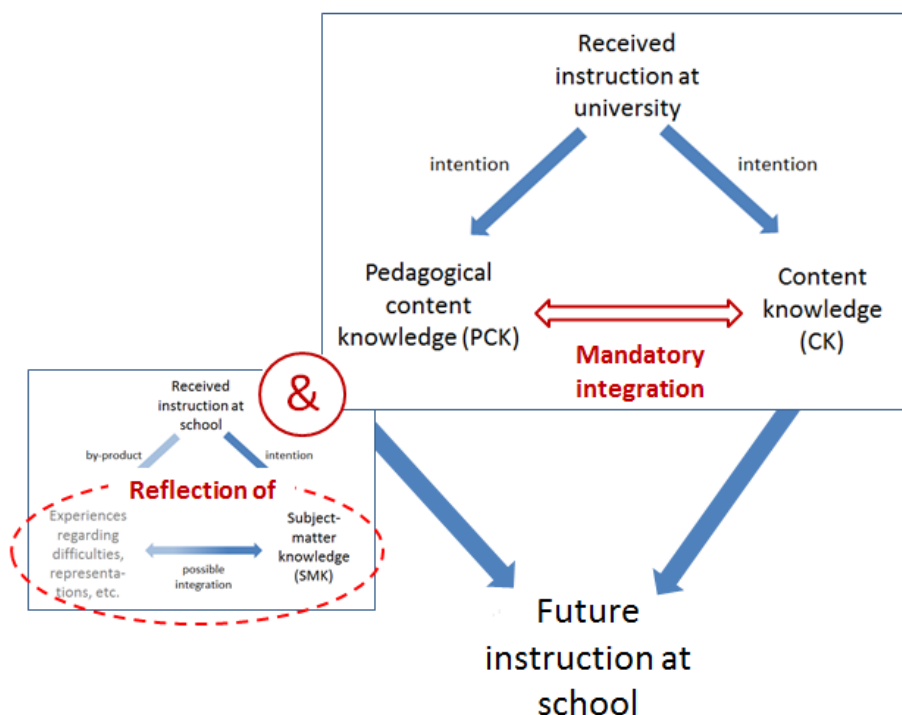


Figure 3: Demand for an integration of university knowledge and reflected school knowledge as well as experiences for pre-service teacher's role as a future teacher

Since pre-service teachers often report that they “feel pedagogically under-qualified [and] subject specifically over-qualified” (SCHUMACHER & LIND, 2000, p.25) and since this impression is also shared by some experienced teachers, especially when thinking of pre-service teachers aiming for a Gymnasium³ teaching degree (ZECHERLE, 2011), it is conceivable that many pre-service teachers ignore much of what they have learnt at university upon returning to school in order to teach. It has been noted that beginning teachers often struggle because they fail to combine university content and school content (MNU, 2004). That is why it is necessary to investigate pre-service chemistry teachers' understanding of university content, which will be attempted in this dissertation with the focus on a secondary level fundamental concept.

Therefore, this research project will take a first step in the direction of assessing the chemistry-related knowledge dimensions. Despite the existence of several models used for a general assessment of content knowledge and pedagogical content knowledge in chemistry (cf. chapter 2.2), a test instrument focusing on a single fundamental concept for both knowledge dimensions does not exist yet. It is intended to create two test instruments that will assess pre-service chemistry teachers' knowledge of the structure-property-concept in the areas of content knowledge as well as pedagogical content knowledge. In a later step, these findings will be combined with those obtained with test instruments on chemical representations and pedagogical knowledge which are being developed alongside by other members of the larger research project.

When developing an instrument that is meant to analyze pre-service chemistry teachers' professional knowledge regarding the fundamental concept of structure-property-relationships, several aspects have to be considered in the theoretical framework (chapter 2). First, the general research context has to be described (chapter 2.1), i.e. presenting the requirements for German students who want to become chemistry teachers. This chapter will present both a brief overview of students' diverse secondary background in chemistry as well as the variety in teacher training programs in tertiary education. The latter will be presented with the German

³ The German education system and the school types will be explained briefly in chapter 2.1.1.

focus, but a comparison with other systems will be included because only this might allow for transferability to other research contexts.

After having established the general situation of German pre-service teachers, an overview of the concept of professional knowledge will be given along with an outline of models how to assess or measure this professional knowledge and more specific competencies (chapter 2.2). The competence definition for this thesis is based on WEINERT (2001a) and KLIEME (2004) and includes cognitive skills and abilities as well as motivational, volitional and social willingness and skill to solve specific problems in certain situations. According to TEODORESCU (2006), competence can be differentiated from competency according to which aspects research focuses on, which results are obtained through models and which objectives they are used for. While *competence* describes “the desired end state for individual performance [...and...] can be used to provide guidelines to success, assess measurable gaps, and direct people to tools, resources, and training” (TEODORESCU, 2006, p.27ff.), models defining *competency* usually provide “the skills, knowledge, attributes, and desirable behavior thought to be required for successful performance for a specific job role [...and can be used...] to build training, hiring, evaluation, and assessment programs” (TEODORESCU, 2006, p.28f.). This distinction shall also be attempted in this thesis: competence and competence models will denote the desired pre-service teacher level of performance in domain-specific knowledge areas, i.e. content knowledge and pedagogical content knowledge, and assess stages on that path, while competencies in both domain-specific knowledge areas will look at more finely-grained skills, i.e. for example the ability to properly explain phenomena to a particular age group of learners.

Finally, it is necessary to zoom in on the content level that is to be assessed. Here, the fundamental concept of structure-property-relationships will be described as the common core for (pre-service) chemistry teacher’s professional knowledge (chapter 2.3).

These descriptions will then lead to research questions with corresponding hypotheses (chapter 3) which this thesis project hopes to answer. Before describing the development of the testing instruments and testing conditions (chapter 4), the integration of this thesis project into a larger research project will be described and followed by methodological considerations. Then the results of the testing phases

(chapter 5 and 6) will be presented along the lines of the research questions. After the general discussion of the project and its implications (chapter 7), a practitioner's perspective is added describing how to integrate content and pedagogical aspects of this fundamental chemical concept in modules of tertiary education for pre-service teachers (chapter 8).

Most sub-chapters in the theoretical framework will depict an advanced organizer, telling the reader which chemistry-specific knowledge area, which educational level and which role of the pre-service teacher the chapter will refer to. The left picture of Figure 4 will be used, for example, when describing the situation of the *secondary school students* in chemistry classes dealing with a combination of content and experiences; the middle picture states that the chapter will focus on the pre-service teacher in his or her *role of a learner* of knowledge about how to use chemistry methods, strategies, analogies, etc.; while the right picture implies that the information presented refers to the pre-service teacher *applying* this pedagogical content knowledge in practice teaching situations.



Figure 4: Advanced organizers quickly allowing the reader to deduce which knowledge domain, educational level and role of the pre-service teacher a particular chapter will focus on (left: school student in chemistry class; middle: pre-service teacher learning pedagogical content knowledge; right: pre-service teacher applying pedagogical content knowledge)

2 THEORETICAL FRAMEWORK

2.1 Research context

To fully understand this thesis it will be important for the reader to know the situation of future chemistry teachers in German education. Therefore, their situation will be described in the following sections.

2.1.1 Secondary education in Germany

GENERAL ASPECTS

CK	secondary	learner
PCK	tertiary	teacher

Usually, parents have to choose after grade four which school type they want their children to go to for secondary education (cf. Figure 5). Their choice is between *Gymnasium*, equipping the child with the requisites for an academic education and possible career, and, depending on the German state they live in, *Hauptschule/Realschule/Mittelschule/Oberschule*, preparing the child for a vocational career. The *Gesamtschule* or *Gemeinschaftsschule* often has a somewhat special role as it combines other non-academic-school types but still sometimes has its own upper secondary track and therefore can also prepare children for university.

The German Educational System

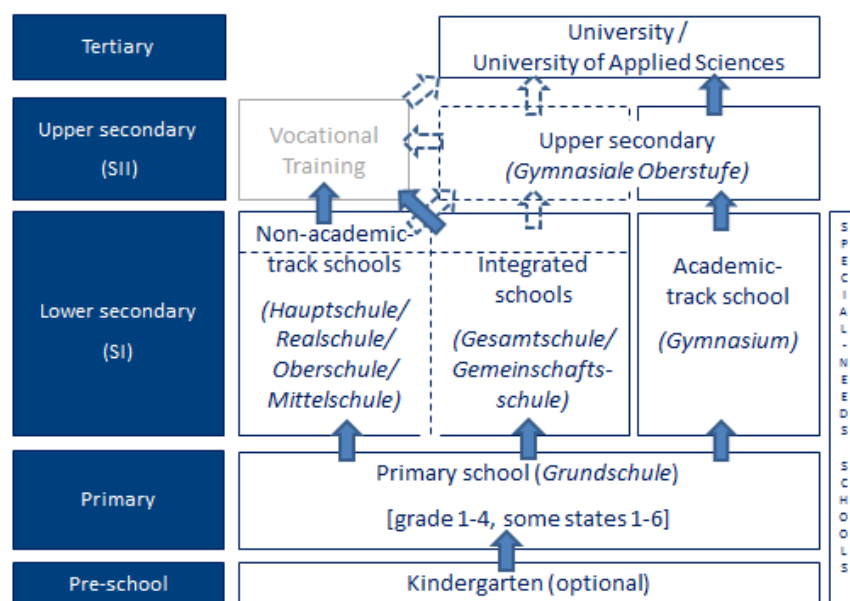
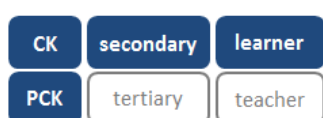


Figure 5: German educational system, based on KMK, 2015 (blue arrows indicate the main path through the system while white arrows show where transitions within the system are also possible)

Even though possibilities to transfer between the school types are inherent in the system, there are main paths that most students from the different school types take (blue arrows in Figure 5). However, the picture of school types becomes even more diverse than Figure 5 suggests, and that is because in Germany, education is not a federal but a state-regulated issue: Each state government regulates how many different school types there are. While all states offer a *Gymnasium*, states usually offer many non-academic track school types which can differ between states both in name or concept (i.e. *Gemeinschaftsschule* in one state might be called *Sekundarschule* in a second and *Mittelschule* in a third state).

As each school type differs in depth and breadth of each subject offered to students, the educational landscape in Germany is quite diverse. Apart from the sovereignty of the federal states, a second factor adds to the heterogeneity in the educational system: school autonomy. Most states offer their schools a considerable degree of autonomy regarding decisions in areas of finances, staff, organization and administration, pedagogical tasks or school quality (VEREINIGUNG DER BAYERISCHEN WIRTSCHAFT EV, 2010). Therefore, even within one state, neighboring schools of the same school type might not begin to teach a certain subject in the same grade, might use different school books and are likely to have – independently of other schools – formulated their own school-internal curriculum for each subject.

ASPECTS RELATED TO CHEMISTRY



For chemistry at the secondary level this means that each state or even each school per state regulates at which grade students in school start learning this subject, whether chemistry is taught as a distinct subject or in the form of integrated science classes and which curricula there are to be followed (RISCH, 2010b) as well as which options there are to drop or intensify the subject in upper secondary. Usually, the non-academic track schools teach chemical aspects in an integrated science subject while in academic track schools this might be true for 5th and 6th grade⁴ at most. Typically for the latter school type, all sciences are taught as individual subjects.

⁴ In some states, 5th and 6th grade still belong to primary education.

Nevertheless, in the scope of this research context, it is necessary and possible to draw some general conclusions as to what chemical aspects are taught at the lower secondary level: students will know something about air and water, liquids, salts, metals, and gases at a more or less basic level. They will have learnt to describe observable properties of these substances and to explain these based on some level of submicroscopic characteristics applying simple models. Students will have had opportunities to plan, perform and interpret experiments focusing on separation methods and qualitative analysis of certain elements or compounds. Learners will also have dealt with the concept of chemical reactions with examples from inorganic and organic chemistry and will have covered reaction types such as acid-base reactions and simple redox reactions, possibly also going into simple reaction mechanisms. As stated before, the extent to which this is done differs considerably among school types and even individual schools of the same type. Considering that even at the Gymnasium students have a choice regarding their chemistry classes at upper secondary (ranging from no chemistry at all to an advanced level where physical chemistry and in-depth organic chemistry are covered), one might understand why students arrive at university with very heterogeneous prior knowledge in chemistry (BUSKER, 2010; or KLOSTERMANN, 2014). Among other goals, it was hoped to reduce this diversity through the introduction of National Educational Standards (NES) for chemistry in 2004 by the Standing Conference of the Ministers of Education and Cultural Affairs of the Federal States in the Federal Republic of Germany (KMK, 2005). Based on the concept of scientific literacy (NEUMANN, FISCHER & KAUERTZ, 2010), the NES describe various competences that secondary students should have mastered when finishing mandatory education. For chemistry, these individual competence standards are grouped into four competence areas: (i) subject knowledge, (ii) epistemological knowledge and subject methods, (iii) communication, and (iv) judgement (SCHECKER & PARCHMANN, 2007). A short description of each competence area is given in Table 1.

Table 1: Overview of competency areas for chemistry

Label of competence area (SCHECKER & PARCHMANN, 2007)	Description of standards per competence area (adapted from NEUMANN ET AL., 2010)
subject knowledge	Students should possess a profound, well-structured knowledge based on the basic concepts. Students should be able to use analogies to solve tasks and problems.
epistemological knowledge and subject methods	Students should be able to formulate hypotheses for standard examples. Students should be able to plan, perform, and analyze standard experiments.
Communication	Students should be able to research a topic using different sources. Students should be able to present results of their work using appropriate terminology and representations.
Judgement	Students should be able to compare and judge alternative (technical) solutions considering subject-specific, economic, social, and ecological aspects. Students should be able to name implications of subject-related findings with respect to historical and social contexts.

The standards and competence areas will be relevant again in chapters 2.2.2 & 2.2.3 when various competence models will be looked at.

2.1.2 Tertiary education regarding a teaching career in Germany

2.1.2.1 General aspects



As Figure 5 suggests, there are at least two possibilities for an academic career at tertiary level: students can attend more specialized universities of applied sciences (so-called *Fachhochschulen*) or general, full-spectrum universities. Since the sovereignty of the German states in regard to education is not only true for secondary but also for tertiary education, some German states offer a third university type, universities of cooperative education (so-called *Berufsakademien*) (TRAUTWEIN ET AL., 2006), while one state offers a fourth type, i.e. universities of education (co-called *Pädagogische Hochschulen*). When aiming at a career as a future teacher, students in Germany could choose between 67 institutions in the year 2016 that offered a teaching degree: six universities of education (only in one state and only for non-academic track schools) and 61 full-spectrum universities (CENTRUM FÜR HOCHSCHULENTWICKLUNG, 2016). One could reasonably argue that the vast majority of pre-service teachers are, therefore, educated at a general university. But similar to

schools, universities are not only regulated by state laws but have their own autonomy in regard to many aspects, from budget considerations to questions regarding the degrees offered to contents to be taught (VEREINIGUNG DER BAYERISCHEN WIRTSCHAFT E.V., 2010). That is why across the universities, the teaching degrees can differ with regard to a Bachelor/Master or so-called *Staatsexamen* (i.e. state examination) certificate, with regard to issuing a teaching license for academic-track schools or non-academic track schools, and of course, with regard to the contents taught.

Pre-service teacher education in Germany is divided into two separate phases: I) a more theoretically-based phase at university or equivalent teacher training institution which generally lasts four to five years, and II) a practical phase with on the job training at school with a duration of usually 18 months⁵. According to TERHART (2002), the two phases of teacher education have different intentions: The first phase is focused on knowledge, reflection of and communication about certain contents, while in the second part, in which students teach actual classes at primary or secondary school while being supervised, counseled and reviewed by experienced teachers, the focus is on diagnosing and assessing learners and conveying content to them. As it is attempted in this thesis to generate test instruments assessing pre-service teachers' professional knowledge, this thesis will focus on the university-part of teacher education.

In Germany, most students decide right after school whether they want to pursue a certain subject in order to either become a teacher of that subject or to work in academia or industry. If they decide to aim for a teaching degree, they generally have to select two subjects. The possible combinations of subjects can be up to the student to decide or can be restricted by university regulations. Alongside these two subjects, pre-service teachers have to take several courses in pedagogy. Again, the proportion of each subject and the pedagogy part is defined by the tertiary institution. In almost all teacher training institutions, both subjects and pedagogy courses are studied side by side throughout the course of the curriculum (cf. Figure 6). Several universities have switched to a Bachelor's/Master's structure with the so-called *polyvalent* Bachelor preparing the university students for a variety of specializations in their Master's studies of which the teaching degree is only one. In practice, however, pre-service teachers have to have taken certain courses in their Bachelor's degree specifically

⁵ In a few German states, this phase lasts 24 months.

designed to the teaching specialization in their Master's degree in order to be admitted to a Master's program for teacher education. In their Master's studies, pre-service teachers also have to take courses in pedagogy as well as in both subjects, thus also continuing their subject-specific education. Therefore, in practice, teacher education in Germany follows the concurrent model, as do Denmark, Belgium, the Netherlands and Poland, for example (EURYDICE, 2002). In the opposite approach, the consecutive model, students first complete a degree in their subject(s) and then continue to add some kind of a teaching certification (either offered as an additional degree from the university or a certificate from other institutions). This approach is used in France, Spain and Italy, for example (EURYDICE, 2002). Other European states that provided data offer both approaches. In countries where both models coexist, such as Great Britain, Norway, Sweden and Finland, the actual model used for a particular teaching program sometimes depends on which type of school or level of secondary education future teachers will be teaching at (RISCH, 2010a).

In German higher education institutions, pre-service teachers have to pass a number of content courses for each subject as well as some (considerably fewer) courses focusing on how to teach each subject; in the pedagogy part, studies include content courses in pedagogy, adjacent disciplines such as psychology and teaching practice courses (cf. Figure 6).

For the content part of each subject, pre-service teacher students usually attend courses originally designed for and visited by those students with a one-subject-focus planning to go into academia or industry. As they have to manage two subjects and pedagogy courses while the students with an academia or industry aim have to take only one subject, it goes without saying that teacher students have to take fewer content courses for each subject.

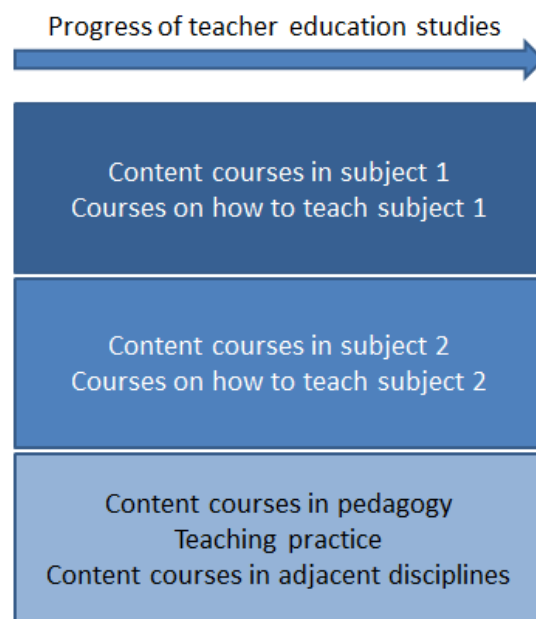
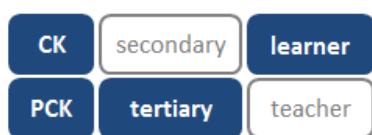


Figure 6: Three-part division in teacher education

2.1.2.2 Aspects related to chemistry



A comparison of universities regarding the possible extent of chemistry content knowledge learning opportunities (KFC, 2014) reveals that in the states of Saarland and Thuringia students in an upper secondary chemistry teacher education program can gain most ECTS-points⁶ in the content-part of their degree while in Bremen students receive the least amount of ECTS-points, with 115 and 72 points, respectively (Figure 7). This general overview shows the difference in time and courses that are offered regarding chemistry content.

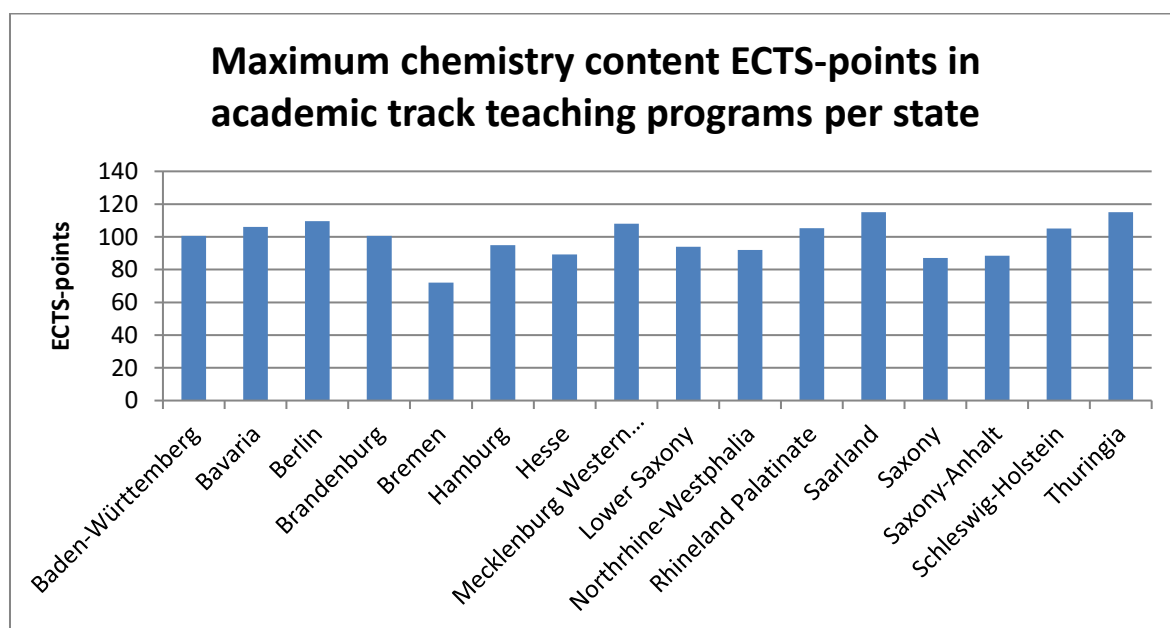


Figure 7: Maximum chemistry content ECTS-points in academic track (Gymnasium) teaching program per state (based on KFC, 2014)

When it comes to the actual content of chemistry courses, the picture is even more varied. As the length, depth and even specializations of any topic covered are up to universities or even individual professors, many experts call for an agreement on certain content aspects. The *Gesellschaft Deutscher Chemiker* (the German Chemical Society with members from academia, education, industry and other areas, acronym GDCH) suggested such a canonical content description for teacher education institutions (GDCH, 2008). An excerpt of these content propositions can be seen in Table 2. Based on the GDCH proposal, the *Secretariat of the Standing Conference of*

⁶ European Credit Transfer System; points are assigned to each course based on the time frame a student needs to spend for a course (including time in lectures, seminars or lab sessions as well as time for preparation and follow-up, preparing for and attending exams, writing a paper or thesis or practical sessions); one point equals a workload of 30 hours.

the Ministers of Education and Cultural Affairs of the Federal States in the Federal Republic of Germany (acronym KMK) issued a resolution specifying contents that teacher education should provide (KMK, 2008).

Table 2: Suggestion regarding content that should be covered by students aiming to teach chemistry; here contents are only listed for lower secondary programs (GDCh, 2008, p.8; KMK, 2008, p.21)

Suggestions for teacher education (GDCh)	Demands for teacher education (KMK)
<p>General and Physical Chemistry</p> <ul style="list-style-type: none"> • Structure of atoms and chemical bonds • Molecular Compounds • Solids • Stoichiometry • States of matter and gas laws • Chemistry in aqueous solutions • Methods of analytical chemistry • Separation procedures • UV/Vis- and IR-spectroscopy • Energy and entropy • Chemical equilibrium • Rate of chemical reactions <p>Inorganic and Organic Chemistry</p> <ul style="list-style-type: none"> • Acid-base concepts and acid-base reactions • Oxidation and reduction, redox-reactions • Chemistry of metals and non-metals • Structure of organic compounds with regard to stereochemistry • Selected organic classes of substances incl. macromolecular compounds • Fundamental types of reactions of chemical compounds • Structure-property-relationships • Structure-function-relationships • Technically important inorganic and organic compounds and methods • Sustainability as basic principle of chemical research and production <p>Teaching Chemistry</p> <ul style="list-style-type: none"> • Teaching-related reflection and structuring of fundamental concepts in chemistry • Teaching-related reflection of scientific methods • Teaching-related research and positions • Approaches and curricula • Learner-oriented experimenting (demonstration experiments and student experiments) • Learning processes, learning difficulties (diagnosis), motivation and interest • School-relevant aspects of the history of chemistry • Sustainable development • Practical phases aiming to investigate and analyze job fields in and out of school, incl. preparation and follow-up 	<p>General Inorganic and Analytical Chemistry</p> <ul style="list-style-type: none"> • Chemistry of the main group elements • Chemistry in aqueous solutions • Molecular and solid chemistry • Chemistry of metals • Methods of analytical chemistry • Spectroscopy • Synthesizing methods / industrial applications • Atoms and bonds • Structure-property-relationships <p>Organic/Biological Chemistry</p> <ul style="list-style-type: none"> • classes of substances, functional groups • Selected reaction mechanisms and syntheses • UV/Vis- and IR-spectroscopy • Natural and synthetic macromolecules • Aromatic compounds • dyes and dying procedures • Basics in metabolism and energy transition • Structure-property-relationships, also in biological chemistry • Sustainability as basic principle of chemical research and production <p>Physical Chemistry</p> <ul style="list-style-type: none"> • Energy and entropy • Rate of reaction and Chemical equilibrium (law of mass action) • Electrochemistry • Kinetic gas theory <p>Teaching Chemistry</p> <ul style="list-style-type: none"> • Teaching-related reflection of fundamental concepts in chemistry • Teaching-related research and positions • Approaches and curricula • Learning processes, diagnosis of learning difficulties, motivation and interest • School-relevant aspects of the history of chemistry • School-oriented experimenting

Even though it becomes apparent that the KMK resolution contains many of the suggestions proposed by the GDCh, both resolutions are quite general and do not reveal anything in regard to how in-depth the aspects should be covered or what specific ability is required of the student. Hence, all universities could well say that they do cover these aspects.

For the courses on chemistry teaching, a more detailed description has been formulated by the *Deutscher Verein zur Förderung des mathematischen und naturwissenschaftlichen Unterrichts* (a German association supporting quality and advances in the school subjects of math and sciences, acronym MNU). In that description, the MNU proposed that teaching-related aspects of – here – chemistry should be offered in three equally-sized modules (Table 3).

Table 3: Suggestions for contents regarding chemistry-education courses (MNU, 2004, p.18)

Module 1 Chemistry-related competencies in reflection and communication	Module 2 Chemistry-educational competencies regarding the fundamentals of teaching	Module 3 Chemistry-educational competencies regarding the teaching and assessment process
<ul style="list-style-type: none"> • Chemistry-educational accompaniment of content lectures, chemistry-educational reflection of fundamental concepts in chemistry, school-relevant aspects of the history of chemistry • Practical phases with the aim of investigating and analyzing employment fields in schools and outside schools, including preparation and follow-up • Techniques for communicating and instructing chemical content according to the targeted group of learners (e.g. techniques of moderation or presenting information or use of media) 	<ul style="list-style-type: none"> • Fundamentals of chemistry-education to turn chemistry teachers into professionals • Chemistry-educational reconstruction of teaching-relevant knowledge and ways to derive at chemical knowledge • School-oriented experimentations (demonstrative experiments and experiments performed by students) 	<ul style="list-style-type: none"> • Selected approaches to theory and research in chemistry education • Chemistry-educational seminar on methods • School practical session with corresponding seminar: selective analysis, trialing and evaluation of teaching

Even though the GDCh and MNU suggestions are not mandatory for universities to follow, the KMK (2008) demands which are intended to be used as a basis for the accreditation and evaluation process of teaching degree programs at all tertiary

institutions are. To what extent (to name just a few facets: depth, length of time, integration into more than one course, level of demands on students) these aspects are likely to be taught at different teacher training institutions cannot be determined easily by looking at the module manuals of each course at any particular university – rather, only numerous qualitative discussions with various lecturers or professors of different institutions might actually reveal the extent to which it is really taught.

Overall, there is a large variety in teacher education programs regarding the actual number of chemistry and teaching chemistry courses covered, regarding the actual contents covered in those courses and regarding the depth or detail of individual content aspects.

2.2 Teachers' professional knowledge

2.2.1 Structure of professional knowledge for teaching



With its division into content courses for each subject, courses on how to teach each subject and pedagogy courses (cf. chapter 2.1.2), teacher education in Germany is structured analogous to an idea that SHULMAN (1986) initially proposed for an ideal teacher's knowledge: He argued that teachers need sound knowledge in at least three main areas belonging to an overarching content knowledge dimension: *Subject Matter Content Knowledge*, *Pedagogical Content Knowledge* and *Curricular Knowledge* (p.9)⁷. For the area of subject matter content knowledge, SHULMAN (1986) requires the teacher to be knowledgeable about different ways of organizing the disciplines' ideas, to have an idea about how knowledge is generated within the discipline and to possess subject-matter knowledge that is "at least equal to that of his or her lay colleague, the mere subject matter major" (p.9). For SHULMAN (1986), pedagogical content knowledge entails those knowledge facets that are necessary to teach a subject: (i) being knowledgeable about possible and proper representations, including analogies, pictures, examples, explanations and demonstrations, (ii) being

⁷ Simultaneously SHULMAN suggests other knowledge types to be important as well, such as "knowledge of general pedagogy, knowledge of learners and their backgrounds, principles of school organization, finance and management, and the historical, social, and cultural foundations of education" (SHULMAN, 1986); he simply does not elaborate on the knowledge types at this point.

knowledgeable about potential difficulties and pre- as well as misconceptions that learners might hold regarding a certain content aspect, and (iii) being knowledgeable about various teaching strategies and using this knowledge to effectively promote learning. By curricular knowledge as an area of content-knowledge, SHULMAN (1986) means that a teacher must know advantages and disadvantages of different ways of incorporating his or her subject into the school context (as could be the differentiation between an integrated science subject or individual science subjects); at the same time a teacher has to be knowledgeable about different ways of assessing student achievement as well as being able to reasonably choose from the array of “materials for [...] instruction, the alternative texts, software, programs, visual materials, single-concept films, laboratory demonstrations, or ‘invitations to enquiry’” (p.10). In addition, being knowledgeable about what their students cover in other subjects at any particular time is also part of curricular knowledge (SHULMAN, 1986). SHULMAN himself afterwards elaborated on these three areas by changing the number and nomenclature of relevant dimensions as well as the re-allocation of different aspects to different dimensions (cf. SHULMAN, 1987). Furthermore, several other propositions have been made to classify teachers' professional knowledge (Table 5). However, the initial three-part division stuck. In many literature reviews (BAUMERT & KUNTER, 2006; BLÖMEKE, KAISER & LEHMANN, 2008; BORKO & PUTNAM, 1996; HELMKE, 2003; LIPOWSKY, 2006; MUNBY, RUSSELL & MARTIN, 2001) as well as in resulting studies on teachers' knowledge in different subjects (e.g. *COACTIV*, *MT21/TEDS-M*, *ProwiN*), however, the three dimensions have been reclassified through general consensus as content knowledge (CK), pedagogical content knowledge (PCK) and pedagogical knowledge (PK).

Table 4 presents only a broad overview. A more detailed analysis of content knowledge and pedagogical content knowledge will be presented in the respective chapters on CK and PCK (chapters 2.3.1 and 2.3.2, respectively). As the development of a PK test has not been part of this thesis but has been contributed by a different institution, this dimension will not be further elaborated on in this thesis.

2 THEORETICAL FRAMEWORK

Table 4: Overview of some exemplary models of teachers' professional knowledge*

LEINHARDT AND SMITH, 1985	SHULMAN, 1986	SHULMAN, 1987	GROSSMAN, 1990	GESS-NEWSOME, 1999	BAUMERT & KUNTER, 2006
<i>Subject-Matter Knowledge</i>	<i>Subject-Matter-Knowledge</i>	<i>Content Knowledge</i>	<i>Subject-Matter Knowledge</i>	<i>Content Knowledge</i>	<i>Content Knowledge</i>
--	<i>Pedagogical Content Knowledge</i>	<i>Pedagogical Content Knowledge</i>	<i>Pedagogical Content Knowledge (knowledge of students' understanding, curriculum, and instructional strategies)</i>	--	<i>Pedagogical Content Knowledge</i>
<i>Lesson Structure Knowledge (which includes planning and running a lesson smoothly, as well as clear explanations)</i>	--	<i>Knowledge of learners and their characteristics</i>	--	--	--
--	--	<i>General Pedagogical Knowledge</i>	<i>General Pedagogical Knowledge</i>	<i>Pedagogical Knowledge</i>	<i>Pedagogical /Psychological Knowledge</i>
--	<i>Curricular Knowledge</i>	<i>Curriculum Knowledge</i>	--	--	<i>Organizational Knowledge</i>
--	--	--	--	--	<i>Counseling Knowledge</i>
--	--	<i>Knowledge of educational contexts</i>	<i>Knowledge of Context</i>	<i>Knowledge of Context</i>	--
--	--	<i>Knowledge of educational ends</i>	--	--	--

*Even though not all sub-dimensions – not even those of the same name – contain the same aspects, a general grouping into content knowledge (dark blue), pedagogical content knowledge (medium blue) and pedagogical knowledge (light blue) can be made for most approaches to define the cognitive knowledge base of teachers.

A recent and prominent model developed in the German research context and including a widely-used classification of teachers' professional knowledge is the *COACTIV* (acronym of *Cognitive Activation in the Classroom*) project's generic structure model of professional competence (BAUMERT & KUNTER, 2013). The model is presented in Figure 8 and shows that professional knowledge is only one area in the larger competence model. In *COACTIV*, competence is understood along the lines of KLIEME, HARTIG & RAUCH (2008) based on WEINERT (2001b) meaning a "personal capacity to cope with specific situational demands [... for which many aspects, in particular self-regulation and professional knowledge, are ...] learnable and teachable [in university courses]" (BAUMERT & KUNTER, 2013, p.27). According to KLIEME ET AL. (2008) and hence also for BAUMERT AND KUNTER (2013), competence not only incorporates cognitive skills, but also includes motivational aspects and personal attitudes (cf. Figure 8).

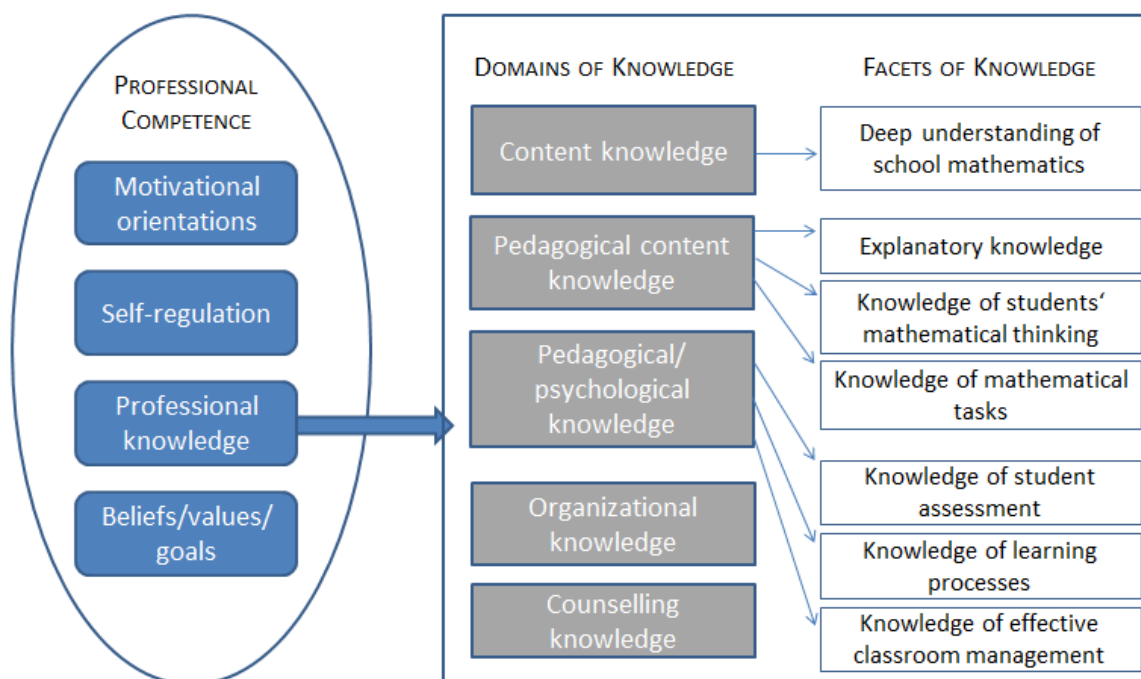


Figure 8: *COACTIV* generic structure model of professional competence (BAUMERT & KUNTER, 2013, p.29)

Competencies and therefore all facets of professional knowledge can be learned in teacher education or teacher training (BAUMERT & KUNTER, 2013). Even though originally, this model has been generated for teaching mathematics, it has been used as the basis for research questions in other disciplines and for different approaches as well (see among others BLÖMEKE ET AL., 2008; DICKE, PARKER, HOLZBERGER, KUNINA-

HABENICHT, KUNTER & LEUTNER, 2015; GOLD & HOLODYNSKI, 2015; MATTERN & BAUER, 2014; RIESE & REINHOLD, 2012; SEIDEL & PRENZEL, 2008; TEPNER ET AL., 2012; WINTHER, 2010).

Even though for the expertise approach⁸, knowledge and metacognitive knowledge⁹ learned in academic settings are necessary as well (BROMME & HAAG, 2008), and “area-specific declarative and procedural knowledge are a prerequisite for being successful as a teacher” (BROMME & HAAG, 2008, p.806f.), the characteristics that a teacher possesses are seen to be equally important. The individual characteristics that turn a person into an expert are not definite, however, and vary depending on researchers or contexts (BROMME, 1997). There are several combinations of teachers’ characteristics that can be agreed upon (BROMME & HAAG, 2008):

- a) Experts are said to have a categorical perception of teaching situations. This means that they hold concepts of teaching situations, generally view the class as an entity, are responsive to those things which are relevant for the teaching process and grasp more complex scenarios when analyzing a teaching situation.
- b) Experts have a stable set of goals and pursue them in a flexible manner. In order to be able to do this, experts know of a variety of goals and can adequately select from their repertoire of these teaching goals in each teaching situation.
- c) Experts exhibit routinized as well as flexible actions. These actions are domain-specific and are based on prior experience and are then used in a flexible way depending on the situation.

These characteristics of an expert teacher have been found in empirical studies to facilitate student success (BROMME & HAAG, 2008).

These two approaches can be reconciled, however. BROMME AND HAAG (2008) stress the relevance of initial subject-specific knowledge, just as BAUMERT AND KUNTER (2013) assume professional knowledge to be learned, the latter being explicitly built on the expertise research and extending it by adding motivational orientations and

⁸ an approach differentiating competencies of experienced, thus expert, teachers from those of inexperienced, thus novice, teachers.

⁹ Defined by PINTRICH (2008, p.219) as follows: “Metacognitive knowledge includes knowledge of general strategies that might be used for different tasks, knowledge of the conditions under which these strategies might be used, knowledge of the extent to which the strategies are effective, and knowledge of self (Flavell, 1979; Pintrich et al., 2000; Schneider & Pressley, 1997).”

self-regulatory skills (p.41). The first two characteristics elaborated on by BROMME AND HAAG (2008), for example, would fall exactly into the area of beliefs/values/goals as proposed in the *COACTIV* model of professional competence. While the *COACTIV* model emphasizes changeability through learning and experience of even these beliefs, BROMME AND HAAG (2008) do not say how these characteristics come to be but simply state that they are there. The third characteristic mentioned by BROMME AND HAAG (2008) is that of experience in the teaching profession and as it is based on prior experience, it must also be learned. In the *COACTIV* model, the actions by a teaching professional are not included in the model directly but rather are a consequence of all aspects of professional competence.

Knowledge acquired during teacher education, however, does not in itself constitute an unchanging knowledge base that teachers can tap into. MESSNER AND REUSSER (2000), for example state that the notion of obtaining 'fully educated teachers' after university education plus practical teaching phase has to be given up. For one, it is reported that teachers leaving tertiary institutions and just beginning their own teaching often feel pressured by, among other things, the complexity of an authentic teaching situation (MÜLLER-FOHRBRODT, CLOETTA & DANN, 1978; or TYNJÄLÄ & HEIKKINEN, 2011, among others) and resort to a rather ineffective teaching style that might be more similar to their own experience as secondary students than to knowledge aspects learnt at university (BROUWER, 2010; LAMPERT & BALL, 1999). Some studies have attempted to analyze how teacher knowledge develops from leaving university to teaching at school: Based on actual teaching experience, teachers' content knowledge changes over time (e.g. ARZI & WHITE, 2008; BAUMERT & KUNTER, 2013; HASHWEH, 2005; HIEBERT, GALLIMORE, & STIGLER, 2002), often due to the school curriculum and context. These changes can both include a negative effect such as forgetting certain content aspects and thus narrowing teachers' content knowledge base; changes can, however, also lead to a more profound understanding of certain content aspects as well as an improved conceptual understanding (ARZI & WHITE, 2008). According to ARZI AND WHITE (2008), this shows that teachers' everyday surroundings and experiences influence their content knowledge and understanding more substantially than content knowledge courses at university. ARZI AND WHITE (2008) state that this shaping process of a teacher's subject matter knowledge takes years of experience. Regarding other parts of professional knowledge apart from content knowledge, fewer studies exist. BLÖMEKE ET AL. (2014) analyzed the content

knowledge and pedagogical content knowledge for math teachers in the last year of their formal education and again four years later to find out whether a development occurred. Expectations had been diverse as researchers had previously hypothesized either that pedagogical content knowledge would increase due to practical experiences while content knowledge would remain relatively steady (LLINARES & KRAINER, 2006; SCHOENFELD & KILPATRICK, 2008) or that no significant changes should result from practical experiences alone (KRAUSS & BRUNNER, 2011). Over the length of these four years, BLÖMEKE ET AL. (2014) found a significant decrease in content knowledge with a small effect and no significant change in pedagogical content knowledge with those teachers showing the highest test results at the end of their education also remaining the teachers with the highest test results later on in content knowledge; this trend being less stable for pedagogical content knowledge. Content knowledge at the end of the teachers' education mediated via that four years later proved to be a strong predictor for the speed of grasping student errors and for teachers' perception and interpretation of and decision-making in relevant situations. For pedagogical content knowledge, a direct effect of teachers' knowledge at the end of their education could be shown on the speed of grasping student errors while for teachers' perception and interpretation of and decision-making in relevant situations this knowledge was predictive when mediated via that measured three years into the teaching profession. From its significant effects on indicators representing situational and behavioral aspects, BLÖMEKE ET AL. (2014) suggest a situation-specificity of professional teacher knowledge.

Furthermore, researchers have also looked at assessing experienced teachers' knowledge of teaching in order to find beneficial aspects. MULHALL, BERRY AND LOUGHRAN (2003) proposed a model of CoRes and PaP-eRs: CoRes stands for Content Representations and provide "an overview of the particular content taught when teaching a topic" (p.17) while PaP-eRs, i.e. Pedagogical and Professional-experience Repertoires, are accounts of teachers specifying the CoRes in descriptions of teaching in a particular classroom setting. For MULHALL ET AL. (2003), teacher knowledge is therefore not only content-specific but also context- or topic-specific. This topic-specificity is also stressed by ROLLNICK, BENNETT, RHEMTULA, DHARSEY AND NDLOVU (2010) and GESS-NEWSOME AND CARLSON (2013). AUFSCHNAITER AND BLÖMEKE (2010) state that literature on teacher competences often lacks precise criteria that allow the researchers to clearly differentiate between general pedagogical and pedagogical

content knowledge and that many facets of general pedagogical knowledge become visible only in the context of teaching the subject.

Before looking more closely into pre-service chemistry teachers' understanding, several models of content knowledge and pedagogical content knowledge shall be investigated. As pre-service teachers arrive at university with a certain content knowledge acquired in school which will then be built on during their tertiary education, models describing school students' content knowledge will be looked at first.

2.2.2 Empirical competence models for students

As the test subjects in this project will be pre-service teachers, before looking at models of teacher competence, thus looking at the more or less final conglomerate of teachers' professional knowledge, it is also interesting to look at models of secondary students' competence, and thereby looking at the starting point of a potential pre-service teacher's knowledge. The first international competence models have been proposed based on *TIMSS* (*Third International Mathematics and Science Study*) or *PISA* (*Programme for International Student Assessment*) data. According to SCHIEPE-TISKA, SCHÖPS, RÖNNEBEK, KÖLLER AND PRENZEL (2013), *PISA* investigates three sub-competencies (Figure 9) to describe science competence of 15-year old students: (i) recognizing scientific questions, (ii) explaining scientific phenomena, and (iii) using scientific evidence. These competencies are assessed in various contexts and necessitate students to activate either type of knowledge: i) knowledge of sciences¹⁰ or ii) knowledge about sciences¹¹. Along with the knowledge types, motivational orientations form the basis for the development of the sub-competencies.

¹⁰ i.e. subject-specific content knowledge learned in class

¹¹ i.e. subject-transcending knowledge of how knowledge is gained in the sciences

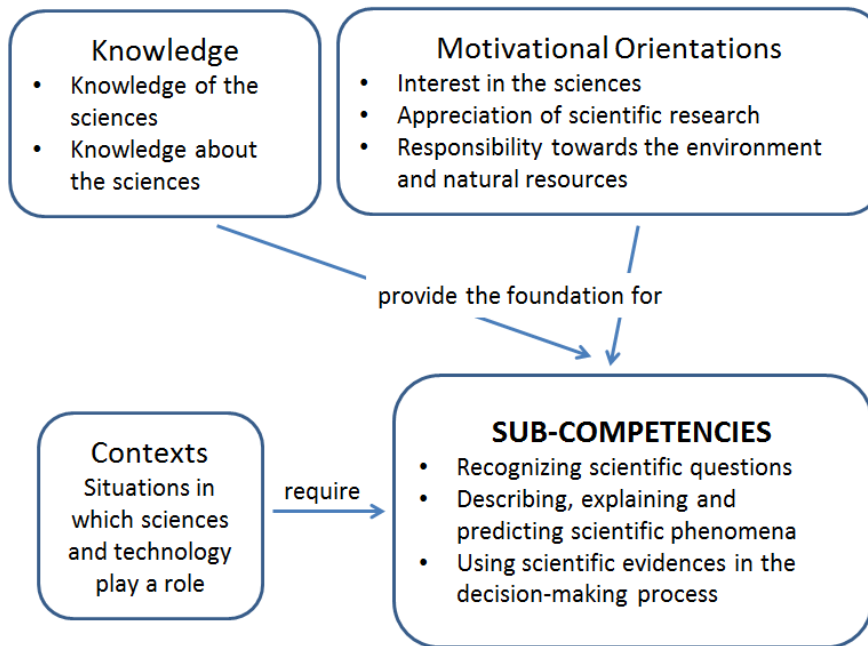


Figure 9: PISA framework for scientific literacy, containing sub-competencies for the sciences (SCHIEPE-TISKA ET AL., 2013)

Even though both the *TIMSS* and *PISA* models are not considered very beneficial for the science research context as they were formulated in retrospect, i.e. after the actual assessment in order to fit the items used and data obtained (cf. NEUMANN, KAUERTZ, LAU, NOTARP & FISCHER, 2007), they provide first insights into the formulation of competences. A number of empirically investigated models concerning students' competences have also been developed in the German research context after these large-scale assessments and after the introduction of national education standards for the sciences in Germany. Many of these models have been developed at about the same time and have used slightly different approaches regarding which facets should be taken into consideration when making assessments of students' competence. As the development of a model normally progresses through several stages, the beginning usually is a theoretical conceptualization of the content to be investigated while the end usually is the empirical validation of a model (MAYER & WELLNITZ, 2014). All of the following models evolving in the German research context have undergone this model development process.

The model by SCHECKER AND PARCHMANN (2006) contains *competence areas* of the NES, *subject specificity* and *content areas or fundamental concepts* (cf. Figure 10). In addition, this model also includes a facet called *context*, in which each task is assigned to a situation of being either solely within-subject or of including aspects belonging to the personal or societal surroundings or of applying to professional life.

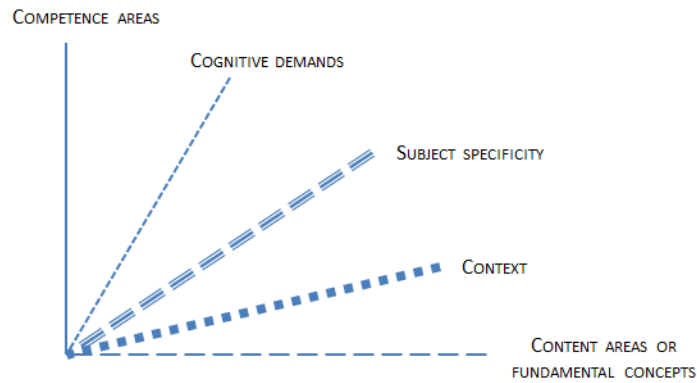


Figure 10: Facets in competence model according to SCHECKER & PARCHMANN (2006)

Also, the facet *cognitive demands* is added. Here it is specified what the students have to do in the respective task. All facets are described in more detail in Table 5.

Table 5: Facets of SCHECKER & PARCHMANN model (2006)

Competence area	Cognitive demands	Subject-specificity	Context	Content areas or fundamental concept
Evaluating facts & situations regarding scientific aspects	Divergent thinking	Every-day environment	Situation within subject	Matter
Communicating within & about sciences	Convergent thinking	Subject-specific, reproducing	Including personal or societal surroundings	Energy
Approaches to arriving at scientific data and knowledge	Applying mental models	Actively applying	Applying to professional life	(Thermodynamics)
Dealing with content knowledge	Applying numbers	Conceptual understanding		...

This model is quite extensive in describing five facets of students' competences. Most researchers who try to empirically verify their competence models, however, deliberately try to keep it simpler by using fewer facets.

For chemistry, the Model of Hierarchical Complexity for Chemistry (*MHCC*) suggests that for each topic or fundamental concept, tasks can be generated and assessed which differ systematically in their complexity. For school chemistry, BERNHOLT (2010) proposed five increasing levels of complexity: (1) everyday experiences, (2) facts, (3) processes, (4) linear causality, and (5) multivariate interdependencies. Declarative knowledge is considered to be integrated into a net-like structure and the degree of competence increases as more facts and eventually relations are integrated into and finally interconnected within this net. The model has been trialed for the contents of combustion, acids and bases as well as redox-reactions (BERNHOLT, WALPUSKI, SUMFLETH & PARCHMANN, 2009). All items or tasks can also be assigned to one or more of the NES as specified by the description of German educational standards (KMK, 2004). The standards are then grouped into broader areas that students should become competent in: (i) subject knowledge, (ii) epistemological knowledge and subject methods, (iii) communication, as well as (iv) judgement (SCHECKER & PARCHMANN, 2007).

A common model for describing students' content knowledge in all sciences was developed by a project that intended to evaluate the science standards in secondary schools; the German project *Evaluation der Standards für die Naturwissenschaften in der Sekundarstufe I*, acronym *ESNaS*. In this model, competence is divided into three facets (Figure 11), with *cognitive processes* describing what students do with presented knowledge aspects, *complexity* describing whether the knowledge aspects require students to deal with facts or relations or the overall concept, and *competency areas* relating to the German NES (WALPUSKI ET AL, 2010):

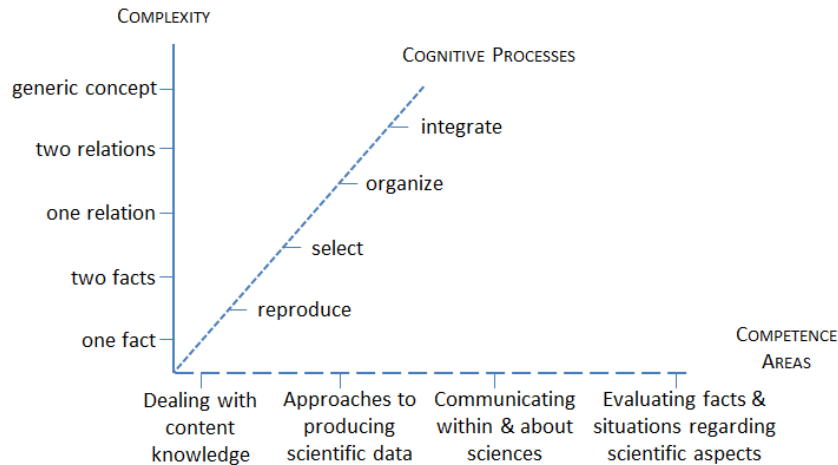
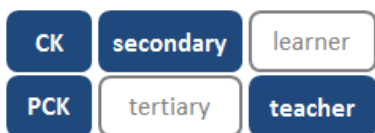


Figure 11: Competence model for all sciences according to ESNaS (according to WALPUSKI ET AL., 2010)

As this model is to be used for all sciences, it can be considered for each fundamental concept in each of the sciences. Fundamental concepts constitute different angles of looking at a variety of topics, can overlap to a limited degree and are an integral aspect of spiral curricula (cf. BRUNER, 1977) in which concepts are revisited and depth is increased. A certain fundamental concept can be but does not necessarily have to be more difficult than a second one. Complexity is used somewhat similar to the MHCC model as a way to structure each content aspect regarding the number and integration of various aspects necessary for answering a task.

2.2.3 Competence models for (pre-service) teachers



While the competence models described so far defined secondary students' competencies focusing largely on content knowledge, they only provide an implicit idea of a future teacher's content competence. When one is interested in a broader picture of teachers' competencies beyond the content knowledge dimension, explicit models on teachers' competence have to be considered. Many of these models have been generated in various mathematic education research projects: *COACTIV* analyzing German secondary teachers' competence, *TEDS-M* (acronym for *Teacher Education and Development Study*) and its pre-study *MT21* (acronym for *Mathematics Teaching in the 21st Century*) focusing on pre-service secondary teachers' competence in

various countries, and projects by the Michigan group around DEBORAH BALL investigating elementary teachers' competence. Apart from mathematics, researchers in other subjects have also set up competence models, such as the *ProwiN* group (acronym for the German project title translating to *Professional Knowledge of Teachers in the Sciences*) for various sciences or RIESE AND REINHOLT (2010) for physics.

MODELS IN MATHEMATICS

In *COACTIV*, for example, in order to examine the mathematical understanding of teachers, content knowledge tasks were mainly designed to assess teachers' theoretical knowledge in this dimension (BAUMERT & KUNTER, 2013). That is why content knowledge has been conceptualized in the following way (Figure 12):

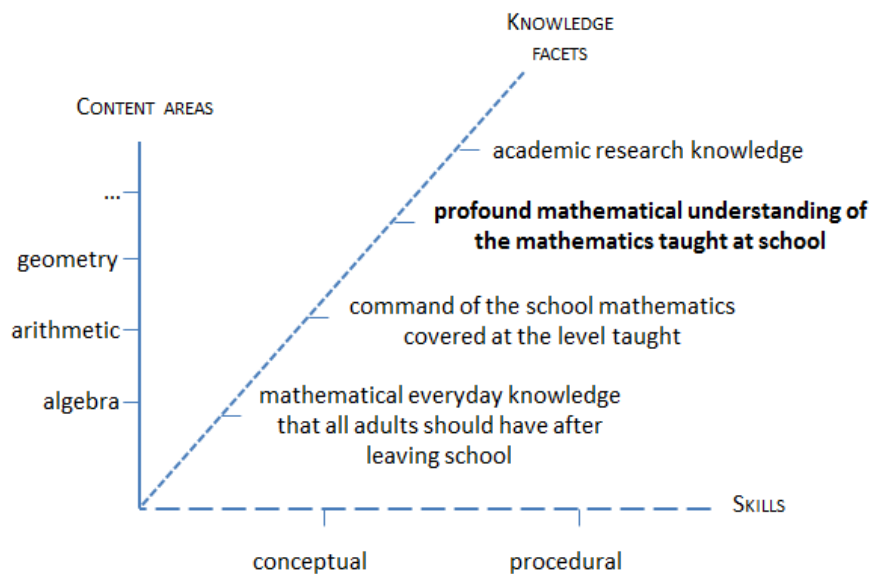


Figure 12: Content sub-dimension in *COACTIV* (KRAUSS ET AL., 2008; bold print is added to stress the exclusive focus of task generation on a single knowledge facet)

As tasks have only been formulated at one level of knowledge facets (bold print in Figure 12) for the *COACTIV* test, and a resulting analysis did not show separable sub-dimensions, content knowledge was eventually considered not to include any sub-dimensions (KRAUSS ET AL., 2008). For pedagogical content knowledge of secondary math teachers, the PCK dimension was conceptualized to assess both

theoretical as well as more practical classroom knowledge and consisted of three sub-dimensions (BAUMERT & KUNTER, 2013, see Table 6).

Table 6: PCK sub-dimension according to COACTIV model (BAUMERT & KUNTER, 2013)

PCK sub-dimensions
Tasks: Knowledge of the didactic and diagnostic potential of tasks, their cognitive demands and the prior knowledge they implicitly require, their effective orchestration in the classroom, and the long-term sequencing of learning content in the curriculum
Student cognitions: Knowledge of student cognitions (misconceptions, typical errors, strategies) and ways of assessing student knowledge and comprehension processes
Representations and explanations: Knowledge of explanations and multiple representations

The Michigan group around BALL AND HILL formulated a model for practicing elementary mathematics teachers of three content areas, i.e. numbers/operations, patterns/functions, and algebra, as well as four knowledge domains as presented in Table 7 (e.g. HILL, SCHILLING & BALL, 2004; BALL, THAMES & PHELPS, 2008).

Table 7: Descriptions of knowledge domains according to BALL ET AL. (2008)

Knowledge Domain	Description
common content knowledge (CCK)	mathematical knowledge and skill that is used in settings other than teaching, such as every-day life or other professions
specialized content knowledge (SCK)	mathematical knowledge and skill that is unique to teaching and not required in other settings, such as why certain operations yield a specific answer or which story problem corresponds to a certain calculation
knowledge of content and students (KCS)	knowledge that combines knowing about students and knowing about mathematics, such as anticipating students thinking and problems in understanding, predicting interest and motivation, assessing student understanding as correct or incomplete/wrong concept
knowledge of content and teaching (KCT)	knowledge that combines knowing about teaching and knowing about mathematics, such sequencing particular content for instruction, choosing appropriate examples or reacting flexibly in a classroom situation while teaching a certain math content

When looking at the knowledge domains, one can see that these include two types of content knowledges as well as two types related to the practical classroom teaching. BALL ET AL. (2008) concede that the knowledge domains had been formulated from a practitioner's view¹², described different foci of mathematics

¹² HILL ET AL. (2004) and SCHILLING (2007) state that the original knowledge domain structure could not be replicated empirically; rather, a three-factor solution emerged with two CK factors (knowledge of patterns, functions, and

knowledge and that two domains (KCS and KCT) might be understood to refer to two PCK-aspects as proposed by SHULMAN (1986): student cognition and instructional strategies.

The international comparative study “Learning to Teach Mathematics: Teacher Education and Development Study” (*TEDS-M*, cf. BLÖMEKE, SUHL & KAISER, 2011) based their items on both the work of the Michigan group and on the study “Teaching Mathematics in the 21st century” (*MT21*, cf. SCHMIDT ET AL., 2011) and conceptualized content knowledge by proposing a cognitive facet consisting of knowing, applying and reasoning, a difficulty facet ranging from novice to expert, and a content area comprising number, geometry, algebra and data.

In a very analogous way, items for pedagogical knowledge in the *TEDS-M* study were designed with the two latter facets being identical to the CK model and with teaching sub-dimensions instead of the cognitive facet (Figure 13).

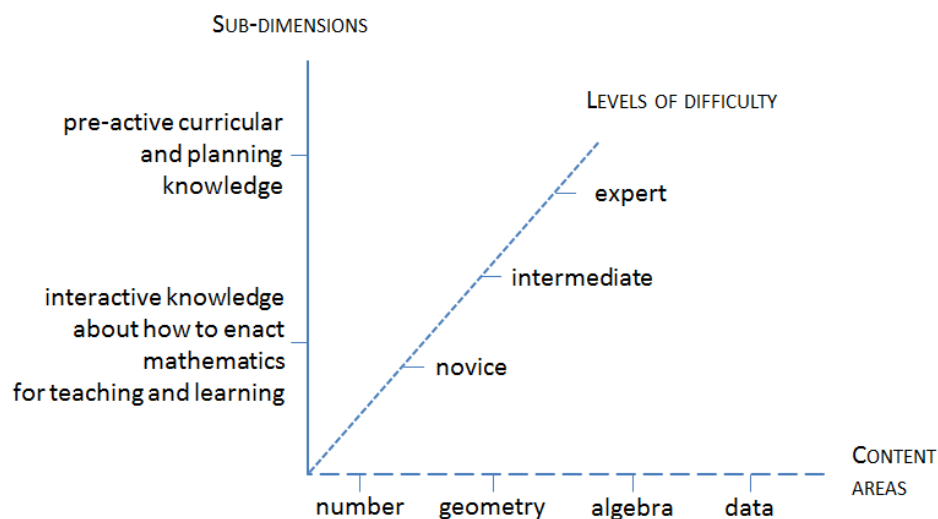


Figure 13: Competency model of mathematical PCK according to TEDS-M study (Blömeke et al., 2011)

algebra as well as covering knowledge of number concepts and operations) and one PCK factor (knowledge of students and content).

Models in Physics

Other research projects have proposed and validated teacher competence models for other subjects, such as RIESE (2009) for physics and the project group of *ProwiN* for each science.

RIESE (2009) also proposed three facets for his model describing competences of physics teachers: content areas, cognitive activities and levels (Figure 14).

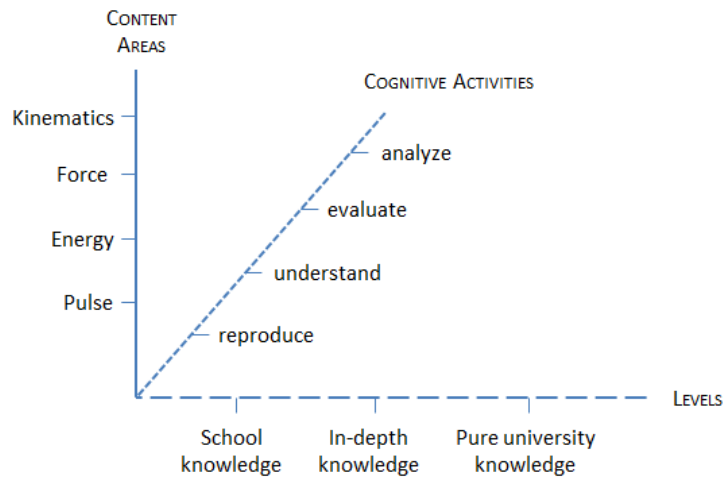


Figure 14: Competency model according to RIESE (2009)

The unique aspect of the RIESE (2009) model is the level called *in-depth knowledge*. It characterizes a profound and conceptual understanding as well as knowledge of multiple approaches to solving content knowledge questions that very good students might have already gained at the end of secondary education.

For pedagogical content knowledge, RIESE (2009) argues that the sub-dimensionality proposed for content knowledge cannot be transferred directly (p.80) and, therefore, PCK is only to contain one dimension in which the individual facets are somewhat ordered according to the facets' increasing complexity (Table 8).

Table 8: Facets of the PCK dimension according to RIESE (2009)

PCK facets
Knowledge about (general) aspects of learning processes in physics
Knowledge about using experiments
Designing and planning learning processes
Evaluating, analyzing, and reflecting on learning processes
Adequate reactions in critical and unexpected teaching situations

Models in Chemistry

As the research project *ProwiN* aims to analyze teachers' professional knowledge and its consequences on the classroom, a joint model for the three sciences was constructed on a theoretical basis encompassing content knowledge, pedagogical content knowledge and pedagogical knowledge (TEPNER ET AL., 2012). For the chemistry part, the model for the content knowledge dimension (Figure 15¹³) as well as the pedagogical content knowledge dimension (Figure 16) were conceptualized as follows (DOLLNY, 2011).

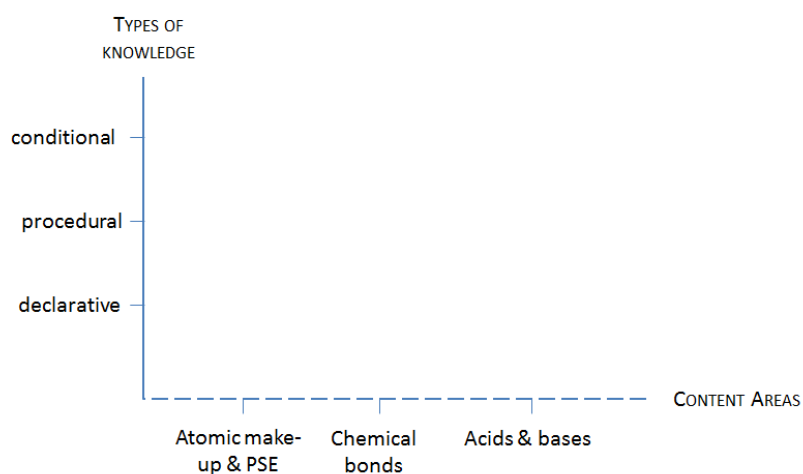


Figure 15: Competence model for chemistry teachers' CK according to *ProwiN* (DOLLNY, 2011)

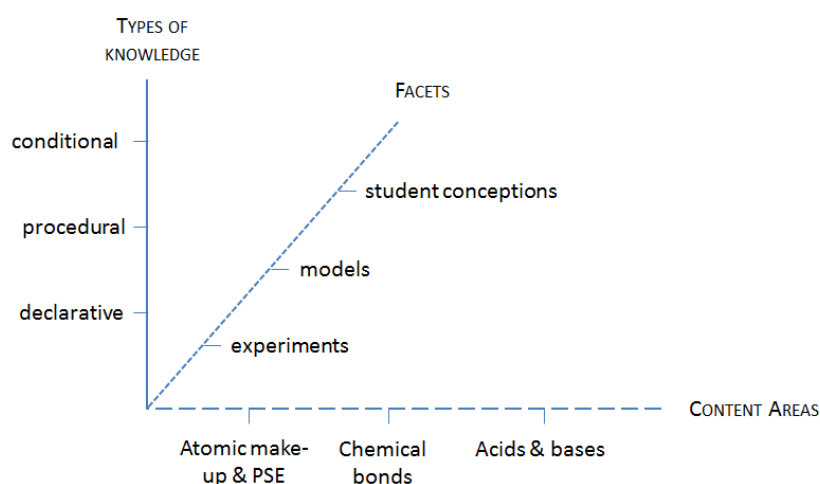


Figure 16: Competence model for chemistry teachers' PCK according to *ProwiN* (DOLLNY, 2011)

¹³ The types of knowledge are classified by PARIS, LIPSON AND WIXSON (1983) into declarative knowledge, i.e. knowing *that* independent facts exist, procedural knowledge, i.e. knowing *how* to perform a task, and conditional knowledge, i.e. knowing in which circumstances and *under which conditions* certain actions should be performed.

When comparing the content knowledge dimension to the pedagogical content knowledge dimension, it becomes apparent that both underlying test instruments focus on the same topics, a necessity when aiming to eventually combine results from both competence models as stressed by WALPUSKI, TEPNER, SUMFLETH, DOLLNY, HOSTENBACH AND POLLENDER (2012).

Structure of Professional Knowledge

Apart from investigating individual dimensions or criteria that make up professional knowledge, researchers have also provided empirical data of how these dimensions are related. This has been done in particular for the three-part division described above: CK, PCK and PK. In the *COACTIV* project, the CK and PCK dimension of practicing math teachers were found to correlate to a large degree (KRAUSS ET AL., 2013, p.161), while the PCK and PK dimensions correlate to a lesser degree and the correlation between the CK and PK dimension is the smallest (VOSS & KUNTER, 2013).

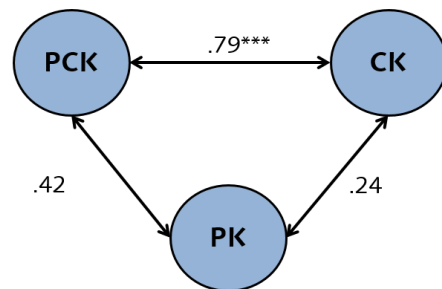


Figure 17: Correlations between latent dimensions in *COACTIV*

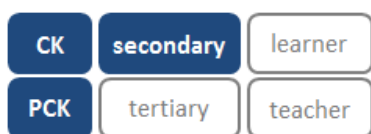
Figure 17 presents the latent¹⁴ correlations between the constructs. While the exact correlation values

differ between researchers, the general tendencies are also shown in the *TEDS-M* group for German secondary math teachers: CK and PCK show a manifest correlation of $r = .70$ (BLÖMEKE, KAISER, DÖHRMANN & LEHMANN, 2010, p.212), while PCK and PK correlate with $r = .30$, and CK and PK show the smallest of the three correlations of $r = .29$ (BLÖMEKE & KÖNIG, 2010, p.275). In the *ProwiN* project, correlations between physics teachers' professional knowledge dimensions show the same tendency: CK and PCK correlate with $r = .45$, PCK and PK¹⁵ correlate with $r = .26$ and CK and PK correlate with $r = .17$ (KIRSCHNER, 2013, p.83). In a later article, KIRSCHNER ET AL. report the following correlations between the CK and PCK dimensions: $r_{\text{PHY}} = .58$, $r_{\text{BIO}} = .29$, $r_{\text{CHEM}} = .42$ (KIRSCHNER ET AL., 2016, p.122).

¹⁴ The term *latent* indicates that latent, underlying and not observable constructs according to item response theory (cf. chapter 5.3) have been used here.

¹⁵ Only part of the PK dimension (i.e. the declarative part) is considered in the structural analysis.

2.3 Structure-Property-Relations (SPR)



Before going into the description of the test development and on to the results, it is necessary to explain briefly what is meant by the concept of structure-property-relations, in the following referred to as SPR, and why particularly this concept has been chosen.

For future teachers the content area evaluated should apply to both secondary as well as tertiary education and include a wide variety of chemistry topics. With this in mind, choosing one of the fundamental concepts proposed in the German National Education Standards for the secondary level (KMK, 2005) made sense. This approach shall be elaborated on in the following paragraphs.

Table 9: Short overview of fundamental concepts for lower secondary education (KMK, 2005)

Basic concept	Exemplary competencies students should acquire
Particulate nature of matter	<ul style="list-style-type: none"> name and describe relevant substances with their typical properties describe the submicroscopic makeup of selected substances describe the makeup of atoms by means of a suitable atomic model use bond models for interpreting aggregates of particles, spatial structures, and intermolecular forces explain the variety in substances due to different combinations and arrangements of particles
Structure-property-relations	<ul style="list-style-type: none"> describe and explain ordering principles of substances, e.g. based on their typical properties or based on typical characteristics of the composition and structure of particles use an appropriate model for interpreting the properties of a substance at the particle level deduce possible applications of a substance as well as advantages/disadvantages based on the properties of that substance
Chemical reactions*	<ul style="list-style-type: none"> describe phenomena of substance conversion and energy transformation during chemical reactions interpret substance conversion and energy transformation as a changing in particles and rearrangement of chemical bonds indicate the transfer of particles in selected donor-acceptor reactions and classify the type of reaction set up reaction equations by using their knowledge of atom conservation and formation of compounds based on constant atomic combination ratios describe that chemical reactions are reversible describe possibilities to control chemical reactions due to a variation in reaction conditions
Energetic considerations	<ul style="list-style-type: none"> state that during chemical reactions the energy content of the reaction system also changes due to an exchange with the environment ascribe energetic phenomena during chemical reactions to a change of part of the energy within a substance into other forms of energy describe that chemical reactions can be influenced by using catalysts

*For upper secondary, the fundamental concept *chemical reactions* is often subdivided into further concepts: donor-acceptor reactions, chemical equilibrium, and reaction rate (NENTWIG ET AL., 2007).

In Germany, the education standards for chemistry are made up of fundamental concepts, with structure-property-relations being one of them (KMK, 2005; NENTWIG ET AL., 2007; cf. Table 9). This idea of fundamental concepts is not unique to Germany, however, as many countries have proposed standards of what students should know or be able to do at the end of certain educational levels (WADDINGTON, NENTWIG & SCHANZE, 2007) and many of those contain concepts and ideas equivalent to structure-property-relations. In the US for example, the Next Generation Science Standards (NGSS) contain three structural elements: science and engineering practices, core ideas, and cross-cutting concepts (NATIONAL RESEARCH COUNCIL, 2012). These cross-cutting concepts are to specify topics or constructs that provide a unifying approach across the sciences. Even though it is not explicitly mentioned as a separate cross-cutting concept in the Next Generation Science Standards, two of the cross-cutting concepts proposed therein relate quite closely to the concept of SPR: i) *cause and effect*, in which students, among other things, “use cause and effect relationships to predict phenomena in natural or designed systems [... or ...] propose causal relationships by examining what is known about smaller-scale mechanisms within the system” (NGSS, 2013, p.83) and ii) *structure and function*, in which students “model complex and microscopic structures and systems and visualize how their function depends on the shapes, composition, and relationships among its parts [... or ...] infer the functions and properties of natural and designed objects and systems from their overall structure, the way their components are shaped and used, and the molecular substructures of their various materials” (NGSS, 2013, p.87). Even though the NGSS cross-cutting concepts present an idea to incorporate all sciences, SEVIAN AND TALANQUER (2014) propose *structure-property-relationships* as one of their chemistry-specific cross-cutting concepts: “Making decisions about what type of substance to synthesize, what reactant to use, or what detection technique to utilize critically depends on the understanding of the relationship between properties and structure at different scales” (p.13).

Even though the cross-cutting concepts in the US NGSS are intended to interrelate understanding of the sciences, while the fundamental concepts in Germany are used for individual sciences, their intentions are comparable: the US standards are “introducing unifying ideas to guide student thinking” (NGAI, SEVIAN & TALAQUER, 2014, p.2439), while the German standards are meant to i) facilitate students’ acquisition of fundamental and integrated knowledge, ii) favor the

systematic and cumulative development of subject-specific competencies, and iii) form the foundation for interdisciplinary linking of knowledge (cf. KMK, 2004, p.8f.).

Several other considerations have led to the selection of particularly this fundamental concept. These shall be presented in the following paragraphs. First, aspects relating clearly to the content knowledge area will be regarded (cf. chapter 2.3.1): As SPR lends itself to looking at various topics through one consistent lens, it is prone to improve students' levels of thinking and can therefore strengthen students' ability to derive at chemically sound explanation processes. Another decisive factor for choosing the SPR concept as the basis for assessing pre-service teachers' understanding is its transferability to the tertiary level. The last reason for deciding on the SPR concept is that numerous problems in connection with this concept have been reported at both school and university level. As knowledge about students' problems and potential pitfalls within a specific content topic is a vital part of PCK, however, this last aspect will be described in the sub-chapter on pedagogical content knowledge (cf. chapter 2.3.2) which will follow the CK sub-chapter.

2.3.1 Content knowledge (CK)



In this chapter, the content-related aspects of the SPR-concept shall be addressed. In literature, content knowledge, in the following referred to as CK, can either denote knowledge learnt in school or knowledge learnt in content courses at university. This differentiation usually depends on the research project cited which in turn is often based on the population sample investigated in that particular project. For *COACTIV*, for example, content knowledge consists of knowledge aspects learnt from the perspective of the discipline, i.e. academic content knowledge, as the project assessed pre-service mathematics teachers (cf. BAUMERT & KUNTER, 2013) while *ProwiN*, for example, specifies CK as content knowledge conveyed at (advanced) school level (cf. WALPUSKI ET AL., 2012). Sometimes, regardless of the reference to either school-learnt or academic knowledge, content knowledge is also termed subject-matter knowledge (SMK, cf. ARZI & WHITE, 2004; COCHRAN & JONES, 1998; or KIND, 2009, for example). As will be explained later, no such distinction is made for the

purpose of this thesis and CK is defined as containing both knowledge aspects relevant for school and/or university level.

2.3.1.1 Topic-transcending consistent perspective

As all fundamental concepts have been designed to provide a certain consistent perspective with which to look at a variety of topics, a short selection of examples for SPR is presented in Table 10. One can see that, as the name of the SPR concept implies, the consistent perspective is that secondary students¹⁶ are expected to describe, explain and deduce properties based on the underlying structure of a particular compound.

Table 10: Selection of examples of how the SPR concept provides consistent chemical perspective (own translation)

Traditional topics	Exemplary aspects of SPR in curricula
structure	<ul style="list-style-type: none"> students are to deduce the specific properties of substances based on the structure of the molecules (FA SH, 2016*, 43)
molecules	<ul style="list-style-type: none"> students are to explain the specific properties of substances with a molecular setup based on intermolecular forces (FA SH, 2016*, 43)
classes of substances	<ul style="list-style-type: none"> students are to differentiate between classes of substances according to their properties and structures and deduce possibilities for their use (KC NDS, 2009**, 58)
reaction types	<ul style="list-style-type: none"> students are to differentiate between substitution, addition, elimination and condensation (KC NDS, 2009**, 58)
salts	<ul style="list-style-type: none"> students are to explain the specific properties of salts with the help of ions, ionic lattices and electrostatic forces" (FA SH, 2016*, 42)
metals	<ul style="list-style-type: none"> students are to explain the specific properties of metals with the help of the metallic bond concept (FA SH, 2016*, 42)
alcohols	<ul style="list-style-type: none"> students are to explain the relations between properties and usages of important organic compounds (alkanes, alkanols, carbonic acids) (FA SH, 2016*, 42)
natural compounds	<ul style="list-style-type: none"> students are to describe the relation between occurrence, usage and properties of important natural compounds (FA SH, 2016*, 42)
...	...

* Ministerium für Schule und Berufsbildung des Landes Schleswig-Holstein (2016). Fachanforderungen Chemie – Allgemeinbildende Schulen, Sekundarstufe I - Gymnasium, Sekundarstufe II. Kiel. Schmidt & Klaunig.

** Niedersächsisches Kultusministerium (2009). Kerncurriculum Chemie für das Gymnasium – gymnasiale Oberstufe, die Gesamtschule – gymnasiale Oberstufe, das Fachgymnasium, das Abendgymnasium, das Kolleg. Hannover. Unidruck.

From this rather short comparison, one can see that for a variety of traditional topics there is at least one competency each explicitly described in state curricula that addresses the SPR concept. This is not surprising as explaining any observable phenomenon from a chemical perspective requires an analysis of the substance at

¹⁶ The transferability of the originally secondary level fundamental concept to the tertiary level will be explained in chapter 2.3.1.4.

the submicroscopic level. Thus thinking along the concept of SPR can facilitate a broader understanding of a variety of chemistry topics.

2.3.1.2 Levels of thinking

As SCHEFFEL, BROCKMEIER AND PARCHMANN (2009) point out, this concept might well be considered the most complex of the fundamental concepts because in order to understand it, the learner needs to be able to link three different *aspects* of chemical understanding:

- (i) the submicroscopic level of functional units of ions, molecules and metals (very often requiring appropriate representations of structures),
- (ii) the consequences of these functional units on the submicroscopic level, such as hydrogen bonds, spatial orientation of molecules in relation to each other, chirality, or even larger units such as lattices or crystalline structures, and
- (iii) macroscopic consequences of these larger submicroscopic units such as boiling temperatures or mechanisms of action of medicinal substances.

These aspects relate back to the levels of chemical representations proposed by JOHNSTONE (1991) and the idea of macro-micro-thinking (see among others GILBERT & TREAGUST, 2009; VAN BERKEL, PILOT & BULTE, 2009) stating that students need to navigate between these levels in order to firmly grasp chemistry. Particularly this task of relating aspects between the observable phenomena and models describing the atomic structure is often perceived as difficult (see among others DAVIDOWITZ & CHITTLEBOROUGH, 2009; VAN BERKEL ET AL., 2009). To improve students' understanding, MEIJER (2011) proposed stressing a further level in the macro-micro-thinking: the meso level as a series of scale-related levels in between the submicroscopic level of atoms and the macroscopic level of every-day products. Among other examples, he explains his idea of a meso level with the help of a jacket, for which the macro level would be the observable properties the jacket displays and the submicro level would be the atoms in the carbon-based molecules. Each larger functional unit would then be one meso level further towards the macro level (see also Figure 18): (i) the way the atoms are coordinated to form the macromolecule, (ii) the way individual macromolecules are coordinated to form a molecular network, (iii) the way these networks coordinate to form fibers, (iv) the way these fibers are coiled into threads, (v) the way these

threads are woven into fabrics, (v) the way the fabrics are made into the textile product (MEIJER, 2011, p.29f.). Even though some of the meso levels are indeed visible, they present intermediate steps for which certain properties arise due to the coordination of smaller building blocks.

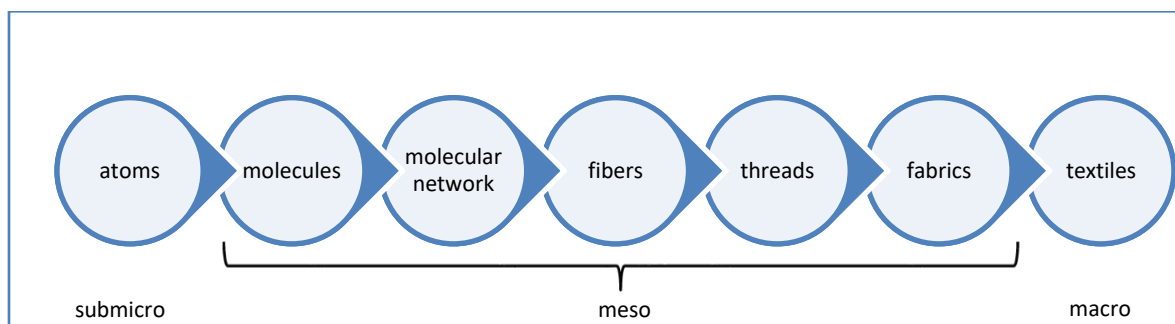


Figure 18: Example of submicro-meso-macro-thinking, according to MEIJER (2011)

Being able to navigate between the different levels in order to explain observable phenomena due to submicroscopic structures presents an essential skill in chemistry and is fostered through using the SPR concept.

2.3.1.3 Chemistry-related reasoning

Students often feel overwhelmed by the large number of apparently altering definitions (e.g. what is an oxidation), various models that build on or seemingly contradict each other (atoms as ranging from the smallest inseparable portion of an element to something that consists of protons, neutrons and electrons, for example) or a large amount of individual facts that they come across in chemistry (see among others TABER, 2001). They explicitly need to understand that they do not need to memorize every reaction they ever encounter but to grasp principles and reasoning within the discipline. Navigating within chemistry, learners need to have a certain factual knowledge and be able to reason at a formal cognitive stage (cf. MCCORMACK, FINLAYSON & MCCLOUGHLIN, 2010). In order to understand which compounds have ionic bonds, for example, students often try to memorize all examples presented to them. Along the lines of the information processing model (cf. MAYER, 1975), for example, this would cause a huge strain on their memory and they would be likely to

fail. What would greatly decrease this strain would be knowing that in general, and for introductory chemistry this is mostly the case, an ionic bond is present when atoms from the left-hand-side of the periodic table are bonded to atoms from the right-hand-side of the periodic table. In order to explain this rather general principle, one needs to refer to models on ion formation and Coulomb forces, and this should be made explicit at certain stages in the learning process. But even before students get to that point, the general trend helps them to avoid the mistake of not recognizing potassium chloride, when given in its empirical formula KCl , as a salt after they have explicitly covered sodium chloride in class, known to students also in its empirical formula $NaCl$. Apart from this trend, there are many more general rules in chemistry: an acid and a base react to a salt and water; when exposed to water, oxides from non-metal atoms form acids and oxides from metal atoms form bases; electronegativity of atoms increases from elements on the left of the periodic table to those at the right of the table and decreases from elements on the top of the periodic table to those at the bottom of the table; just to name a few. Even though these general trends sometimes have exceptions or need modification (especially as the chemistry level increases), generalizations are one way of reducing the number of individual facts that need to be retained at one particular moment. Generalizations need, however, some experience before they can form and become more than just memorized phrases. Another learning strategy quite relevant for chemistry is the chunking of information. When students are asked to compare the solubility of propan-1-ol and octan-1-ol in water, for example, rather than considering every single atom (considering every atom in the structure would itself add up to 39 individual pieces of information¹⁷) they can – in their mind or on paper – chunk each compound into two larger parts: the alkane-part and the alcohol-part (thus reducing the information pieces to four). In order to arrive at a solution to the question of solubility, they would then have to draw on existing knowledge that the alcohol-part is responsible for solubility in water (due to its polarity and the possibility to form hydrogen bonds with the water molecules) and that the alkane-skeleton is the reason for alkanes not being soluble in water (due to it being quite non-polar and not being able to form hydrogen bonds), thus arriving at six pieces of relevant information. From that point, students only need to deduce (which in turn adds one relevant piece of information) that the longer the alkane-similar part,

¹⁷ This number does not even include any information about comparing atoms based on electronegativity and the consequences this has for molecules yet.

the more unlikely the substance is to be soluble in water. This way of chunking information would reduce the cognitive load in the solubility example from over 40 pieces of information to a manageable number of seven. Chunking is frequently used in learning and explaining chemistry, as it is exactly what the concept of functional groups is all about.

Both strategies of generalizing and chunking information are, of course, independent of any fundamental concept. However, these strategies are particularly helpful when dealing with the SPR concept: By applying such strategies to reduce the complexity of a given set of facts and relations, a learner more or less automatically achieves two aims: For one, he/she moves through the submicro, meso and macro levels mentioned in the previous chapter as he/she no longer solely focuses on individual atoms (submicroscopic level) but rather on the relations between atoms and larger functional units (e.g. salts, acids/bases, alcohols) and their effects at the submicroscopic (e.g. electronegativity) or macroscopic level (e.g. solubility in water). Thus, the aspect of strategies for chemical reasoning is very closely connected to the characteristic of levels of thinking. The other aim achieved is the explicit connection inherent in the SPR concept: by considering the consequences that structural aspects have on the submicroscopic and macroscopic level (e.g. electronegativity and solubility in water, respectively), students explicitly reason in the realm of structure-property-relations. As JONASSEN AND IONAS (2008) put it, “[b]eing able to reason causally is an essential cognitive skill that is central to conceptualizing the physical world” (p.291). This skill must be practiced, especially in order to apply the concept later as a teacher: “Without conceptually explaining the underlying mechanisms of causal relationships, learners will not be able to build a coherent conceptual model of domain content, which will preclude problem-solving and other higher-order activities with a domain” (JONASSEN & IONAS, 2008, p.297). Similarly, SEVIAN AND TALANQUER (2014) define *chemical thinking* as a central idea for this skill; to them the term describes “the development and application of chemical knowledge and practices with the main intent of analyzing, synthesizing, and transforming matter for practical purposes” (p.10).

By using strategies that minimize cognitive load of memorizing individual information pieces and at the same time using the reduced amount of information

pieces in a logical reasoning process, SPR can provide many examples that help improve general but also chemistry-specific reasoning skills.

So far it has been stressed that the SPR concept can i) facilitate secondary students' broader understanding of a variety of chemistry topics due to its topic-transcending consistent perspective, ii) foster awareness of different levels of thinking that is necessary to differentiate between submicroscopic structures and macroscopic phenomena, and iii) improve students' reasoning skills when explaining chemical phenomena. These aspects are not only true for secondary learners but can also be transferred to the tertiary level.

2.3.1.4 Transferability to tertiary level

As mentioned above, SPR is a fundamental concept in secondary education (KMK, 2004) and consists of many content aspects embedded in various chemistry topics. These aspects range from characteristic properties of classes of substance such as salts, alcohols, metals, to the differences in bonding types to reaction types and mechanisms.

In university, the content aspects belonging to SPR can be found in many different courses as they are covered in each sub-discipline of organic, inorganic and physical chemistry. At a basic level, students will encounter structure-property-relations in inorganic chemistry when dealing with macroscopic properties of metals, molecules or salts such as melting points, density or reaction behavior, for example. These examples are quite similar, if not identical, to SPR aspects found in secondary education. When, at a more advanced level, considering the color of different metallic solutions with the metal ion at different oxidation states or the notion of reactions of hard and soft acids and bases, these aspects clearly go beyond secondary chemistry teaching. In the area of organic chemistry, students are faced at an introductory stage with a wide range of classes of substances with characteristic reactions which need to be predicated based on the existence of functional groups. Later in their studies, students will cover macromolecules and will have to deal with the fact that macromolecules often have more than one functional group with different reactivities which have to be considered. Students' ability to make predictions about how these

molecules might react under certain conditions is especially relevant in the related areas of medicinal chemistry or environmental chemistry.

In an interview study with twenty-three US organic chemistry educators in tertiary education, DUIS (2011) found out that the aspects provided as fundamental concepts in organic chemistry by at least 22% of the educators are: reaction mechanisms, correlation between structure and properties/reactivity, acid-base chemistry, structure and bonding, 3-D nature of molecules, functional groups, and electron delocalization/resonance (p.348). Apart from explicitly mentioning the correlation between structure and properties, the other aspects can also be included in the SPR-concept as well, as they cannot be taught without relating to at least the submicroscopic (or following MEIJER, 2011: the meso) part of the SPR concept and usually are related to meso or macroscopic properties of the substances under consideration. Similarly, DUIS (2011) reported that the educators considered the following concepts as most important for further learning in chemistry: acid-base chemistry, reaction mechanisms, structure-reactivity, and kinetics (p.348), of which at least the first three concepts can again be included into the SPR-concept without any problems.

A short case study with two organic chemistry textbooks for the tertiary level and adjacent university lecture notes showed that there was a strong focus on the representational and, to a somewhat lesser extent, on the submicroscopic level and that links between the levels are often less explicit and obvious (SCHEFFEL, HERZOG & PARCHMANN, 2011) suggesting that the individual parts of the SPR concept are definitely dealt with in tertiary education but that the actual relations between these parts are often left for the students to grasp themselves. The concept's relevance in tertiary chemistry learning is also emphasized by research on tertiary students' misconceptions in the area of SPR (for a selection see chapter 2.3.2.3).

In Germany, the GDCH not only includes the explicit term *structure-property-relations* into their recommendation as to the tertiary curriculum for future chemistry teachers (GDCH, 2008), they also suggest dealing with various aspects of the concept in their recommendations for general chemistry programs at tertiary level (GDCH, 2015). In the latter document, they state that when dealing with the topic of reaction mechanisms, the learner should manage to predict reactivity and favorable reaction paths from the electronic and steric effects of the molecules, from the electrophilic or

nucleophilic character of reactants and from the conditions surrounding a reaction (GDCH, 2015, p.29). Another example is given for the topic of solid state inorganic chemistry, in which the acquired skills should enable a learner to obtain a deeper understanding of relations between structure, chemical bonds and properties (GDCH, 2015, p.36).

Therefore, it can be concluded that SPR is necessary for understanding chemistry and therefore is dealt with both at secondary and tertiary level. After this demonstration that the secondary fundamental concept of SPR can be transferred to the tertiary level and a preceding look at opportunities this concept can provide for learners in the scope of understanding various topics in and a general reasoning ability for chemistry, the final reason why SPR has been chosen shall be addressed: secondary and tertiary learners show numerous problems with aspects belonging to structure-property-relations. Before doing this, a brief and more general look at the area of pedagogical content knowledge has to be taken.

2.3.2 Pedagogical Content Knowledge (PCK)



Before presenting reported problems that students on secondary and tertiary level have regarding the concept SPR, a very short overview on different models on pedagogical content knowledge, in the following referred to as PCK, will be given, followed by a closer look at possible ways to approach the concept in the teaching process. Both the possible order of teaching certain aspects of the fundamental concept as well as learners' problems with the concept are aspects belonging to the PCK area, as shall be explained in the following paragraphs.

2.3.2.1 Short overview of PCK models

When looking at different conceptualizations of PCK, it becomes obvious that there is quite a variety regarding its suggested (sub-)dimensions. Several articles reviewing the conceptualization of PCK have been published (see among others, NILSSON, 2008; KIND, 2009; KAYA, 2009 & SCHMELZING, 2010), and therefore this venture

shall not be repeated at this point. Rather, the diversity of models shall be stressed again. Table 11 is an aggregate of possible sub-dimensions of pedagogical content knowledge as presented in the review literature, it is not to be seen as an exhaustive list. It needs to be pointed out that most authors mention more aspects in their overall-models on teacher knowledge than are checked in Table 11 – yet they might not consider them to be part of PCK. This is especially true for the sub-dimension of subject matter knowledge, which is sometimes considered a separate dimension (cf. GROSSMAN, 1990; MAGNUSSON, KRAJCIK & BORKO, 1999; SHULMAN, 1987) and sometimes is seen as part of PCK (cf. HASHWEH, 2005; MARKS, 1990; MULHALL ET AL., 2003). The former view stresses that a profound understanding of the discipline's content is a necessary knowledge base alongside PCK, while the latter emphasizes that knowledge about different PCK facets such as representations and analogies beneficial for or hindering learning are always topic- and therefore content-specific. Similarly, SHULMAN (1986) proposes that knowing about “the most regularly taught topics in one's subject area, the most useful forms of representation of those ideas, the most powerful analogies, illustrations, examples, explanations, and demonstrations” (p.9) belongs to the PCK dimension, being a “second kind of content knowledge” (p.9).

In addition, the mere presence of a sub-dimension in Table 11 does not yet say much about the authors' actual understanding of that sub-dimension (e.g. is the sub-dimension *goal or intention* understood in its subject-specific or general meaning), so that this evaluation can only be comparing things at a rather superficial level. Even though Table 11 therefore might be considered limited in its explanatory power, it does show two things: i) a variety in conceptual understanding and in the number of sub-dimensions considered to be part of PCK by different researchers as well as in terminology and ii) a common core of two aspects are included in (almost) every model. One of these aspects is instructional strategies, i.e. knowledge of how to best teach a topic or concept or to use particular examples or representation; the other being student cognition, i.e. knowledge of learners' views and knowledge at particular stages in their learning process as well as learning difficulties that might hinder their understanding of certain topic or concept. In Table 11, no distinction is made between hypothetically proposed or empirically validated models.

Table 11: A closer look at possible sub-dimensions of PCK (based on KIND, 2009 & SCHMELZING, 2010)

	<i>PCK in this model includes knowledge of...*</i>										
	<i>instructional strategies</i>	<i>student cognition & learning</i>	<i>subject-matter</i>	<i>curriculum</i>	<i>purpose/ goals</i>	<i>assessment</i>	<i>context</i>	<i>classroom management/ pedagogy</i>	<i>media</i>	<i>socio-cultural issues</i>	<i>school</i>
SHULMAN, 1986	✓	✓									
SHULMAN, 1987	✓	✓									
TAMIR, 1988	✓	✓		✓		✓					
SMITH & NEALE, 1989	✓	✓			✓						
GROSSMAN, 1990	✓	✓		✓	✓						
MARKS, 1990	✓	✓	✓						✓		
COCHRAN, DERUITER & KING, 1993		✓	✓				✓	✓			
GEDDIS, 1993	✓	✓		✓							
FERNÁNDEZ-BALBOA & STIEHL, 1995	✓	✓	✓		✓		✓				
VAN DRIEL, VERLOOP & VOS, 1998	✓	✓									
GESS-NEWSOME, 1999	✓	✓									
VEAL & MAKINSTER, 1999	✓	✓	✓	✓	✓	✓	✓	✓		✓	
KOBALLA, GRÄBER, COLEMAN & KEMP, 1999		✓	✓	✓			✓	✓			
MAGNUSSON, KRAJCIK & BORKO, 1999	✓	✓		✓	✓	✓					
MULHALL, BERRY & LOUGHRAN, 2003	✓	✓	✓		✓		✓	✓			
HASHWEH, 2005	✓	✓	✓	✓	✓	✓	✓	✓			
BANKS, LEACH & MOON, 2005	✓	✓	✓	✓				✓			✓
BAUMERT & KUNTER, 2006	✓	✓									
PARK & OLIVER, 2008	✓	✓		✓	✓	✓					
LEE & LUFT, 2008	✓	✓		✓	✓	✓			✓		
ROHAAN, 2009	✓	✓			✓						
RIESE, 2009	✓	✓		✓	✓	✓					
ROLLNICK, BENNETT, RHEMTULA, DHARSEY & NDLOVU, 2010		✓	✓				✓	✓			
VAN DIJK & KATTMANN, 2010	✓	✓									
DÖHRMANN, KAISER & BLÖMEKE, 2010	✓	✓		✓	✓	✓					

* Certain sub-dimensions might be part of the overall model but are not specified as that models' conceptualization of PCK.

Based on these different models, a group of researchers working in the field of PCK research and modelling, some of whom having published their individual models before, came up with a joint model of pedagogical content knowledge; this so-called *PCK summit consensus model* was presented by GESS-NEWSOME AND CARLSON at the ESERA conference in 2013 (cf. Figure 19).

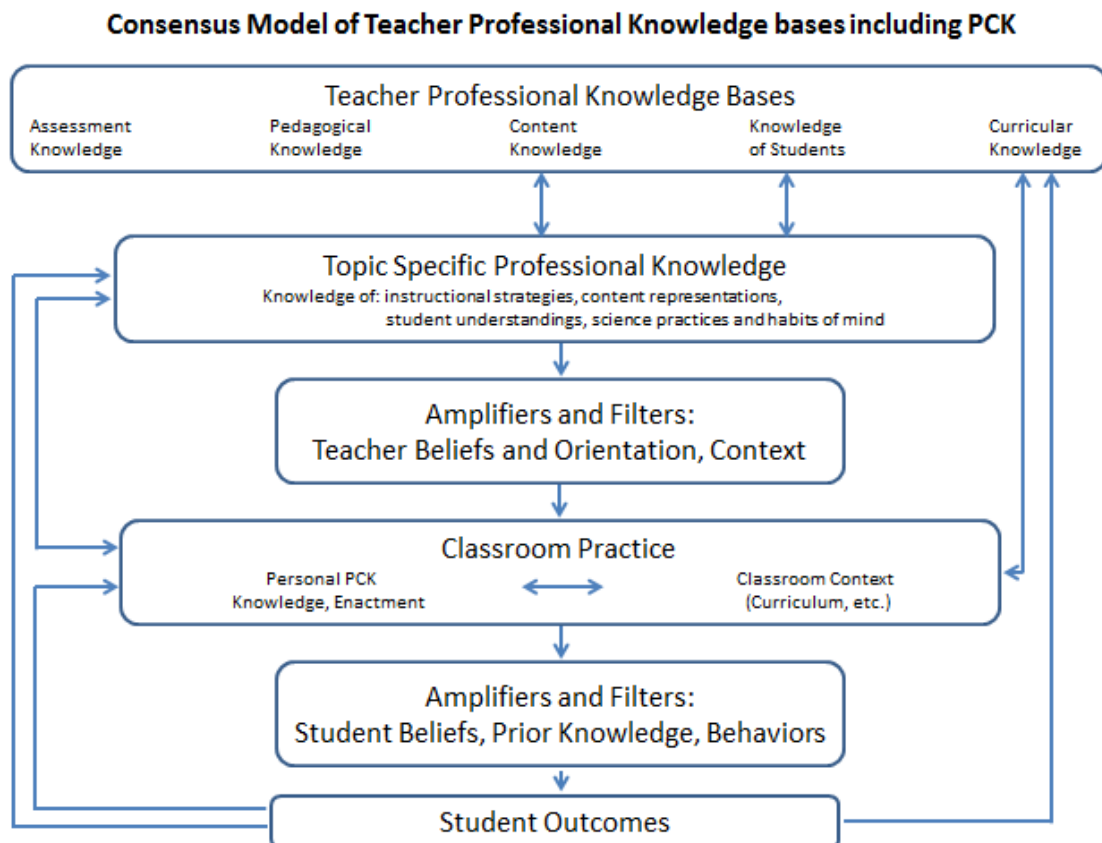


Figure 19: Consensus Model of Teacher Professional Knowledge Bases including PCK (GESS-NEWSOME & CARLSON, 2013)

Even though other researchers before have also stressed that PCK is not a stand-alone knowledge area and is influenced by other areas inherent in a teacher's professional knowledge, the PCK summit consensus model puts a strong focus on interdependencies between PCK and other knowledge areas (cf. GESS-NEWSOME & CARLSON, 2013): it builds on the so-called *teacher's professional knowledge bases*, including more general knowledge aspects of assessment, pedagogy, content, curriculum and students. When applying those general knowledge areas to a more topic-specific level, teachers tap into their *topic-specific professional knowledge*,

including knowledge of instructional strategies, content representations, student understanding, science practices and habits of mind. Either amplified or filtered through a teacher's beliefs and orientations, his/her own misconceptions as well as the teaching context, this topic-specific knowledge then boils down to the teacher's personal PCK. This is defined in an approach that might be called holistic as it stresses both knowledge and performance aspects:

“**[k]nowledge** of, **reasoning** behind, and **planning** for teaching a particular **topic** in a particular **way** for a particular **purpose** to particular **students** for **enhanced student outcomes** (Reflection *on* Action, Explicit)

[and]

[t]he **act of teaching** a particular **topic** in a particular **way** for a particular **purpose** to particular **students** for **enhanced student outcomes** (Reflection *in* Action, tacit or explicit).”

(GESS-NEWSOME & CARLSON, 2013, slide 15¹⁸)

The first part of the definition is usually used in paper-and-pencil competence testing and will also be used in this project while the second part is more appropriate for performance testing. This consensus model has already been designed with two main thoughts in mind:

- i) The thought of assessing teachers' PCK: from general to specific knowledge bases, which could be tested relatively easy with a paper-and-pencil test, to a performance test in an actual teaching situation which would be considerably more complex to assess.
- ii) The thought of professional development, as the model may also describe a teacher's own learning process: from acquiring knowledge, which changes from general to topic-specific over the course of the learning process, to performance according to that knowledge.

While most of the research studies in Table 11 focus on the assessment part, and of those, the vast majority evaluate knowledge instead of performance, some, usually

¹⁸ color-coding as in original

more qualitative, studies attempt to evaluate learning processes. Again, the first part of thought i) is the one this dissertation is based on: assessing (prospective) teachers' knowledge base via paper-and-pencil tests aiming at performance-related activities. Other researchers such as BLÖMEKE ET AL. (2014), MULHALL ET AL. (2003) or VAN DRIEL ET AL. (1998) have focused on thought ii), i.e. the development of teachers' competences. As the latter requires valid and reliable instruments, the development of these instruments regarding a particular fundamental concept have to come first and are the goal of this dissertation. Before, a lot of models have been proposed with one of these thoughts in mind and because of that it has been rather difficult to compare models effectively. As this PCK model combines both ideas, the model truly is a consensus model.

GESS-NEWSOME AND CARLSON (2013) also states some assumptions regarding PCK, based on what research in that area has shown. She argues that every teacher has a certain expression of PCK ranging from weak to strong and that each teacher's PCK can be increased through several strategies: continuing professional development, teaching interventions and teaching experience; the latter, however, not being a guarantee for improvement. She also points out that a teacher's increase in PCK needs to be based on already existing sound content knowledge or needs to be accompanied by studying this content knowledge, a fact that is supported by some studies according to the researcher. Furthermore, she assumes that a teacher with stronger PCK is "better able to improve student learning" (GESS-NEWSOME & CARLSON, 2013, slide25) even though this aspect is somewhat ambiguous in research data.

After having looked very briefly at PCK in general, a closer look at how the concept of SPR can be developed to foster learners' understanding becomes necessary before describing relevant research on student problems regarding SPR.

2.3.2.2 Curricular development of SPR

There are theoretical approaches for describing the curricular development of the structure-property-relations concept.

One such approach is the notion of macro-micro-thinking in the Netherlands (group around VAN BERKEL): “The scientists’ macro-micro thinking is to understand and to use the relations between the observed phenomena at the macro level and the models of the invisible particles at the submicro level. In other words, scientists try to describe, understand and predict the properties of materials by relating these to the scientific models of structures at a submicroscopic level” (MEIJER, 2011, p.11). As MEIJER pointed out, the content-inherent switching between the levels should be made apparent to the learner (cf. Figure 18 in chapter 2.3.1.2).

NGAI, SEVIAN AND TALANQUER formulated a curricular sequence for secondary education and beyond (cf. Figure 20) that all discipline-specific concepts should be developed “by fostering sophisticated reasoning at the macroscopic and submicroscopic levels in a parallel and coordinated manner” (p.2456f.).

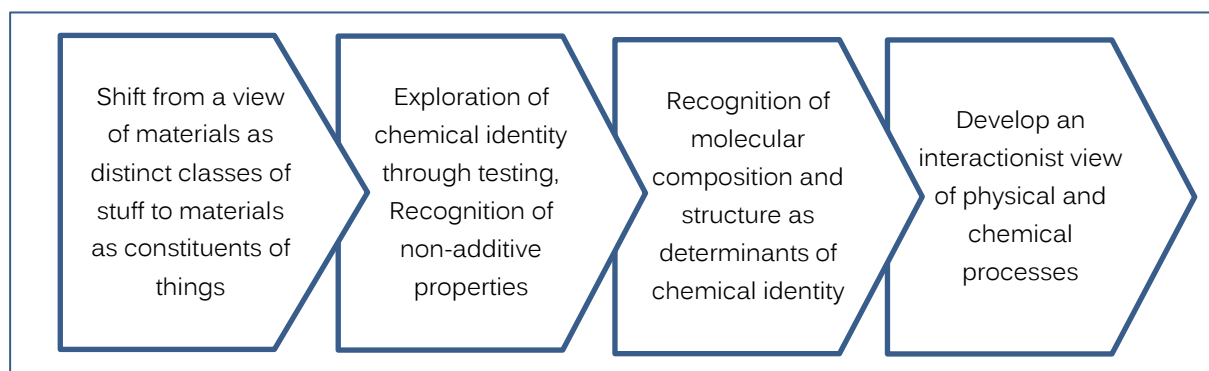


Figure 20: Excerpt of curricular sequence for students’ understanding of a discipline-specific cross-cutting concept, here chemical identity (NGAI ET AL., 2014)

Another learning progression has been proposed by the research group around COOPER (e.g. COOPER, GROVE, UNDERWOOD & KLYMKOWSKY, 2010; COOPER, UNDERWOOD & HILLEY, 2012a/b). Their learning progression focuses on the fundamental concept called *structure and function*, being considered “an overarching concept in chemistry” (COOPER ET AL., 2012a, p.195) that helps provide a “clear understanding of how molecular-level features can be used to predict macroscopic physical and chemical properties” (COOPER ET AL., 2012b, p.1351). These different features should

be navigated through by the learner in a particular sequence, thus delineating their learning progression for this concept (cf. Figure 21): Students should first be introduced to the central idea that every reaction and state change takes place due to energetic reasons. Only afterwards are students taught how to navigate logically from structure to explaining the structure's properties by following a concrete step-by-step description of reasoning: First, physical properties should be introduced which need to be related to the molecular structure of the substance under consideration. In order to visualize this structure students should understand three-dimensional models and as a subsequent step, be able to translate these into two-dimensional drawings. The backward translation process needs to also be addressed in order to then arrive at more complex structures. Finally, Lewis-structures should be introduced from which the physical properties having started the learning cycle can be explained.

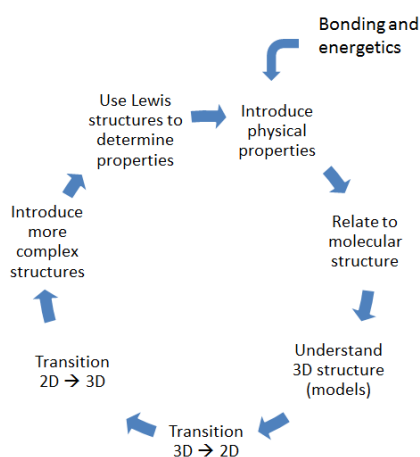


Figure 21: Learning progression on structure and function as proposed by COOPER ET AL. (2012b)

For the German educational standards of chemistry, curricular developments have also been proposed (PARCHMANN, BÜNDER, DEMUTH, FREIENBERG, KLÜTER & RALLE, 2006; SCHWARZER, RUDNIK & PARCHMANN, 2013). Figure 22 suggests such a curricular development for SPR: Students are usually first introduced to properties of substances such as magnetism or hardness. From there, they will encounter the fact that certain properties are the result of particles, for example that metals have certain characteristic features such as metallic shine and relative hardness. As a next step, they come into contact with the fact that everything is made up of atoms and they

look at properties resulting, for example, from intermolecular bonds: e.g. explaining why a lake freezes from top to bottom. Later students will also take into account the fact that the size of the particles themselves leads to different properties, such as the whole area of nano chemistry. From there, students investigate properties that go beyond the idea of particles in isolation: properties can also depend on the system surrounding certain particles, e.g. copper ions in an aqueous solution result in a different color of that solution than copper ions in ammonia solution. The last two realms of chemistry are sometimes dealt with in advanced upper secondary courses and are more often reserved for university courses. Even though this development is portrayed in a somewhat hierarchical way, this curricular development (Figure 22) it is by no means exclusively linear. It is rather to be understood as a spiral curriculum meaning that not only the steps build upon each other but also that each step can be revisited for certain questions, topics or explanatory models.

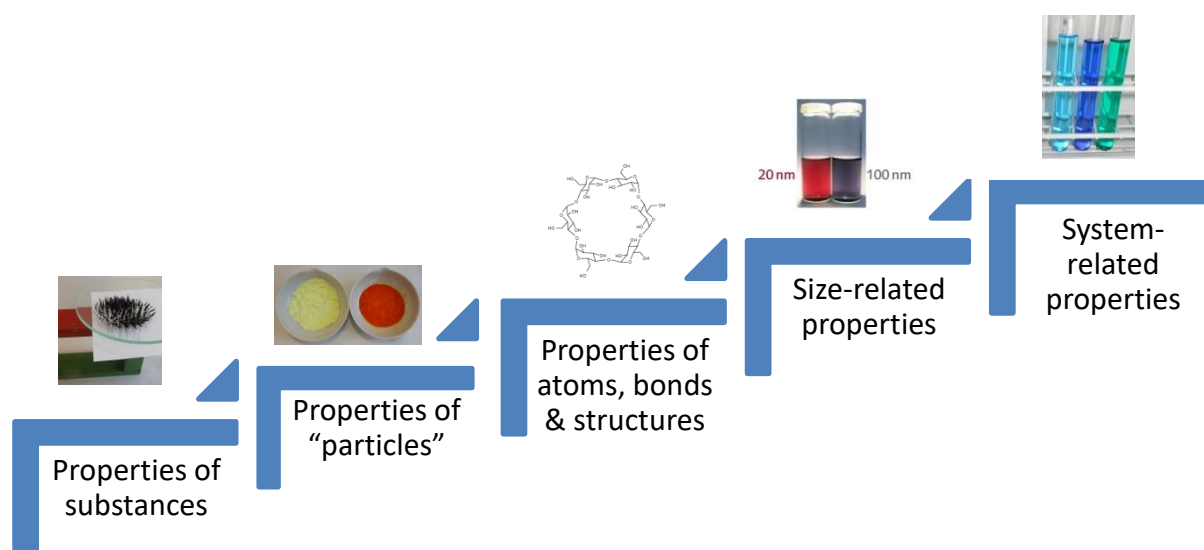


Figure 22: Curricular development of the SPR-concept (adapted from SCHWARZER ET AL., 2013)

A pre-service teacher needs to know about these different curricular developments that are suggested to be beneficial for learning. In addition, he/she needs to reflect on these in order to find suitable examples, representations, experiments and models that illustrate certain stages in the curricular developments.

2.3.2.3 Problems regarding student understanding reported in connection with SPR

As has been already mentioned quite briefly in chapter 2.3 (reasons for choosing SPR as fundamental concept), there are many problems reported that secondary and tertiary students have with aspects of SPR. As it is a vital part of PCK to know about students' problems and potential pitfalls within a specific content topic or concept, these studies focusing on students' difficulties will be presented at this point.

One reason for looking into the research on secondary and tertiary students' difficulties and misconceptions is that researchers have pointed out that teachers often have similar misconceptions as their students (LAWRENZ, 1986; WANDERSEE, MINTZES & NOVAK, 1994). Furthermore, pre-service teachers are themselves learners of chemistry content in their university education, so that some of the problems reported for chemistry majors or minors might be found in this population as well. Another argument for looking into problems shown by both secondary and tertiary students is that there is no distinct line dividing content aspects into those only covered in school and those only covered at university. Rather, the transition is somewhat blurred, especially as students progressing to university have very heterogeneous prior knowledge, due to a large variety in time spent on the subject and depth of selected chemistry courses at school (cf. chapter 2.1). Some students arrive at university, for example, with only the rudimentary understanding that there are atoms and different types of bonding, while other students have already dealt with photochemistry and azo compounds. Whole books have been written on students' general misconceptions and problems in various areas of chemistry (cf. for example BARKE, HAZARI & YITBAREK, 2009; KIND, 2004 or TABER, 2002). As the focus of this project is SPR, students' problems relating more closely to this fundamental concept shall be presented in the following paragraphs.

Problems observed with secondary and tertiary students are often comparable. The most fundamental problem with structure-property-relations is that both secondary and university students exhibit problems with switching between the levels of thinking as they hardly use the submicroscopic level at all when having to explain macroscopic properties (NAKHLEH, SAMARAPUNGAVAN & SAGLAM, 2005). Similarly, university students commonly show problems with "explain[ing] physical properties such as boiling- and melting-point trends [or] solubilities of organic

compounds” (HENDERLEITER, SMART, ANDERSON & ELIAN, 2001, p.1126). Even though most lecturers would most likely say that their teaching definitely contains aspects for which macroscopic findings need to be interpreted through submicroscopic structures, some research suggests that this reasoning often is not made explicit enough to the students (cf. BHATTACHRAYYA & BODNER, 2005; TAAGEPERA & NOORI, 2000). This might be one reason for many of the problems that both secondary and university students have with the concept of structure-property-relations.

In order to organize the more specific problems with SPR comparable at secondary and tertiary level, the reported aspects are categorized according to the curricular development structure presented in Figure 22 above. Note that only the first three stages of the curricular development structure are used here as *size-related properties* and *system-related properties* are not necessarily covered in secondary education.

When it comes to the curricular development step of *properties of substances*, the most common misconception in school is that macroscopic properties are often assigned to individual atoms of an element (cf. among others BEN-ZVI, EYLON & SILBERSTEIN, 1986; LEE, EICHINGER, ANDERSON, BERKHEIMER & BLAKESLEE, 1993; MULFORD & ROBINSON, 2002; ÖZMEN, 2004) or that substances contain non-material properties, implying that properties could be transferred from one substance to another (SANMARTI, IZQUIERDO & WATSON, 1995; STEFFENSKY, PARCHMANN & SCHMIDT, 2005; among others). Describing properties of substances without referring to the submicroscopic level is usually done when school students first come into contact with the subject of chemistry. This misconception is not observed very often in university students, which might be because the latter group is asked to reasons mostly within the submicroscopic level.

In regard to the next step *properties of 'particles'*, university students are said to have difficulty in grasping larger functional units, treating an “ester as if it were a carbonyl attached to an ether, rather than [...] a single unit that would have a characteristic reactivity” (BHATTACHRAYYA & BODNER, 2005, p.1405). This problem is not mentioned very often in research literature, possibly because tertiary students are rarely required to identify whether they think of a functional unit as an entity or as the

sum of its parts. Rather, university tests often ask students to complete organic mechanisms, which they even seem to be able to do more or less correctly without comprehending the underlying ideas (BHATTACHARYYA & BODNER, 2005).

For the subsequent step of *properties of atoms, bonds & structures*, the largest number of research studies dealing with students' problems in the area of SPR can be found. From Figure 23 one can see that the problem of confusing changes in aggregation states with the breaking of bonds can be found in both groups of students. For example, secondary students often regard physical changes and chemical reactions as equal (cf. LEE ET AL., 1993; or ANDERSSON, 1986). This is especially true for the boiling process when water is believed to dissociate into hydrogen and oxygen (GARNETT, GARNETT & HACKLING, 1995). The same can be said for university students, as TAAGEPERA AND NOORI (2000), for example, found in their organic class that their students did not distinguish between "boiling and burning, thinking that covalent bonds are broken on boiling" (p.1227).

When it comes to bonding, several researchers describe students' difficulties regarding this aspect. They have investigated secondary students' difficulties in distinguishing between covalent and ionic bonds, which becomes especially apparent when asked to draw either type (TABER, 1997). In a study by LEVY NAHUM, MAMLOK-NAAMAN, HOFSTEIN & KRONIK (2008), differentiating between inter-molecular and intra-molecular bonding was the most frequent difficulty students had in the area of bonding: When asked to define a hydrogen bond, students could choose the correct definition but when asked further, they wrongfully identified a hydrogen bond as either the bond between the oxygen and hydrogen atom within one water molecule, or as a bond appearing between two oxygen atoms or two hydrogen atoms of two separate water molecules, thus showing that the notion of hydrogen bonds was not understood thoroughly. This phenomenon is something VINNER (1997) calls pseudo-conceptual thinking and states that certain processes "are simpler, easier, and shorter than the true conceptual processes. Under these circumstances, it is only reasonable to assume that many students will prefer the simpler, easier, and shorter processes to the more complicated conceptual processes" (p.101). Simply put, students will recognize definitions or recall certain elements of a definition without actually having understood the definition.

In regard to hydrogen bonding, university students also show problems: they demonstrate an inability to recognize the need for unpaired electrons and the necessity of hydrogen being directly bonded to an electronegative atom; rather, students tend to focus only on chain lengths instead of considering molecular shape (SCHMIDT, 1996).

Bond polarity is another aspect for which problems could be observed in both secondary and tertiary students. When asked to locate the bonding electrons in a polar covalent bond, secondary students often positioned the electrons exactly in between the two bond partners as they thought that electrons in a covalent bond would be shared equally (PETERSON, TREAGUST & GARNETT, 1989), disregarding the aspect of electronegativity. Similarly, university students seemed to think that "bond polarities depend on absolute electronegativities of atoms only" (TAAGEPERA & NOORI, 2000, p.1227) and not on differences in electronegativity of two neighboring atoms. This lack of fully understanding bond polarity is likely to lead to "fundamental problems when students are faced with their first university organic course when they will [have to deal with] mechanisms of reactions in some detail" (HASSAN, HILL & REID, 2004, p.47).

Further problem areas for university students are in the realm of organic chemistry. GRAULICH, HOPF AND SCHREINER (2010) pointed out that university students have difficulties in finding similarities in organic structures, more specifically in predicting reactions and mechanisms of cyclic six-electron systems. Related to the six-electron systems is the more general idea that electron density can be used to predict physical properties. TAAGEPERA AND NOORI (2000) stated that university students in an organic chemistry class rarely used electron density in analyzing submicroscopic properties such as bond polarity or macroscopic properties such as solubility. In regard to predicting reaction mechanisms, students also had problems with using the appropriate justifications when reasoning for or against the reactivity of certain compounds. RUSHTON, HARDY, GWALTNEY AND LEWIS (2008) found that the results to their study "seem[ed] to indicate that steric hindrance elsewhere in the molecule was an over-riding consideration" (p.127), demonstrating that students considered steric hindrance even when it was not an appropriate aspect: they named steric hindrance as the only reason for a reduced chemical reactivity in S_N1 reactions

instead of considering the stability of carbocations as intermediates which should have been the relevant feature in the requested explanation (RUSHTON ET AL., 2008).

The reported problems with SPR have been categorized according to the curricular development structure presented above (Figure 22). Figure 23 visualizes a synthesis of these problems for both educational levels.

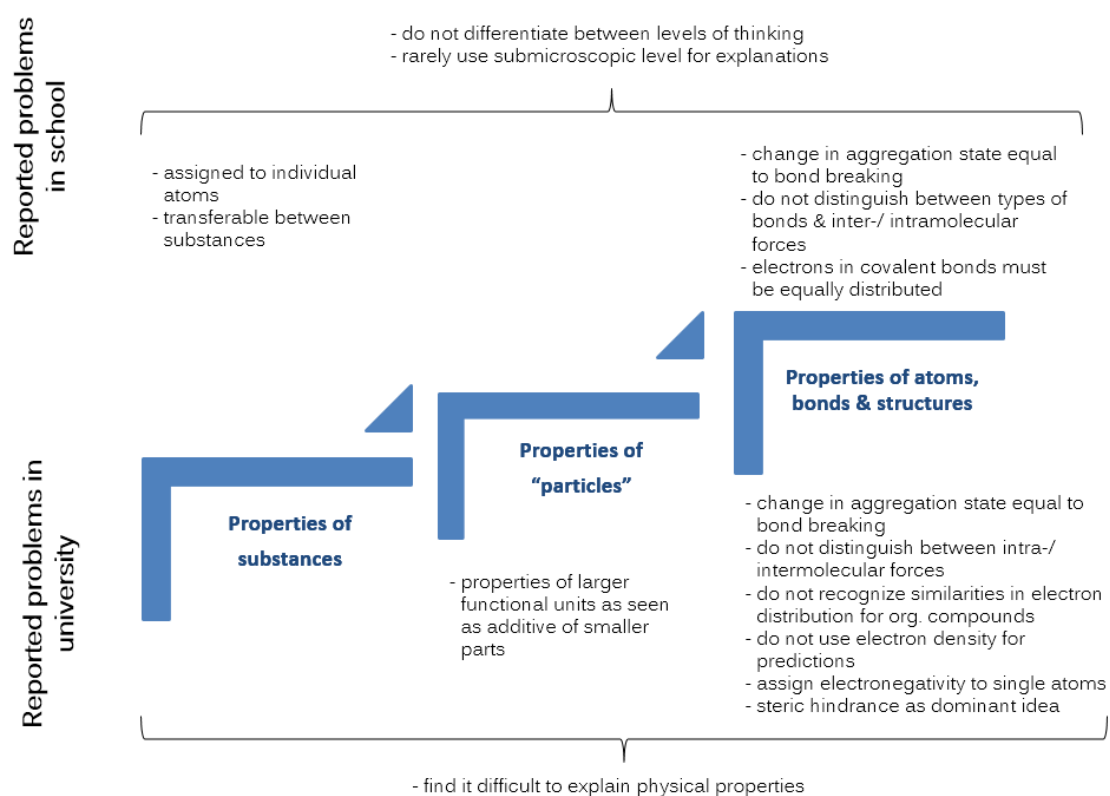


Figure 23: Overview of students' problems

To sum up, there are many levels of problems reported for secondary and tertiary learners regarding SPR. The one most often reported is the lacking integration of information from the submicroscopic level to explain macroscopic properties. A second level is the particle level in which problems with atoms, molecules and functional groups become apparent: students often assign macroscopic properties to individual atoms of an element or think that substances contain non-material properties that can be transferred from one substance to another. Furthermore,

students display problems with grasping larger functional units, e.g. when treating the properties of a functional unit merely as the sum of its components. A third level of problems with the structure-property-concept is the area of inter- and intramolecular bonding. Not being able to differentiate between the consequences of physical and chemical changes falls into this group of problems. The same is true for problems of distinguishing between covalent and ionic bonds as well as understanding hydrogen bonds. A fourth level of challenges can be found in the handling of explanations for structure-property-relations: Reasoning necessary for understanding and applying this concept is often not made explicit enough to students and teachers are sometimes satisfied with pseudo-conceptual thinking, meaning that the students arrive at a correct answer without having mastered the underlying concept. Future teachers need to be aware of all of these problems and difficulties as they have to i) realize that their own teaching could cause some of them, ii) find out if their learners exhibit any of them, and if so, explicitly address them in order to use them as learning opportunities.

3 RESEARCH PROJECT AND RESEARCH OBJECTIVES

RESEARCH PROJECT

This thesis is embedded into a research project planning to evaluate pre-service teachers' professional knowledge in the science subjects and mathematics, called *KiL* (*Kompetenzen in der Lehrerbildung* translated as “competences in teacher education”). Even though the title of the project includes the term “competences”, it has been operationalized by project members as knowledge because it is actually knowledge that is assessed through paper-and-pencil tests (cf. also first part of the PCK definition by GESS-NEWSOME & CARLSON, 2013). As this is an attempt to assess competence through paper-and-pencil tests, the working definition of competence proposed by KLIEME ET AL. (2008) will be used here: “competencies as context-specific cognitive dispositions that are acquired by learning and needed to successfully cope with certain situations or tasks in specific domains” (p.9). In the domain of PCK, this definition is widened to also include hypothetical performance-related real-life activities that are authentic in the context of teaching (cf. SHAVELSON, BAXTER & PINE, 1991), as far as test items allow for this.

As professional knowledge has been considered to be the foundation for successful teaching both from a theoretical (BROMME, 1997; or SHULMAN, 1986) and an empirical perspective (BAUMERT & KUNTER, 2013; or HILL, ROWAN & BALL, 2005), the project's goal was to develop test instruments measuring the professional knowledge during the university phase of teacher education. As stated above, the three main facets of this professional knowledge agreed on by researchers (cf. chapter 2.2.1) are content knowledge, pedagogical content knowledge and pedagogical knowledge. In order to analyze pre-service teachers' knowledge in these facets, each subject¹⁹ created its own test instruments to assess its content knowledge and its pedagogical content knowledge. In addition, one test instrument for pedagogical knowledge was also developed in the *KiL*-project. As the instruments were intended to be used for further research studies, the tests were designed to address a wide range of knowledge areas and difficulty. And instead of focusing on school-relevant content

¹⁹ i.e. biology, chemistry, mathematics and physics

knowledge, all subjects decided to put the main emphasis on aspects that could only be answered with university knowledge. For chemistry, two main areas of competence have been chosen as the basis for both content knowledge and pedagogical content knowledge tests²⁰: 1) chemical representation as a vital means of communication in and about chemistry and 2) structure-property-relations as a fundamental concept necessary for understanding chemical phenomena (cf. chapter 2.3). The former has been developed by a fellow doctoral student and is described in TASKIN (2014) while the latter is presented in this thesis. Due to this separation, some of the aspects suggested by COOPER ET AL. (2012b) in their learning progression for structure and function are not being addressed in this thesis but in the adjacent dissertation focusing on the competence area of chemical representations (Figure 24, gray areas are dealt with in the adjacent dissertation on chemical representation). Even though the development of a learning progression was not the intention of the *KiL*-project, the aspects presented by COOPER ET AL. (2012b) can be used to show which aspects belonging to the larger SPR concept are dealt with in the adjacent dissertation. The distinguishing criterion whether a test item belonged to this dissertation or the one on chemical representations was its necessity to apply different levels of thinking (cf. chapter 2.3.1.2): If all necessary interpretations remain at the structural or submicroscopic level (e.g. deducing how atoms are connected within the compound when given a certain sum formula), then that task

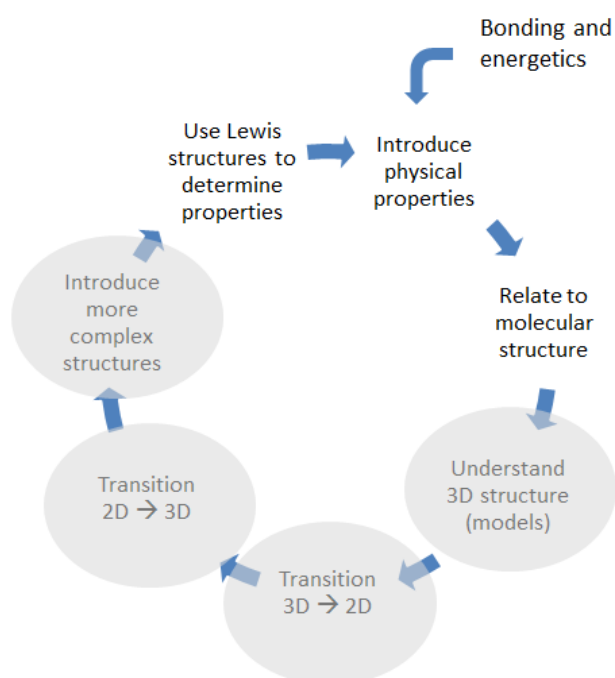


Figure 24: Learning progression on structure and function as proposed by COOPER ET AL., 2012b, used here as the basis for differentiating between items belonging to dissertation on chemical representation (gray) or to this dissertation on structure-property-relations

²⁰ Following the idea that the topic under investigation should not differ between content knowledge test and pedagogical content knowledge test (cf. WALPUSKI ET AL., 2012)

belongs to the chemical representation dissertation. If solving an item also requires the pre-service teacher to use macro-micro-thinking (cf. VAN BERKEL, PILOT & BULTE, 2009) or any combination of the macro-meso-submicro-thinking (MEIJER, 2011), then that item belongs to this dissertation on structure-property-relations. In chapter 6.4.4., the resulting test instruments from TASKIN (2014) and this dissertation project will be looked at jointly.

Apart from assessing pre-service teachers' professional knowledge, a central intention of the *KiL*-project was to investigate how the three areas of professional knowledge were related (KLEICKMANN ET AL., 2014).

At this point it should be made clear (again) that the *KiL*-project is not attempting to assess teaching performance in the classroom. Rather, it is attempting to get at the underlying knowledge that might later influence certain behavior in class²¹. As it is essential, especially for PCK, to have the task reflect a real-life situation as closely as possible in order to get a response from the test taker that might actually reflect his/her response in real life (SHAVELSON, 2010), the test instruments should include performance-oriented tasks. However, routinized teaching performance cannot be expected of pre-service teachers in the university part of their education. That is why the pedagogical content knowledge items cover only the knowledge part of the PCK consensus model (cf. chapter 2.3.2.1) and not its performance aspects. The test instruments can only assess hypothetical procedural as well as hypothetical conditional knowledge, thus shedding light into the areas of how one could go about something as well as when and why one could adopt certain actions (PARIS, LIPSON & WIXSON, 1983). Both content knowledge and pedagogical content knowledge tests on structure-property-relations therefore contain these hypothetical knowledge aspects.

²¹ Apart from knowledge, many other factors may of course also influence a teacher's behavior in class; among others: his/her beliefs, attitudes, expectations, assumed role in learning process, surrounding conditions, school environment, individual students or classes, etc.

RESEARCH OBJECTIVES

The overall intention of the *KiL*-project and therefore, this dissertation, was to create two valid, reliable and objective test instruments assessing pre-service teachers' content and pedagogical content knowledge as well as to investigate the structure of professional knowledge of pre-service teachers. In order to shed light into this area, two main research objectives arose:

- O_{study I}: To create valid, reliable and objective instruments with which to assess pre-service teachers' chemistry content knowledge and chemistry pedagogical content knowledge focusing on structure-property-relations and to investigate possible influencing factors.
- O_{study II}: To describe the overall structure of pre-service chemistry teachers' professional knowledge and to find aspects that influence pre-service teachers' chemistry content knowledge and chemistry pedagogical content knowledge.

Two separate studies involving different pre-service teachers have taken place to follow these research objectives. Therefore, the studies and corresponding research objectives will be described separately in the following chapters. Before going into the analysis of the test instruments, however, their development has to be described.

4 ITEM DEVELOPMENT

In the following pages, the approach to item development will be described first for the content knowledge and then for the pedagogical content knowledge area of SPR. This description will include the reasoning behind the selection of certain content, describe further factors used for a variation in items, aspects regarding validation as well as presenting some of the items.

4.1 Content Knowledge

CK	secondary	learner
PCK	tertiary	teacher

In this dissertation, the focus will be on the logical inferences between observable properties on the macroscopic level and invisible structural makeup on the submicroscopic level. As the adjacent dissertation project on chemical representations (cf. Tasin, 2014) might be considered a prerequisite of the latter, the items of both projects were developed in close communication wherever possible. The items were not developed to validate an existing competence model, as is done when items are to fill the grids for model testing (for example ROPOHL, 2010). Similar to the *ProwiN*-project (cf. TEPNER ET AL., 2012), the resulting model (cf. Figure 26) is seen as a tool to generate a wide variety of items.

4.1.1 Sources for and content of items

Based on a content analysis of research on SPR and a comparison of different secondary and tertiary curricula (cf. chapter 2.1), possible descriptions of a curricular development for this concept (cf. chapter 2.3.2.2) as well as a review of student problems relevant for the concept of SPR (cf. chapter 2.3.2.3), the content of the test items was chosen (Figure 25).

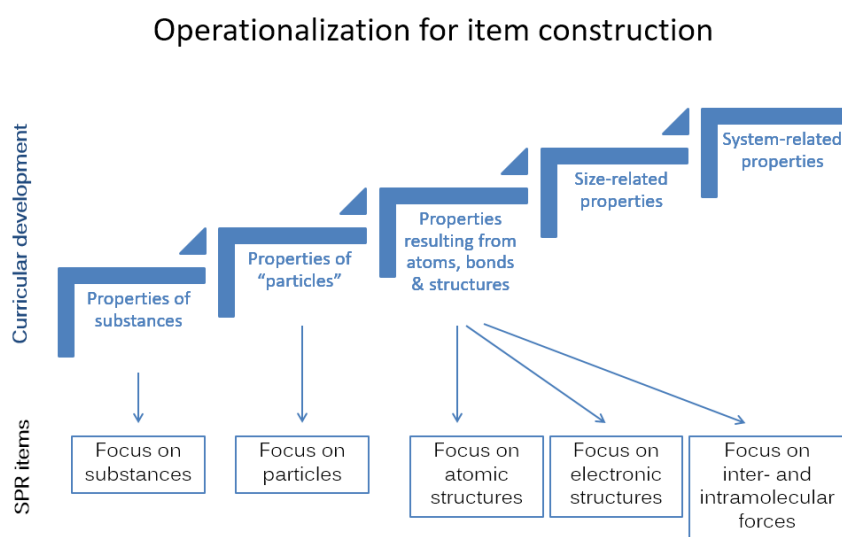


Figure 25: CK item foci based on possible curricular development and described difficulties

For examples on the phrasing of items belonging to SPR, several other sources were also consulted: standard chemistry textbooks that are used in upper secondary education in Germany²², introductory and advanced university chemistry books²³ as well as organic and inorganic chemistry university exams. The exam tasks that had been collected were analyzed in more detail according to what level they were dealing with (cf. chapter 2.3.1.2: Levels of thinking). The two central lines of demands became visible:

- a) students having to define, explain or assign reaction possibilities on a meso level (cf. MEIJER, 2011), such as reaction mechanisms, and
- b) students having to define, explain or assign certain macroscopic properties based on a submicroscopic structure (or based on the systematic name of a compound implying a certain submicroscopic structure), such as boiling point, solubility or color.

²² *Chemie heute S II* (Publisher: Schroedel), *Chemie im Kontext S II* (Publisher: Cornelsen), *Elemente Chemie S II* (Publisher: Klett)

²³ *Mortimer – Chemie* (Publisher: Thieme), *Chang – Chemistry* (Publisher: Mac Graw Hill), *Clayden et al. – Organic Chemistry* (Publisher: Oxford University Press), *Peter et al. – Organische Chemie* (Publisher: VCH)

Even though in order to explain solubility, for example, knowledge about the meso level is also needed, it is possible to still divide items into these two lines of demands, initially, as students are not always asked to *explain* properties in the analyzed exam tasks – sometimes they only have to put them in order or identify them. It also emphasizes the focus more clearly when separating these two groups: thinking in mechanisms and possible locations in a molecule where certain other molecules can attach to is more of a challenge for many students than dealing with observable properties (see among others HENDERLEITER ET AL., 2001 or NAKHLEH ET AL., 2005). Research on science teachers' misconceptions suggests that even teachers with a major in chemistry have problems distinguishing between "properties of particles (i.e., bonding, polarity, and intermolecular forces of attraction) [... and] the resulting properties of bulk matter (i.e., crystalline structure, melting point, density)" (KRUSE & ROEHRING, 2005, p.1248).

In order to include the breadth of chemistry, the generated items had to come from the main sub-disciplines of chemistry: general, inorganic, organic and physical chemistry. The resulting instrument therefore has to contain items focusing on substances, particles, atomic and electronic structures as well as inter- and intramolecular forces. Even though the notion of incorporating structure-property-relations into national curricula and suggesting a curricular development regarding this concept makes sense from a chemical perspective, studies have yet to show how an understanding of this concept develops. For that, a longitudinal study is necessary and is already planned in a following project. Before a longitudinal analysis can be conducted, however, test instruments measuring a development for this concept need to be developed. That is where this thesis aims to provide a first attempt in describing such a test instrument that focuses on the development of a structure-property-understanding at tertiary level.

These above-mentioned areas for test construction are described in a little more detail in Table 12. Only after knowing about the relevant content aspects covered by the items, other factors regarding the design can be described.

Table 12: Exemplary content of SPR-items

Items with focus on...	Exemplary content aspect of items
substances	<ul style="list-style-type: none"> • comparing solubility, acidity, aromaticity or coloredness without a detailed description of the submicroscopic level
particles and atomic structures ²⁴	<ul style="list-style-type: none"> • comparing carbon modifications graphite to diamond; • comparing salt to metal; • explaining boiling temperatures through structure; • identifying reaction products
electronic structures	<ul style="list-style-type: none"> • explaining strength of electrophiles or nucleophiles; • explaining filling of electrons in orbitals; • explaining electronic effects; • explaining tautomerism; • defining Lewis-acids or -bases
intermolecular forces	<ul style="list-style-type: none"> • defining hydrogen bonds; • explaining boiling points of different substances; • explaining solubility

4.1.2 Design of items

In order to generate a variety of items and an intended differentiation in item difficulty, several considerations have been made beyond the decision of content aspects. For other tests and other research projects attempting to assess learners' content knowledge (cf. chapter 2.2), both cognitive processes and formal aspects seem to have been important and recurring item features. Similarly, for this thesis, both features will be used and, therefore, described here. The first feature is cognitive processes, divided into complexity and number of perspectives.

4.1.2.1 Complexity

As one aspect, the Model of Hierarchical Complexity adapted for Chemistry (MHCC) (BERNHOLT & PARCHMANN, 2011) was chosen as a basis because it presents a framework that allows for tasks to be rated independently of students' answers. In this model, items are classified according to different levels of complexity as one parameter of item difficulty. While the original model includes five levels of complexity, in this dissertation project only three levels are used: *facts*, *description of processes*

²⁴ As the test was created for university students, the area of particles is not treated as a distinct area but is integrated with the atomic structures.

or procedures, and *causal relations*. Regarding the factual level, items relate to the submicroscopic structure of a compound or to its reactivity. An example for this group of items might ask students to choose the correct description of the bonds in diamond.

To solve items belonging to the complexity level of *description of processes or procedures*, students need to apply certain facts to a given situation. In one item, for example, they have to refer to their factual knowledge of nucleophilic attack, electronegativity as well as electron-pulling or electron-pushing constituents, apply it to the given compounds and assess how these effects will impact the reactivity of the carbon atom of the carbonyl function. An exemplary item belonging to this category can be seen in Table 13.

In the group of items belonging to the third level of complexity, students are required to use chemical reasoning or to explicitly verbalize causal relations, i.e. explain a phenomenon presented. Here, an item might ask students to explain the difference in states of aggregation of two substances.

4.1.2.2 Number of perspectives

Apart from complexity, a second aspect in the domain of cognitive processes contributing to an item's difficulty is the number of perspectives addressed by an item and its answer. This notion of number of perspectives is based on the chemical tetrahedron (MAHAFFY, 2006), in which the four levels of chemical representation are described as the *symbolic*, *macroscopic* and *molecular or submicroscopic* (as proposed by JOHNSTONE, 2000; GABEL, 1999) and the *context* as described by PARCHMANN, BROMAN, BUSKER & RUDNIK (2015) or MAHAFFY (2006) originally labelling it as "human contexts for chemistry" (p.51). In this way, the notion of number of perspectives follows the idea of levels of thinking (cf. chapter 2.3.1.2) without assessing each meso-step as a separate perspective. Items are classified in this dissertation describing the number of perspectives. The rationale behind that is that even though understanding aspects of and confidently maneuvering in any one of the four perspectives might be possible, the level of difficulty increases with the number of perspectives needed to come to a correct solution of the item. Table 13

shows an item addressing two perspectives: students need to make sense of the symbolic level in order to deduce the different compounds' reactivity at a submicroscopic level.

Apart from these cognitive aspects, two more characterizing features were considered: educational allocation and item format, both of which shall also be described briefly.

4.1.2.3 Educational allocation

As mentioned before, the instrument is intended to assess pre-service teachers' understanding of the SPR concept. In order to do that – while keeping in mind the heterogeneity of students' knowledge – items were generated that could be solved at various stages during the students' university education. For the introductory stages, items were constructed that can be solved with a more profound understanding of chemical aspects that are usually taught in upper secondary, while for more advanced stages, items require knowledge that is usually presented at university.

Referring back to Figure 25, in which the operationalization for test generation is presented, each step of the SPR-development can be taught and assessed by using different examples, ranging from rather common ones at the school level to rather complex ones at the university level. For example, for the curricular development step called *properties resulting from atoms, bond & structures*, one of the most prominent compounds used in secondary education is water. Students learn that water molecules consist of two hydrogen atoms and one oxygen atom which are connected by polar covalent bonds, and that the molecule itself has two lone pairs of electrons, resulting in a bent structure of the molecule. Only due to this atomic and structural makeup of the water molecules can individual water molecules be connected through hydrogen bonds, resulting for example in a higher melting temperature than would be seen in hydrogen sulfide. At the university level, the exemplary compound might be, for example, aspartame, a macromolecule having a number of functional groups, all of them contributing to the reactivity of the molecules and the properties of the substance.

Therefore, apart from complexity and perspectives, the items were also assigned an educational allocation, stating whether an item was more likely to be answered by applying school knowledge or university knowledge (cf. Table 13).

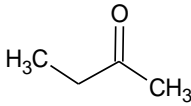
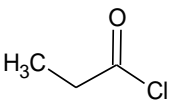
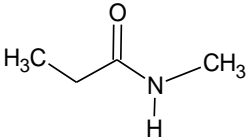
4.1.2.4 Item format

A formal item feature is its format. The items are formulated either in a multiple choice or multiple select format for assessing certain relevant facts, in a two-tier format as a combination of multiple choice and open explanation trying to assess how students justify their choices (cf. TREAGUST, 1988), or in an open format whenever students are asked to draw, describe or explain something. Even though these open items pose more of a problem when it comes to evaluating students' answers, the choice of incorporating open items was made deliberately: it is pointed out that in test situations, the answers to open items allow more insights into a students' thorough conceptual understanding (cf. MARTINEZ, 1999) compared to a pseudo-conceptual understanding (cf. VINNER, 1997). In other studies, interviews have been used in order to gauge this understanding and have provided alternative conceptions in some cases "even when students selected the correct answer [with multiple choice items]" (RUSHTON ET AL., 2008, p.129). By using open items in the test instrument, it was hoped to illicit students' knowledge that would go beyond a pseudo-conceptual understanding (VINNER, 1997). HENDERLEITER ET AL. (2001) also promote open task formats so that students would have to explain their reasoning: "Assessments should, where possible, require students to describe why and how choices were made" (p.1129).

As many of the items in the test instrument are formulated to assess pre-service teachers' understanding behind certain chemical facts, relying purely on multiple choice items did not seem the best option even if distractors had been formulated based on other students' answers. Another reason for using mainly open items is to explicitly find out about common misconceptions held by pre-service teachers and not being limited by the choice and wording of distractors. Furthermore, an observed mismatch between students' correctly chosen definitions and flawed own explanations or drawings of certain content aspects, in particular relating to hydrogen

bonding, has been reported by LEVY NAHUM, HOFSTEIN, MAMLOK-NAAMAN AND BAR-DOV (2004).

Table 13: Classification of an exemplary item according to the varying factors

Exemplary item			
Please arrange the following compounds according to their carbonyl reactivity towards a nucleophilic attack, starting with the <u>most reactive</u> compound.			
 A	 B	 c	
Complexity	Perspectives	Educational allocation	Item format
<i>description of processes or procedures</i>	<i>two (submicro + symbolic)</i>	<i>university knowledge</i>	<i>closed</i>

Summing up the above-mentioned points, and following the depiction of other competence models (cf. chapter 2.2.2 & 2.2.3), the resulting model for SPR-content knowledge ensues (cf. Figure 26), including complexity, educational allocation, number of perspectives, and content features of SPR as varying factors:

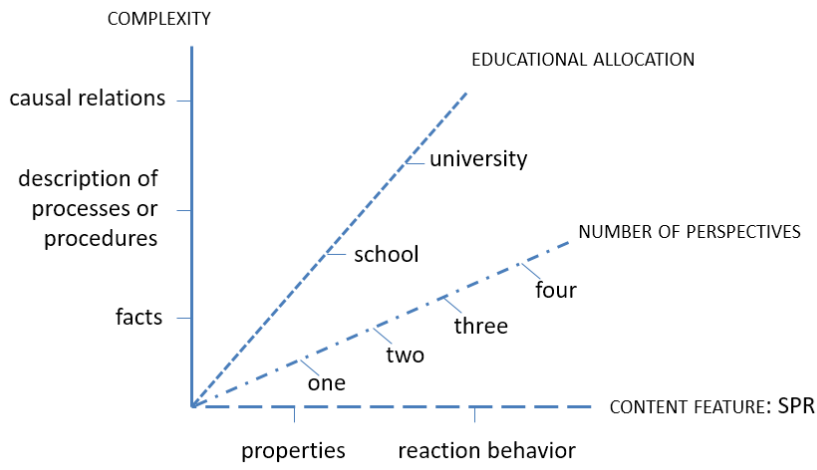


Figure 26: Resulting model characterizing CK items

4.2 Pedagogical Content Knowledge

Similar to the CK items, the items on PCK are developed in order to assess students' knowledge and not in order to test an existing model. Thus the variation in items has not been generated with the intention to fill all existing cells in the resulting model.

4.2.1 Sources for and content of items

Many of the PCK models from Table 11 have been analyzed by researchers collaborating in the *KiL*-project to arrive at a model that could be agreed on by all. As MAGNUSSON ET AL. (1999) have conceptualized their model with the science subjects in mind, it was chosen as a common base by the participating researchers of the subjects biology, chemistry, math and physics. The extent to which researchers in the project then chose to use all facets of this model (cf. Table 11) for their respective subjects was, however, left up to them. For chemistry, the two facets *student cognition* and *instructional strategies* were focused on. There are three reasons for doing this:

- (i) These are the two facets that nearly all models on PCK contain with different content foci (BAUMERT & KUNTER, 2006; BLÖMEKE ET AL., 2008; GROSSMAN, 1990; MARKS, 1990; SMITH & NEALE, 1989; TAMIR, 1988; VAN DRIEL, VERLOOP & DE VOS, 1998). They have also been selected by the *ProwiN*-group as their choice to empirically evaluate their PCK dimension as well (TEPNER ET AL., 2012). So by using them in this project, it is hoped that most researchers in the field would agree that what is tested regarding SPR in chemistry could safely be considered PCK;
- (ii) Due to the development of PCK items for both chemistry dissertation projects (cf. chapter 3), there was a test economy reason – it was decided that it would be better to aim at fewer sub-dimensions with about five to seven items each rather than aiming at the original MAGNUSSON ET AL. (1999) model's five sub-dimensions with only a couple items each;
- (iii) For other facets in the MAGNUSSON ET AL. (1999) model, for example, such as *curriculum* or *assessment*, the educational system in Germany was

considered a complicating factor because of the sixteen federal states being independent entities when it comes to curricula or school types (cf. chapter 2.1.1.). It would seem necessary to generate state-specific items that students educated in a certain system could reasonably answer. This would however defeat the purpose of a comparative large-scale assessment of PCK across the federal states.

As VEAL AND MAKINSTER (1999) pointed out, there is a distinction between the knowledge bases for PCK which address the specificity of PCK: there is *general PCK*, comprising all knowledge facets of a particular area such as science, *domain-specific PCK*, containing all knowledge facets of a particular subject such as chemistry, and *topic-specific PCK, including all knowledge facets of a particular concept or topic such as thermodynamics*. As this thesis deals with the fundamental concept of structure-property-relations in chemistry, which allows teacher and learner to consider many chemistry topics from a common perspective, the scope is in between domain-specific PCK and topic-specific PCK.

After having selected the two central facets of PCK, the literature review on secondary and tertiary students' misconceptions regarding SPR (cf. chapter 2.3.2.3) as well as oral exams of pre-service teachers and discussions in pedagogical content modules at the universities of Oldenburg and Kiel formed the basis for PCK items. In addition, PCK aspects to be used for test items arose from discussing the concept of SPR with school teachers and university lecturers as well as analyzing video sequences of pre-service chemistry teachers teaching a real Gymnasium class²⁵ in a lab setting. Even though these video sequences were taped in the 1980s and were, therefore, somewhat outdated regarding the teaching approaches used, the problems identified dealt with topics that are still relevant in teaching today: how to

²⁵ The following videos were analyzed: Class 7, topic "Reactions", © 1986; Class 9, topic "Metal-metal-reaction – formation of alloys", © 1989; Class 10, topic "Introduction to organic chemistry", © 1987; Class 11, topic "Reaction of carbonic acids in water", © 1982; Class 13, topic "Introduction to complex chemistry", © 1990; All video clips were produced by the Medienzentrum der Universität-GH Essen and were in German.

deal with incorrect student answers, how to come up with good blackboard layouts, or how to guide students to the correct answer.

In contrast to TEPNER ET AL. (2012), who incorporated experiments and models into their *instructional strategies*-facet of PCK, it was decided to create a separate facet in this study specifically focusing on problems with experiments and models, as those two aspects play a central role in the teaching of chemistry. Even though it is not specifically mentioned in the MAGNUSSON ET AL. (1999) model, this aspect can definitely be considered part of PCK as it is necessary for teachers to think of the best way to represent a certain phenomenon or content aspect or to anticipate students' problems with a certain experiment or model (cf. also SHULMAN, 1986). In these items, participants are asked to list suitable experiments for given student problems, to provide advantages and disadvantages for certain experiments or to justify the use of certain models.

In contrast to the CK items, all PCK items were to focus on problems that could authentically come from secondary chemistry classes. As PCK items cannot and should not be formulated without a specific reference to a content knowledge aspect (cf. the knowledge part of the PCK consensus model definition, chapter 2.3.2.1), the PCK items contain school-relevant content knowledge, as this would be the context in which pre-service teachers would later encounter PCK issues.

4.2.2 Design of items

In the area of *student cognition*, items focus on pre-service teachers' knowledge about student problems as well as recognizing a potential problem from a given classroom scenario. Items aiming to assess *instructional strategies* deal with certain scenarios or blackboard sketches that pre-service teachers are asked to complete, analyze or explain, thus simulating a future teacher's possible actions in a certain classroom situation (cf. SHAVELSON ET AL., 1991). Similarly, many researchers use so-called vignettes (among others RIESE & REINHOLD, 2010). With these vignettes, a classroom situation of a fictitious teacher is usually presented and (pre-service) teachers are commonly asked to describe how they would directly continue the lesson or why they would not have reacted the way the fictitious teacher had (REHM

& BÖLSTERLI, 2014). Some items in the *KiL*-study are indeed designed as short vignettes; however, they are usually written with a short item stem in order to minimize the influence of reading competence and maximize the possible number of items being tested. Other items are formulated to assess declarative knowledge, i.e. asking pre-service teachers to define or classify student answers. The third area, *experiments and models*, requires pre-service teachers to select experiments in order to explain a certain concept or to provide alternatives to certain models.

Coding PCK questions often appears less clear than coding CK items because answers cannot be divided as strictly into the categories of false and correct. As the PCK items are constructed with real-life situations in mind, however, it was decided to solely use an open item format²⁶. This format stresses the real-life orientation of the tasks as “life rarely presents itself as a question with a set of possible answers, with only one being correct” (SHAVELSON, 2010, p.47). In addition, it was hoped to receive more authentic insights into pre-service teachers’ understanding regarding pedagogical content knowledge aspects when giving them the opportunity to provide answers in their own words: Assessing introductory university students with the Chemical Concepts Inventory, POTGIETER, ROGAN AND HOWIE (2005) concluded from their findings that “[a]lmost half of correct answers provided for the multiple-choice questions were followed by an inadequate or incorrect explanation” (p.127).

As many of the items are constructed resembling authentic teaching situations, several kinds of knowledge are tapped into: apart from declarative knowledge, a stronger emphasis is laid on pre-service teachers’ procedural and conditional knowledge (PARIS ET AL., 1983). Simply assessing declarative knowledge is not enough in the realm of PCK, as one teacher stated when confronted with her own alternative conceptions, “knowing and teaching are two different things” (KRUSE & ROEHRING, 2005, p.1248). Pre-service teachers may be well aware that students have, for example, numerous problems with switching back and forth between the macro and submicro levels of chemistry; they may not, however, know how to proceed in a teaching situation in order to overcome these problems. Therefore, items assessing the procedural aspects of pre-service teachers’ PCK are added. Even more important

²⁶ The initial intention was to eventually formulate closed items gained from the answers provided by pre-service teachers, but that was not meant to be done within this dissertation.

when it comes to teaching situations is the third kind of knowledge: conditional knowledge. In this category, the conditions of classroom situations are analyzed and the questions of when and why (PARIS ET AL., 1983) certain actions might be hindering or promoting student understanding are addressed. The three kinds of knowledge were also used by the *ProwiN*-project as one element in their conceptualization of teachers' professional knowledge (TEPNER ET AL., 2012). This three-part division can roughly parallel the adaptation of the model of Hierarchical Complexity adapted for Chemistry (cf. BERNHOLT & PARCHMANN, 2011), as presented in the description of the CK knowledge area (cf. chapter 4.1.2.1): i) declarative knowledge corresponds to the facts level, as both deal with knowing that principles or laws exist; ii) procedural knowledge can be considered analogous to the description of processes or procedures, as both areas deal with knowledge centering on the idea of how to do something; and iii) conditional knowledge parallels the causal relations-level as both aspects put into focus the conditions under which certain actions could be reasonably carried out (Table 14).

Table 14: Comparison of kinds of knowledge to complexity levels for PCK

Kind of knowledge (according to PARIS ET AL., 1983)	Complexity level (adapted from BERNHOLT & PARCHMANN, 2011)	Cognitive task in PCK items
declarative	facts	stating that teaching approaches, misconceptions or concepts exist; providing or choosing correct definitions or examples
procedural	description of processes or procedures	stating how to do something, act in a given situation, or use a certain teaching approach
conditional	causal relations	stating under which conditions certain actions could be reasonably carried out

For both categorization schemes it is commonly understood that within one content area, the complexity increases from declarative knowledge or facts to conditional knowledge or causal relations, respectively, as the mere number of aspects taken into account increases along with those aspects' interrelations. Just as

the different factors in the CK dimension of this thesis are not varied in a systematic way solely to fill all cells in a model, the division into the three kinds of knowledge is not used to yield an approach to model testing; rather it is used in order to generate a variety of items.

Based on the above-mentioned aspects, the resulting model for pedagogical content knowledge is presented in Figure 27.

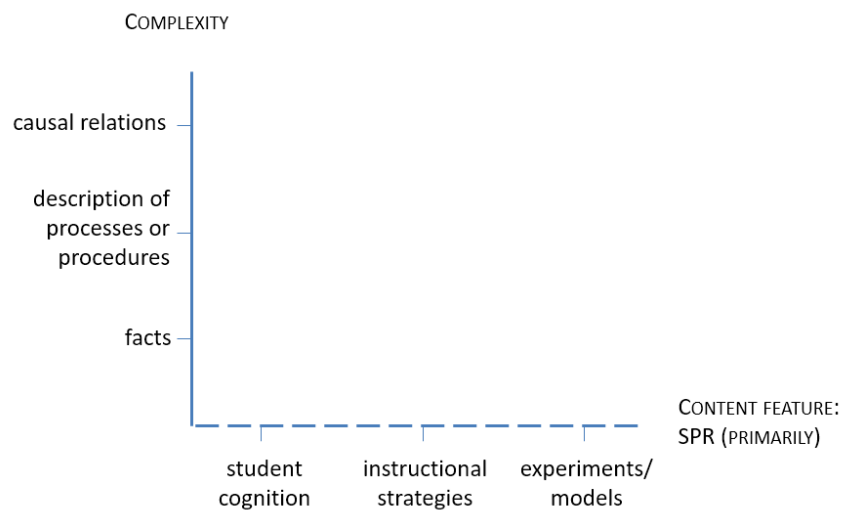


Figure 27: Resulting model including relevant aspects characterizing PCK items

By comparing the models on CK and PCK, one can see that the content aspects of the SPR field are different ones: properties and reaction behavior for the CK area and student cognition, instructional strategies and experiments/models for the PCK area. While the factor *complexity* is included in both models, the factor *number of perspectives* is only found in the CK model as that factor is strictly content-related. The factor *educational allocation* is eliminated from the PCK model as all items refer to classroom situations of pre-service teachers' future career.

4.3 Communicative item validation

CONTENT KNOWLEDGE ITEMS

Groups of researchers, lecturers and teachers discussed the CK-items and possible solutions. In addition, the items were also assessed by three external university professors lecturing in organic or inorganic chemistry at different institutions. The experts were asked to assess whether the phrasing of an item was understandable for students, whether the item touched a central aspect of the SPR concept, whether in their opinion the item addressed more of a factual or more of a conceptual level, whether the lecturers themselves would cover the content or even the explicit example in their teaching and whether they considered the overall item to be – by trend – easy or difficult for university students to solve. Whenever the experts saw problems in the understandability, missing relevance or extremely high level of difficulty in an item, items were revised or discarded. It should be noted at this point that the experts expected more prior knowledge on the part of pre-service chemistry teachers than had been intended when generating the items: For example, when the name of a compound was provided along with the compounds' structure, the experts expressed that the structure would be providing the students with too much information as tertiary level students should be able to make that translation from name to structure by themselves. At this particular point, it was decided not to follow the experts' advice because it would be adding a potential further pitfall to answering an item. Instead, pre-service teachers should be given the opportunity to work with explicit structures instead of names, as the concept under consideration is exactly the relation between structures and properties.

A second phase of CK-item revision took place before administering the test, when about a third of the items were given to five pre-service teachers in the middle of their university education. These students were asked to verbalize their understanding, so that problems in phrasing or content became clear. Those items were then also rephrased to eliminate difficulties. Furthermore, it was checked that the anticipated time frame for solving the items would be sufficient. The number of items in study I was chosen according to this time calculation.

The issue of open items being less easy to assess was addressed to a certain extent by having experts provide exemplary answers that they would expect pre-service teachers to give, and by discussing the coding of answers both within the research group and with other researchers working in the field of testing pre-service teachers' knowledge.

A further aspect of validating the relevance as well as approaches to coding the CK-items was to discuss them with international researchers. Two rounds of discussion took place with research groups in Ireland. One group had – among other things – gained expertise in judging content and skill levels relevant for and manageable by both secondary level chemistry teachers and students through doing this as part of several EU-projects and lecturing at university level. A consequence resulting from this discussion was the exchange of some items in the item pool that were asking the students to identify rather specific functional groups. For those items, others were constructed that made the structure-property-link more explicit. A second research group had generated research on areas of difficulties in the sub-discipline of organic chemistry for secondary students (CHILDS & SHEEHAN, 2009; O'DWYER & CHILDS, 2011) as well as on chemistry misconceptions that are still present in pre-service teachers (SHEEHAN, CHILDS & HAYES, 2011). Even though these studies had already been considered in item collection and generation, the personal exchange with these researchers provided helpful information in item revision after the piloting phase: the explicit linking between structures and resulting properties was stressed by inserting structural formulae in some of the items containing less well-known compounds, for example. With both groups of researchers, the wording of the (translated) items was also analyzed, pointing out that occasionally items could be phrased in a more precise way.

PEDAGOGICAL CONTENT KNOWLEDGE ITEMS

Regarding the items on PCK, experts and international researchers working in the field were also asked to provide feedback. Teachers as well as university professors for teaching chemistry were asked to fill in the same item rating form that was used for CK. Therefore, items were assessed regarding their wording, regarding their central

role in the SPR concept, regarding the items factual or conceptual level, reading the item being explicitly used in the teachers' or professors' teaching and regarding the item's difficulty. Based on this rating of items, some items were revised to increase understandability, relevance or alter the level of difficulty of an item.

A second phase of item revision involved a small think-aloud study with a non-representative sample of four pre-service chemistry teachers in their first Master semester. This study focused mainly on understandability and also on the allocated test time. Items that the majority of these students could not answer were discarded or specified whenever wording was considered the underlying problem.

In order to assess the teaching-specific character of the PCK items, they were also answered by two PhD students who had received a master degree in chemistry and had neither taken nor given any courses in teacher training. Even though this was only a very small and non-representative sample, their answers pointed out that they did not know specific experiments relevant for the teaching of chemistry at school, were not aware of possible misconceptions resulting from certain experiments and instructional decisions and could not extract underlying problems from hypothetical student answers. Instead of naming advantages or disadvantages of given general experiments, they focused on a mere content-oriented explanation: When asked, for example, to state two aspects of the experiment *Burning of iron wool* that could be discussed with beginner students and with more advanced students, they did not know possible distinctions but rather replied that iron oxide formed, triggered only by the content part of that item. After answering all PCK items, one of them admitted “*I have no idea what's going on in students' minds*” (Chemistry PhD student A). Pedagogical content knowledge items that had a strong CK part could, in general, be answered by both chemistry majors, such as listing a certain number of alkanes that students know from every-day life or providing an experiment that showed that both graphite and diamond are made of carbon atoms.

As with the CK items, the PCK items were also presented to international researchers active in the field of PCK. At this point, particularly the integration of some CK aspects into the PCK items was discussed in order i) to make the items topic- or concept-specific, and ii) to enable a more specific coding of prospective student

answers because general or stereotypic answers such as “*If a student doesn't understand something, I would explain it again*” would not justify a full score. In addition, some items were rephrased to be more authentic for teaching situations so that answers would give insight into how these future teachers might react in an actual teaching situation (cf. SHAVELSON ET AL., 1991).

5 STUDY I: ANALYSIS OF TEST INSTRUMENTS

5.1 Research objective

The main goal of this first study is to find out how the test items would work, i.e. evaluate whether the range of items was able to assess pre-service chemistry teachers' understanding of content knowledge and pedagogical content knowledge regarding SPR. From the large pool of items developed, the goal was to select as many items as possible that fulfilled the standard test criteria (i.e. proper level of difficulty, high degree of variance and an acceptable item-total-correlation) and that could be used in a second study focusing more strongly on the relationship between the dimensions of pre-service teachers' professional knowledge. It was also intended to get first ideas of features influencing item difficulty. Therefore, the resulting overall research objective guiding this first study is the following:

- O_{study I}:** To create valid, reliable and objective test instruments with which to assess pre-service teachers' chemistry content knowledge and chemistry pedagogical content knowledge focusing on structure-property-relations and to investigate possible influencing factors.

This overall research objective can be further divided into smaller goals, separated according to their focus on content knowledge (G1-G3) and pedagogical content knowledge (G4-G5):

- G1:** To create a valid and reliable CK instrument.

Hypothesis: A reliable and valid instrument assessing pre-service teachers' content knowledge on structure-property-relations can be developed. This is the basis for the following goals.

- G2:** Based on the CK instrument, to identify factors that influence item difficulty.

Hypothesis: According to BERNHOLT AND PARCHMANN (2011), CK items with higher complexity level will prove more difficult. Similarly, the more

perspectives are required to solve an item, the more difficult the item will be (MAHAFFY, 2006). Items requiring university knowledge will be more difficult than items with a secondary educational allocation.

- G3:** Based on the CK instrument, to assess how pre-service teachers score on the CK test.

Hypothesis: As the test includes items on school and university level, and all participants will be enrolled in varying semesters of educational programs for chemistry teachers at tertiary institutions, pre-service teachers in a higher semester will score better on the instrument than those beginning their university education as they should have had more learning opportunities.

- G4:** To create a reliable and valid PCK instrument.

Hypothesis: A reliable and valid instrument assessing pre-service teachers' pedagogical content knowledge on structure-property-relations can be developed. This is the basis for the following goal.

- G5:** Based on the PCK instrument, to assess how pre-service teachers score on the PCK test.

Hypothesis: As the test only includes items on school level, some of which the pre-service students might have encountered themselves in their secondary education, and all participants will be enrolled in varying semesters of educational programs for chemistry teachers at tertiary institutions, pre-service teachers in a higher semester will score better on the instrument than those beginning their university education as they should have had more opportunities to consider problems from a PCK perspective.

5.2 Test administration

Regarding the administration of the paper-and-pencil tests, pre-service teachers of eleven educational institutions throughout Germany were simply asked to come to a lecture hall at the university they were attending. The students themselves were at different stages in their programs and therefore had taken a variety of courses in chemistry and chemistry education: students were allowed to participate at any stage of their studies. Tests were administered in the summer of 2012 and alternating members of the KiL-project were present at each university in cooperation with local university staff, providing test instructions. As can be seen from the organization slide that was shown at the beginning of each testing session (Figure

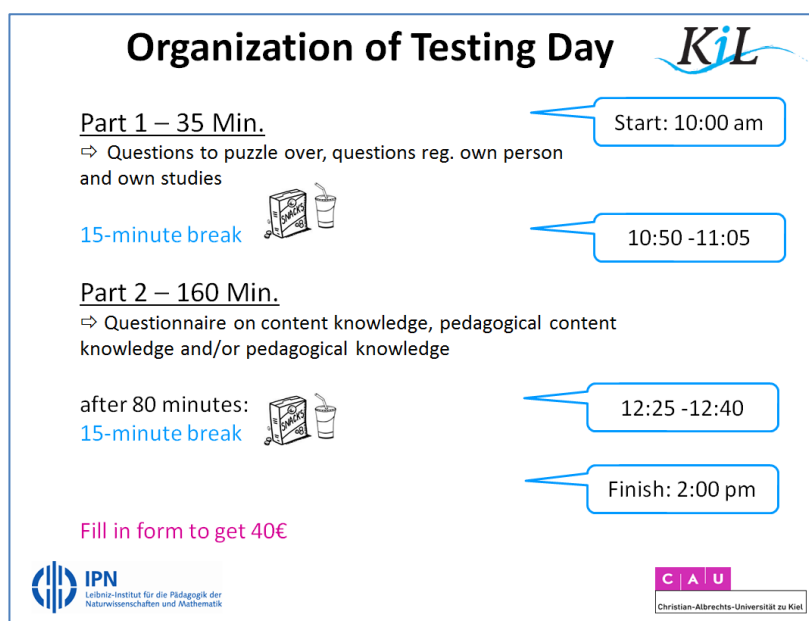


Figure 28: Overview of testing session

28), the four-hour testing day started with the cognitive ability test (for motivational purposes disguised as *questions to puzzle over*) and questions on demographic data. As students' cognitive ability is considered to be an indicator of their intelligence, predicting learning progress and performance in many studies (e.g. HATTIE & HANSFORD, 1982) and are said to have relatively high correlations to science competencies (STEINKAMP & MAEHR, 1983), students were asked to provide two cognitive ability measures: their grade in their school leaving certificate (GPA of their Abitur-exam; ranging from 1 being an excellent grade to 4 being a poor grade) and a cognitive ability test (HELLER & PERLETH, 2000). Regarding learning opportunities at university, students were also asked to provide their semester at university. The semester was expected to correlate with increased scores on both test instruments as it should be assumed that pre-service teachers increase their understanding of content and pedagogical content knowledge throughout their university education.

After a 15-minute break, students answered those items pertaining to the subject-specific professional knowledge (and pedagogical knowledge). This session was again interrupted by a 15-minute break. As the project's testing sessions were to not interfere with regular university classes, Saturdays were commonly used as test days. In order to still attract a sufficiently large number of students, each student was issued a financial incentive of 40 Euros.

Regarding the test time, each subject was given one hour each for content knowledge and pedagogical content knowledge. Due to the division into the two chemistry areas (i.e. chemical representations and structure-property-relations), the actual test time for SPR was about half an hour each.

5.3 Data preparation & statistical measures

Students' answers were digitalized and coded according to their appropriateness and reasonability. As described before, standard phrases were not rated as high as specific answers. When asked how to react to a wrong student answer regarding the formulation of a net equation, for example, a possible answer such as *"I would explain it again"* would not be classified as appropriate as the hypothetical statement: *"I would ask the student to write down all participating compounds, ask him to take the compounds apart into ions, then cross out those ions that are identical on both sides of the equation"*. What makes the second answer more appropriate is not its detailed description but its specificity to the given situation – the first answer is transferable to any situation in which a student does not understand a particular aspect.

For most of the CK and all of the PCK items, a partial credit coding was chosen, with 0 being coded for a false, 1 being coded for a partially correct and 2 being coded for a completely correct answer. For closed items, a mere differentiation of false/true resulted in a 0/1 coding. As had been checked in prior rounds with several students in the middle of their studies, the allocated time for answering the items was sufficient for more medium to advanced students. Therefore, missing data were assumed to be due rather to a lack of knowledge (or unwillingness to answer) than to time restriction. As a result, answers that had not been filled out by pre-service students in this study were assigned a 0 (false) coding.

After analyzing the data, a second person coded roughly 20% of the answers to each item. The resulting Cohen's kappa values for the CK categories were $\kappa_{\text{Complexity}} = .89$, $\kappa_{\text{Perspectives}} = .91$ and $\kappa_{\text{Educational allocation}} = .83$, all of which suggest moderate to good inter-rater reliability. For PCK, the inter-rater reliability showed an overall Cohen's kappa value of $\kappa = .92$, showing a good inter-rater reliability.

In order to arrive at a reliability assessment, items have to first be selected based on a psychometric processing of the obtained data, which is evaluated in regard to the items' difficulty, variance, and item-total correlation coefficient. This is done through classical test theory with the program *SPSS 23.0*. Classical test theory is based on the assumption that every observed score on a test is the sum of a true score and an error score (ROST, 2004). From the observed data, a score is calculated that contains an error term, which cannot however, be quantified. For a first approach to the data, this approach is quite useful.

The difficulty of an item reveals the percentage of the test population having correctly solved that item; thus a higher value would indicate an easier item (KELAVA & MOOSBRUGGER, 2008). Values for item difficulty should be in the range of 0.25-0.75 for conceptual tests. The item-total correlation coefficient of a test item is a correlation between a person's scores on an individual item and his/her score on the entire test. It can assume values in a range of -1 to +1. A high positive value means that a person with a high total score is more likely to answer the item correctly than a person with a low total score. The item-total correlation coefficient should be greater than or equal to 0.2; however, items with a value lower than 0.2 may remain in a test as long as there are not too many of them (DING, CHABAY, SHERWOOD & BEICHNER, 2006).

For evaluating the reliability of the scales, the values for Cronbach's α are used. Cronbach's α is a measure of a test's internal consistency; i.e. it states whether similar items have been answered in a similar way. For cognitive tests, a Cronbach's α value of .8 is generally considered acceptable, while a cut-off value of .7 is more suitable for ability tests. In addition, a content-driven selection process compliments this statistical analysis.

In order to compare the means in test scores or outcomes, a *t*-test is used when two groups are considered. When wanting to find out whether the means of more

than two groups differ in a statistically significant way, an ANOVA (i.e. analysis of variance) is calculated. The values provided by this calculation do not provide insights into how the various groups differ, however. To obtain information on the actual differences between the groups, a post hoc test according to Scheffé will be performed.

When looking at the data in more depth, item response theory (also called latent trait theory) becomes more appropriate. As the name implies, this approach assumes i) the presence of not directly observable, i.e. latent, traits for the constructs under consideration and ii) that items in a test are answered based on a person's ability in those latent, i.e. unobserved, traits. As the difficulty of items enters into the mathematical calculation of a person's ability regarding the trait, abilities of test takers and difficulties of items can be considered simultaneously, pointing out items that are too easy or difficult for the tested population. This theory allows statements about the probability that a person might correctly answer a certain item. The majority of items should be in the range of \pm one standard deviation around item parameter and person parameter mean score (KAUERTZ, 2014). Through this approach, conclusions about the extent of that trait in that population can be drawn. For analyses based on item response theory, the software package *ConQuest* was used.

In order to provide consistent guidance to the reader, the term *(test) score* will be used when measures of classical test theory are looked at focusing on a person's sumscore of a test instrument, while in analyses based on item response theory the term *(test) outcome* will be referred to as person ability regarding the items in a test instrument are evaluated.

5.4 CK and PCK instruments

From the item pools for content knowledge and pedagogical content knowledge regarding SPR, a certain number of items had to be selected for study I. Several considerations guided this first round of item selection: i) a maximum allocated test time for each subject, ii) the possibility of coordinating items of the two chemistry areas SPR and Chemical Representations (the adjacent dissertation by TASKIN, 2014), and iii) testing as many items as possible. As the area of chemical representations

focused on various ways of representing chemical formulae, e.g. sum formulae, ionic lattices, Lewis structures, and ball-and-stick models, these structural considerations were an intended overlap to part of the larger SPR concept (cf. COOPER ET AL. (2012b) and Figure 24). The focus of the items in this dissertation was more on the actual interaction, i.e. logical inferences, between structures and properties. This was especially true for the content knowledge items: one item asked the students to name limitations of a certain representation of a compound (in the chemical representations-part) and a second item would require the students to order this and other compounds based on their boiling temperatures (in the SPR-part), for example. However, whenever one item could be assumed to be a direct prerequisite for another item, thus making those two items dependent on each other, the items were assigned to different test booklets. The use of multiple test booklets enabled a trialing of many items. As the goal was to reach a total of 150 pre-service chemistry teachers, it was decided to use three test booklets. With the three booklets being assigned randomly to the participants, each item would be answered by about 50 students. Through this number, enough information about each item was meant to be obtained.

Each booklet contained a certain number of content knowledge as well as pedagogical content knowledge items from both areas (Table 15, items on chemical representations, called CRep, are also mentioned in order to give an overview of the complete subject-specific test booklets). A portion of content knowledge items regarding SPR, i.e. 25 items, overlapped between test booklet A and B.

Table 15: Assignment of items to test booklets (items on CRep are added in gray here for the sake of completeness)

Test booklet A	Test booklet B	Test booklet C
PCK (CRep): 24 items	PCK (SPR): 22 items	CK (SPR): 26 items
CK (CRep): 28 items	CK (SPR): 42 items	PCK (SPR): 20 items
CK (SPR): 51 items	CK (CRep): 31 items	PCK (CRep): 22 items

Regarding content knowledge, items were presented in a variety of formats; the detailed characterization of all items can be seen in Table 16. For SPR, 30.9% of the

content knowledge items were in a short answer format, 39.4% were in a multiple choice format, 7.4% were two-tiered tasks thus asking a selection and subsequent explanation or drawing, and 4.3% were multiple select items, where more than one of the provided options was correct. Regarding what pre-service teachers actually had to do in an item, 6.4% asked the students to put compounds in a logical order or assign them to certain groups and 11.7% involved some kind of marking or drawing. Each test booklet contained a fairly even mix of items in regard to the two content aspects: focusing on properties (items with *E*– from German *Eigenschaften*– in their name, cf. chapter 5.5.2) vs. focusing on reaction behavior (items with *R*– from German *Reaktionsverhalten*– in their name, cf. chapter 5.5.2). Furthermore, all sets included a comparable proportion of organic vs. inorganic chemistry as well as a comparable distribution regarding complexity and perspectives.

Table 16: Item formats per test booklet (TB)

	TB A		TB B		TB C		Total (%)	
	CK	CK _{net} ²⁷	CK	PCK	CK	PCK	CK	PCK
Multiple choice	29	18	14	-	5	-	39.4	-
Multiple select	2	1	1	-	2	-	4.3	-
Short answer	9	3	13	21	13	18	30.9	90.7
2-tier	2	2	3	1	2	2	7.4	7.0
Ordering/matching	4	1	5	-	1	1	6.4	2.3
Marking/drawing	5	1	6	-	2	-	11.7	-

For pedagogical content knowledge, items on SPR are only included in test booklets B and C due to test design issues (i.e. the combination of SPR items and those on chemical representations by the other dissertation project). In both remaining test booklets, the vast majority of items (90.7%) regarding SPR are in an open format, asking either for short explanations or for examples. In both test booklets, there is an even distribution of items belonging to the three sub-dimensions (i.e. *student cognition*, indicated as *S* in item name; *instructional strategies*, receiving

²⁷ CK_{net} is the number of items reduced by the overlapping 25 items between test booklet A and B. These numbers are used to calculate percentages.

an / in the item name, and *experiments and models*, being referenced by *EM* in the item name).

5.5 Results

5.5.1 Demographic data of the participants in study I

220 participants from eleven tertiary education institutions across Germany took part in the chemistry-part of the *KiL*-study in 2012. The average participant's GPA was 2.3 ($SD = 0.5$)²⁸. At the time of the data collection, participants were on average in their 5th chemistry semester ($SD = 3.0$). Students were on average 23.1 years old ($SD = 3.0$) and most of them were female (60.5%).

²⁸ with 1.0 being the best and 4.0 being the worst passing grade

5.5.2 CK – individual item analysis

A total of 94 content knowledge items regarding SPR were given to pre-service teachers in this study. The first step in the process of generating a reliable instrument was to look at how difficult or easy the items were for the students. In order to do this, Figures 29 to 31 present the statistical difficulty of each item in the study, grouped into the respective test booklets. From a test theory point of view, all items with difficulty values below .2 should be discarded as they are too difficult and items with a difficulty value above .8 are excluded from further analysis as well because they are too easy for the students. For test booklet A, this meant that items²⁹ E12, E13, E15, R9, R10, R23, R24_2, R27_3u4, R31_1, R31_3, R31_4, R31_6, and R31_8 should have been eliminated. For test booklet B, items E22, E12, E26, R2, R10, R18, R23, R24_1, R24_2, and R27_3u4 should have been excluded due to this criterion. And for test booklet C, items E11, E21, E25, E34, R19_1, R19_2, and R36 should not have been considered further. Even though these items should have been eliminated from a test theory perspective, content-driven aspects needed also be taken into account: by keeping some of the items that have a difficulty value of $<.2$ (i.e. rather difficult items), the resulting test instrument can also be used to measure a development in CK as part of possible longitudinal studies. For test booklet B, for example, such items were E22, R18 and R23.

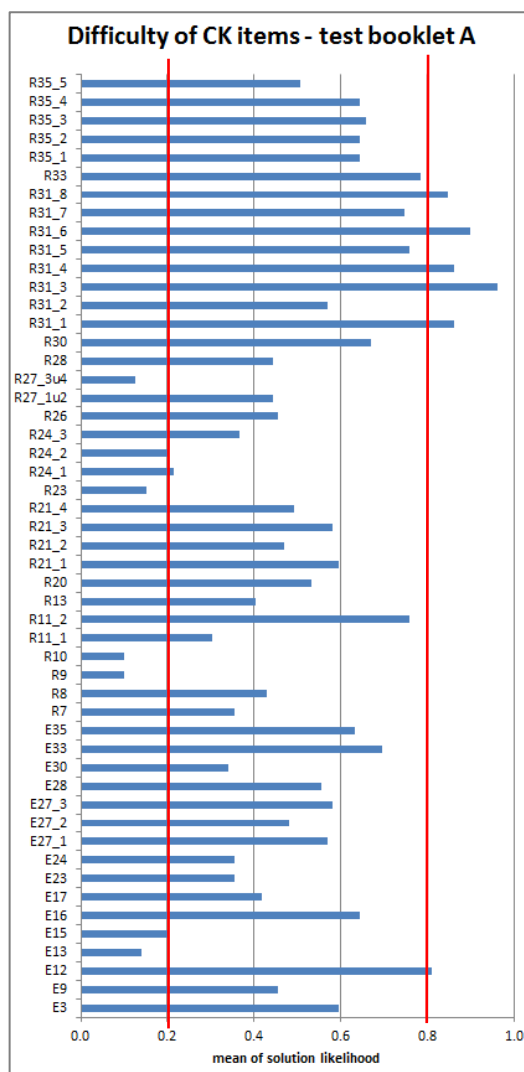


Figure 29: Difficulty of CK items from test booklet A

²⁹ As the items are currently being used in a subsequent research project, not all items can be shown in this dissertation. Exemplary items can be seen in Tables 13 and 26 and are described in chapter 5.5.2 (for CK) and 5.5.5 (for PCK).

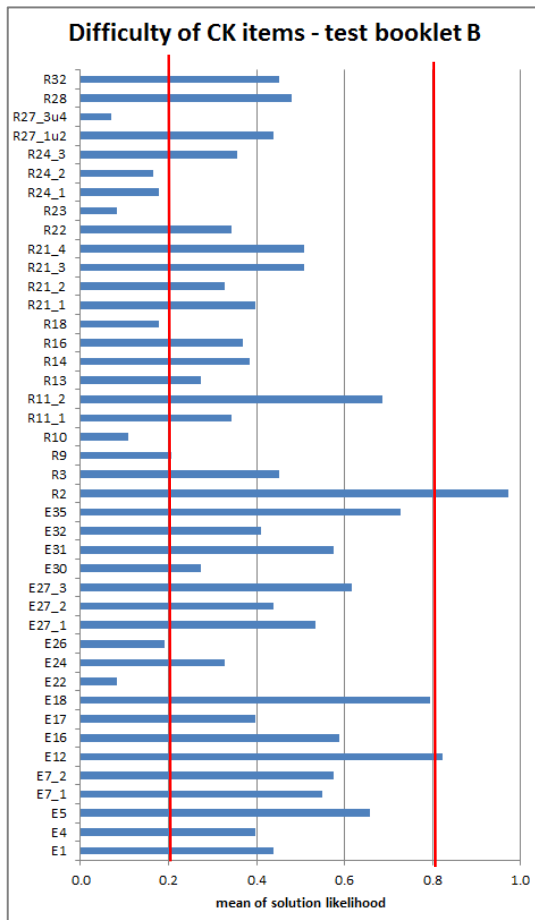


Figure 31: Difficulty of CK items from test booklet B

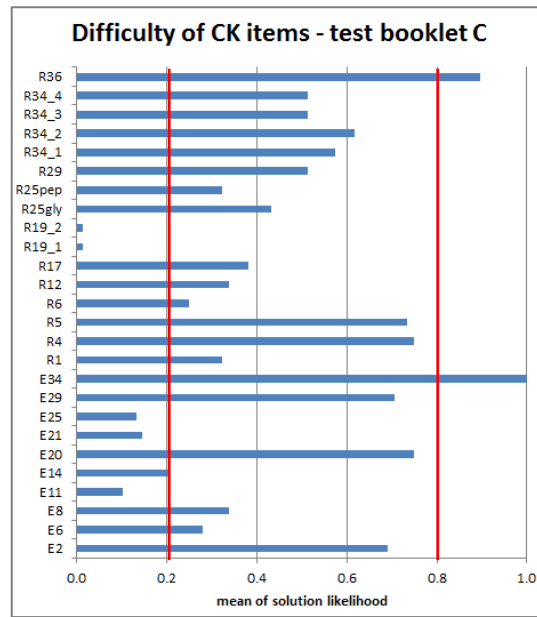


Figure 3130: Difficulty of CK items from test booklet C

While Figures 29 to 31 present the statistical difficulty for each item, the question on how students performed in the resulting instrument will be looked at in chapter 5.5.7. At this point, exemplary items with corresponding student answers shall be presented. The selection of items was made based on two criteria: i) a content-driven criterion, meaning that items of a large subarea of the SPR-concept were chosen which included a variety of items which differ in their difficulty, and ii) a research methodological criterion, meaning that items were chosen that had already been presented at conferences and for which the categorization can therefore be

considered publicized³⁰. A selection of items focusing on intermolecular bonds in general and hydrogen bonds in particular is chosen for this closer analysis.

QUALITATIVE ANALYSIS OF SELECTED CK ITEMS

This quite general look at the items does not yet reveal anything about why certain items might be easy or difficult. In order to get closer into students' difficulties, and thus receive information about the instruments' content validity, a more detailed look at selected items and in particular at students' answers was necessary. For the assessment of correctness, answers were evaluated from an in-depth content perspective, so that school knowledge was sometimes not enough to get full credit (e.g. drawing an ethane molecule without the hydrogen atoms but with the bonds to the omitted hydrogen atoms still in place is commonly taught as a correct depiction in school but is wrong from a chemical perspective). The perspective was general enough, however, so that highly specialized knowledge was neither needed nor asked for in answers and therefore not categorized as correct if contrary to common examples and not explained further (e.g. stating that ether molecules can form hydrogen bonds with other ether molecules is, without further explanation by the participant, considered wrong even though there might be instances of highly specific reaction conditions that might allow for exceptions).

In test item E1, for example, students were asked to put the given alcohols butan-1-ol (A), tert-butanol (B), and butan-2-ol (C) in the correct order of descending boiling temperatures, and to briefly explain their ordering. From the given names of the alcohols students should have deduced the structural make-up of each molecule (with butan-1-ol having the most linear and tert-butanol the most spherical structure) and should then have inferred based on the structure which molecule would lead to the largest intermolecular forces with a second molecule of the same substance (two molecules with linear structure will have largest overall sum of intermolecular forces as a longer chain can form van-der-Waals forces over a larger surface area while a

³⁰ As the test instruments are currently being tested in a longitudinal study, the explicit coding for as few items as possible should be publicly available at this point.

spherical molecule can only form van-der-Waals forces over a relatively small surface area).

Students' answers showed a variety of answer patterns, which were classified and then categorized as correct, partially correct and incorrect, according to Table 17. In Table 17 – and all subsequent tables providing exemplary student answers – pre-service teachers' original German explanations are given and then translated into English, remaining as closely to the original as possible. 21.8% of students provide a correct explanation together with a correct ordering of the alcohols. Almost as many (19.1%) provide at least a partially correct explanation. Over half of the students (59.0%) do, however, give an incorrect answer. Of the three categories, the one classified as “incorrect” might be most interesting for further analysis: 19.2% of students (this subgroup being on average in their 6th chemistry semester, range: 2 – 11, \bar{X} 5.6) gave an either incorrect or correct ordering but no explanation. The remaining explanations in the “incorrect” category (i.e. regardless whether the students put the alcohols in correct or incorrect order) could be assigned to one of four main error groups: I) reference to polyhydric alcohols (10.2%), II) reference to bond breaking within molecules (9.0%), III) reference to polarity of molecules (1.3%), and IV) other explanations not making any sense (15.4%). Students in all three of these groups are on average in their 4th chemistry semester (range_i: 2 – 10, \bar{X} 4.3; range_{ii}: 2 – 10, \bar{X} 4.0; value_{iii}: 4).

Table 17: Answer patterns and coding scheme for item E1

Answer pattern	Exemplary student explanation	%		%*
no order, no explanation	--	3.8	incorrect	59.0
incorrect order, no explanation	--	17.9		
incorrect order, incorrect explanation (various)	„Je "kugelförmiger" das Molekül ist, desto besser können sich van-der-Waals-Bindungen ausbilden.“ (P0589) [the more spherical the molecule, the better van-der-Waals-forces can form]	12.8		
incorrect order, incorrect explanation (bond breaking)	„Butan-1-ol lässt sich am einfachsten spalten, da die funktionelle Gruppe am Rand ist.“ (P0185) [butan-1-ol is the easiest to break off as the functional group is at the end]	9.0		
incorrect order, incorrect explanation (polyhydric alcohols)	„tert-Butanol = tertiärer Alkohol. Hat mehr Substituenten, mehr Wechselwirkungen, die das Molekül stabiler/träger machen.“ (P0593) [tert-butanol = tertiary alcohol, has more substituents, more intermolecular forces which make the molecule more stable/inert]	9.0		

5 STUDY I: ANALYSIS OF TEST INSTRUMENTS

correct order, no explanation	--	1.3		
correct order, incorrect explanation (various)	„A hat einen Liganden an erster Stelle, wodurch Siedetemperatur schneller erreicht wird als bei C, da dort der Ligand an zweiter Stelle steht und bei B.“ (P0290) [A has a ligand in the first position, thus reaching boiling temperature faster than C because there the ligand is in the second position and B]	2.6		
correct order, incorrect explanation (polarity)	„A hat eine höhere Polarität → Wasserstoffbrückenbindungen und somit einen höheren Siedepunkt als die anderen angegebenen Alkohole.“ (P0139) [A has a higher polarity → hydrogen bonds and thus a higher boiling point than the other alcohols]	1.3		
correct order, incorrect explanation (polyhydric alcohols)	„A ist ein einwertiger Alkohol, C ein zweiwertiger und B ein dreiwertiger Alkohol“ (P0333) [A has one hydroxyl group, C has two hydroxyl groups, and B has three hydroxyl groups]	1.3		
correct order, explanation using rule-based knowledge	„Je verzweigter der Stoff desto geringer die Siedetemperatur.“ (P1132) [the more branched a substance, the lower its boiling point]	3.8		
correct order, partially correct explanation	„Durch die lange Kette bei A verzweigen sich die Moleküle, weshalb sie nicht so leicht in Gasphase zu bringen sind. B ist kleiner und kompakter, daher kann es nicht so leicht verzweigen.“ (P0597) [Due to the long chain in A, these molecules are more branched, so they are not as easily transferred to the gaseous state. B is smaller and more compact, so it cannot branch as easily]	11.5	partially correct	19.1
incorrect order, correct explanation (partially or rule-based knowledge)	„Umso länger Kohlenwasserstoffketten werden, desto höher werden Siedetemperaturen.“ (P0844) [the longer carbohydrate chains get, the higher their boiling temperatures get]	3.8		
correct order, correct explanation	„Der Alkohol mit der höchsten Siedetemperatur muss die größten zwischenmolekularen Kräfte haben. Da der primäre Alkohol am wenigsten "sterisch" gehindert ist und sich die Moleküle gut aneinander lagern können hat er die höchste Siedetemperatur gefolgt vom sekundären und dann erst der tertiäre Alkohol.“ (P0613) [The alcohol with the highest boiling temperature has to have the largest intermolecular forces. As the primary alcohol is 'sterically' hindered to the least extent and molecules can align well, it has the highest boiling temperature. It's followed by the secondary alcohol and then the tertiary alcohol.]	21.8	correct	21.8

*The deviance from 100% is a result due to rounding to the first decimal place.

A second item (E5) asked students to decide whether butan-2-ol or diethyether is better soluble in water and to provide an explanation for their choice. In the example, the decisive factor for solubility in water is the formation of hydrogen bonds. Therefore, students need to know that molecules of both compounds will be able to form hydrogen bonds with water molecules. As an ether molecule has only the free electron pairs at the oxygen atom for forming hydrogen bonds with water molecules,

two hydrogen bonds can be created per ether molecule. Butan-2-ol, having identical free electron pairs at the oxygen atom but having in addition one positively polarized hydrogen atom of the hydroxyl group, can therefore form a total of three hydrogen bonds per molecule. That is why based on the structural makeup, butan-2-ol is better soluble in water than diethylether.

Again, students' answers showed a variety of answer patterns, which were classified and then categorized according to Table 18.

Table 18: Answer patterns and coding scheme for item E5

Answer pattern	Exemplary student explanation	%		%*
no choice, no explanation	--	3.9	incorrect	36.4
incorrect choice, no explanation	--	2.6		
incorrect choice, incorrect explanation (various)	„Diethylether löst sich besser (keine OH-Gruppe vorhanden)“ (P0063) [diethylether is better soluble (has no OH-group)]	2.6		
incorrect choice, incorrect explanation (bond breaking)	„Diethylether, da die Esterbindung zwischen den beiden Ethylgruppen gespalten werden kann und so 2 Ethanol-Moleküle entstehen, die aufgrund ihrer kleinen Größe besser löslich sind als das größere Butan-2-ol.“ (P4378) [diethylether, because the ester bond between both ethyl groups can be split and 2 ethanol molecules form which are easier soluble in water due to their smaller size than the larger butan-2-ol]	1.3		
incorrect choice, incorrect explanation (polarity)	„Diethylether löst sich besser in Wasser, da es durch die Sauerstoffbrücke leichter polarisierbar ist“ (P0002) [diethylether is better soluble in water as it can be polarised easier due to the oxygen bridge]	5.2		
incorrect choice, incorrect explanation (steric reasons)	„Butan-2-ol: polar OH-Gruppen, aber sterisch gehindert. Diethylether: müsste sich besser in Wasser lösen → geringere sterische Hinderung“ (P1201) [butan-2-ol: polar OH-groups, but sterically hindered; diethylether: should be better soluble in water → lesser sterical hindrance]	2.6		
correct choice, no explanation	--	1.3		
correct choice, incorrect explanation (various)	„Butan-2-ol lässt sich besser in Wasser lösen, da die OH-Gruppe nur 1-fach gebunden ist.“ (P0549) [butan-2-ol is better soluble in water because the OH-group is only bonded once]	2.6		
correct choice, incorrect explanation (bond breaking)	„Butan-2-ol, da dort besser ein H-Ion an der Hydroxidgruppe angreifen kann, womit es dann zu einer Eliminierungsreaktion kommen kann.“ (P1265) [butan-2-ol, as a H-ion can attack the hydroxyl group better and therefore leads to an elimination reaction]	9.1		

correct choice, incorrect explanation (steric reasons)	„Das Butan-2-ol, da die Hydroxygruppe besser zugänglich ist als das Sauerstoff Atom des Ethers und damit besser Wasserstoffbrückenbindungen ausbilden kann [...]“ (P0102) [butan-2-ol, for the hydroxyl group is more easily accessible than the oxygen atom of the ether and can therefore form hydrogen bonds]	5.2		
correct choice, explanation using rule-based knowledge	„Diethylether lässt sich schlechter lösen → Ether. Butan-2-ol dagegen besser → Alkohol“ (P0562) [diethylether is less soluble in water → ether. Butan-2-ol is better soluble → alcohol]	14.3	partially correct	61.1
correct choice, partially correct explanation (having an OH group or forming hydrogen bonds)	„Butan-2-ol löst sich besser, da OH-Gruppe vorhanden, Wasserstoffbrücken können sich bilden“ (P0151) [butan-2-ol is better soluble, as it has an OH-group, hydrogen bonds can form]	35.1		
correct choice, partially correct explanation (polarity)	„Butan-2-ol, da hier die Polarität des Moleküls höher ist als bei Diethylether“ (P0194) [butan-2-ol, as the polarity of the molecule is higher than that of diethylether]	11.7		
correct choice, correct explanation (alcohol can form one hydrogen bond per molecule more than ether)	„Butan-2-ol ließe sich besser lösen, da zusätzlich ein positiv polarisiertes H-Atom vorhanden ist → bessere Wasserstoffbrückenbildung“ (P0090) [butan-2-ol would be more soluble as it has an additional positively polarized H-atom → better formation of hydrogen bonds]	2.6	correct	2.6

* The deviance from 100% is a result due to rounding to the first decimal place.

When looking at the categories found in students' answers to this solubility question, themes quite similar to those found in E1 (boiling points of different alcohols) can be observed: 10.4% of all students (range of chemistry semester: 4 – 11, Ø5.3) use some sort of bond breaking argument in their explanation. Again, the key term 'polarity' is also found in the solubility argumentation (E5), both within the incorrect category (5.2%) when arguing for the ether and within the partially correct category (11.7%) when explaining the better solubility of the alcohol. While for this item, using 'polarity' does make some sense as the relevant fact is a positively polarized hydrogen atom, the reasoning of most students does not go that far but rather stops at the mere mentioning of the term or at one of the molecules being described as more polar. Students also compare the polar makeup of the molecule in question to that of water, thus arguing along the lines "like dissolves like", i.e. rule-based knowledge (14.3%; range of these students' chemistry semester: 2 – 12, Ø7.8). An additional key term that is wrongfully used in explaining this item is 'steric hindrance' (total of 7.8%, students' range of chemistry semester: 2 – 13, Ø6.3).

As both compounds provided have an identical number of carbon atoms in their chains, and show a similar linear makeup with the oxygen atom in a non-terminal position, steric hindrance is not the decisive factor in this explanation. Interestingly, 35.1% of all students justify their selecting the alcohol as the compound that is more easily soluble in water with the mere existence of the hydroxyl group in the alcohol and their consequential argumentation that only due to this, hydrogen bonds can form with water molecules. Students providing this kind of answer pattern are on average in their 5th chemistry semester (range: 2 – 10, Ø5.0).

A third item dealing with hydrogen bonds is item E17, in which students were asked to select – based on their own drawing of chemical structures – whether ether or terminal alcohol of identical carbon chain length would form hydrogen bonds among its own kind more easily. In order to answer this question, students needed to realize that they had to first draw at least two of each molecule and then draw (correct) hydrogen bonds between the alcohol molecules. As the ether molecules do not contain polarized hydrogen atoms, no hydrogen bonds can be formed between ether molecules. In this item, both the drawing of the chemical structures and the written selection were assessed. As the task was not explicitly asking for an explanation, a mere correct selection together with a correct chemical drawing would qualify as a correct answer. However, whenever students did provide some reasoning for their selection, this reasoning was included in the assessment.

When comparing answers (cf. Table 19) with the other items on hydrogen bonds (i.e. E1 and E5), the most interesting answers are the ones using key terms such as polarity and steric hindrance. 7.6% of all answers (range of these students' chemistry semester: 4 – 12, Ø7.5) base their selection on the general aspect of polarity, without going into further detail what they mean by it or how it ultimately leads to the positively polarized hydrogen atom in alcohols which is one prerequisite for hydrogen bond formation and the decisive factor when comparing an alcohol's ability to form hydrogen bonds to that of an ether.

Table 19: Answer patterns and coding scheme for item E17, verbal answers

Answer pattern	Exemplary student explanation	%		%*
no selection/ distinguishable expl.	--	6.3	incorrect	49.4
incorrect selection, no expl.	--	2.5		
incorrect choice, incorrect explanation (free pairs of electrons)	„Ether bilden leichter H-Brücken aus aufgrund von 2 freien Elektronenpaaren am Sauerstoff“ (P4361) [ethers form hydrogen bonds more easily because of the two free pairs of electrons at the oxygen atom]	2.5		
incorrect choice, incorrect explanation (lecture knowledge)	„Ether bilden leichter Wasserstoffbrückenbindungen aus, da es schon in AC behandelt wurde“ (P0200) [ether form hydrogen bonds more easily because we already covered this in inorganic chemistry]	1.3		
incorrect choice, incorrect explanation (polarity)	„Ether sind weniger polar, da ihre Gesamtsymmetrie immer noch höher ist als bei endständigen Alkoholen → bessere Ladungsverteilung“ (P0194) [ethers are less polar because their overall symmetry is still higher than in terminal alcohols → better distribution of charge]	2.5		
incorrect choice, incorrect explanation (steric reasons)	„Ether bilden leichter Wasserstoffbrückenbindungen aus, da sie sich räumlich besser anordnen können und dadurch häufiger H- Brücken“ (P0610) [ether form hydrogen bonds more easily because they can align better and therefore form hydrogen bonds]	3.8		
correct choice, incorrect explanation (free electrons)	„endständige Alkohole, da mehr freie Elektronen“ (P4342) [terminal alcohols because more free electrons]	1.3		
correct choice, incorrect explanation (various)	„Alkohol bildet eher H-Brücken aus, da das H-Teilchen eine Teilladung wie beim Wasserstoff besitzt“ (P0597) [alcohol will form hydrogen bonds more easily because the hydrogen particle has a partial charge similar to hydrogen]	6.4		
correct choice, incorrect explanation (steric reasons)	„Da der Ether durch die beiden Reste sterisch gehindert ist, bilden Alkohole eher H-Brücken aus“ (P4364) [As ethers are sterically hindered by both rests, alcohols form hydrogen bonds more easily]	22.8		
correct choice, partially correct explanation (polarity)	„Alkohole haben höheren Dipolcharakter → Ausbildung H- Brücken möglich“ (P1230) [alcohols have larger dipole moments → formation of hydrogen bonds possible]	1.3	partially correct	10.2
correct choice, partially correct explanation (polarity + steric reasons)	„endständige Alkohole sind stärker polarisiert. Zudem ist die sterische Hinderung geringer. Also: Alkohole“ (P0970) [terminal alcohols are polarized more strongly; in addition the sterical hindrance is less, therefore: alcohol]	3.8		
correct choice, partially correct explanation (polarization/partial charges)	„leichter endständige Alkohole, weil die negative Partialladung am O größer ist.“ (P0565) [easier for terminal alcohols, as the negative partial charge at the oxygen atom is higher]	5.1		
correct choice, no explanation**	--	21.5	cor	40.5

correct choice, correct explanation	„Alkohole können leichter WBB ausbilden, da an dem Sauerstoff ein polarisiertes H-Atom gebunden ist. Dies fehlt dem Ether komplett → keine WBB“ (P4400) [alcohols can form hydrogen bonds more easily because the oxygen atom is attached to a polarized hydrogen atom. This is completely missing in the ether → no hydrogen bonds]	19.0		
-------------------------------------	---	------	--	--

* The deviance from 100% is a result due to rounding to the first decimal place.

**As the item did not demand an explanation, this category is assessed as "partially correct".

Therefore, simply arguing that one of the two molecules is more polar than the other does not suffice at this point. A similarly incorrect argument for this item is steric hindrance. A total of 30.4% of students' answers refer to this key term as being the only or main reason for selecting either ether or alcohol. On average, students that answer this way are in their 5th chemistry semester (range: 2 – 13, Ø5.2). As in item E5, steric hindrance is a key term often referred to in organic chemistry but is not the relevant factor here. Even more interestingly, 3.8% of students (range of these students' chemistry semester: 4 – 13, Ø7.0) refer to the number of free electrons as a selection criterion, not knowing or realizing that ethers and alcohols (at least those with one hydroxyl group) are structural isomers, thus having an identical number of free electron pairs.

When looking at the chemical structures drawn to illustrate the selection of ether or alcohol in this item, the following numbers can be obtained (Table 20):

Table 20: Answer patterns and coding scheme for item E17, drawing of chemical structures

Category	%
No drawing or incorrect structural formulae	21.5
Correct structural formulae for one ether and alcohol molecule each	49.4
Correct structural formulae for more than one ether and alcohol molecule each, but with no or incorrect depiction of a hydrogen bond	16.5
Correct structural formulae for more than one ether and alcohol molecule each, and with correct depiction of a hydrogen bond between alcohol molecules	12.7

From these results one can see that about one fifth of the students (21.5%) have not even attempted to draw an ether and alcohol molecule with the same chain length. In addition, almost half of the students (49.4%) draw only a single ether and alcohol molecule each. Even more interestingly is the third category, which contains

correctly drawn structural formulae for more than one ether and alcohol molecule each but without any or with an incorrect depiction of a hydrogen bond. Apart from not having drawn any hydrogen bond at all (3.8%), the vast majority (10.1%) of drawings in this category correctly depicted the hydrogen bond between alcohol molecules but incorrectly used the hydrogen atoms of the methyl groups in ethers as allegedly polarized hydrogen atoms in hydrogen bonds (cf. Figure 32, P0778):

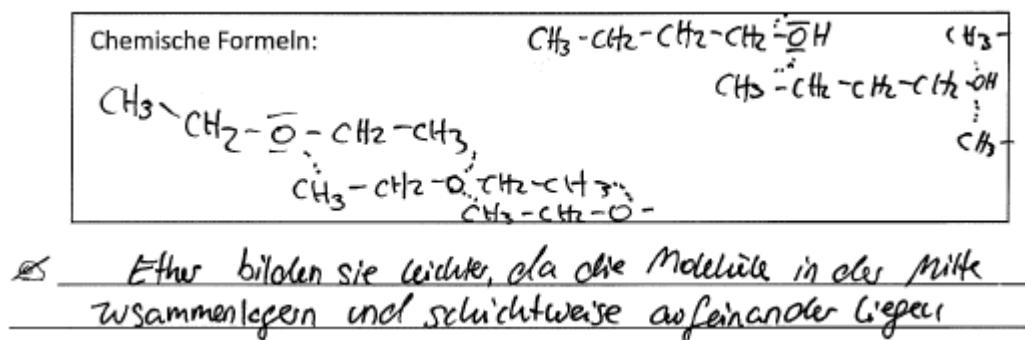


Figure 32: Drawing and explanation³¹ of person P0778, illustrating incorrect hydrogen bonds for ether

All structural drawings falling into this last group coincide with an argumentation referring to steric reasons or steric hindrance.

When looking at the items explicitly asking students to define and depict a hydrogen bond (i.e. R11_1 definition), the following categories arise for the written definition (cf. Table 21): only 3.3% of student answers can be considered complete as they explain that a hydrogen bond is an intermolecular force between a partially positively charged hydrogen atom and the free pair of electrons of a partially negatively charged oxygen, nitrogen or fluorine atom of a neighboring molecule³².

³¹ English translation of explanation: Ethers form them easier because the molecules can align in the middle and then form layers.

³² The definition containing two separate parts of one larger molecule so that intermolecular forces would result was, of course, also considered correct as long as it was clear that the partially negatively charged atom was not directly covalently bonded to the partially positively charged hydrogen atom involved in the hydrogen bond.

Table 21: Answer patterns and coding scheme for item R11_1, definition of hydrogen bond [HB=hydrogen bond]

Answer pattern	Exemplary student explanation	%		%*
no definition given	--	22.2	incorrect	85.5
incorrect definition (incl. HB between two H-atoms)	„Imaginäre Bindungen, die z.B. bei Wasser vorhanden sind. Sind geringer als v.d.W-Kräfte“ (P0198) [imaginary bonds that are present in water, for example; they are smaller than van der Waals forces]	23.5		
insufficient definition (only free pairs of electrons)	„Wasserstoffbrückenbindungen sind Wechselwirkungen zwischen H-Atomen und freien Elektronenpaaren“ (P0604) [HB bonds are forces between a H-atom and free pairs of electrons]	2.6		
insufficient definition (only partial charges)	„Im Wassermolekül ist das Wasserstoffatom partiell positiv und das Sauerstoffatom partiell negativ. Es kommt zu Wasserstoffbrückenbindungen.“ (P0568) [In water molecules the hydrogen atom has a partially positive and the oxygen atom has a partially negative charge. HB result.]	5.9		
insufficient definition (only intermolecular forces)	„Eine Wasserstoffbrückenbindung beruht auf statischen Wechselwirkungen zwischen H-Atomen und O-Atomen innerhalb eines Moleküls oder auch zwischen Molekülen.“ (P0589) [a HB is based on static forces between H and O-atoms within a molecule or between molecules]	7.8		
insufficient definition (only describing prerequisites, consequences or both)	„Wasser ist ein Dipol. Die entgegen gesetzt geladenen Teilchen ziehen sich gegenseitig an und stabilisieren sich so.“ (P0579) [Water is a dipole. The opposing charged parts attract each other and stabilize that way]	14.4		
insufficient definition (only describing prerequisites, consequences or both + intermolecular force)	„Intermolekulare Kräfte, die durch Elektronegativitätsunterschiede beeinflusst werden. Zwischen H, O und N“ (P0511) [Intermolecular forces that are influenced by differences in electronegativity. Between H, O and N]	5.2		
insufficient definition (only describing prerequisites, consequences or both + partial charges)	„durch die intramolekulare Polarisierung durch höhere EN-Differenzen wird ein "elektronenarmer" Pol von einem "elektronenreicheren" angezogen. Diese Anziehung ist stark, jedoch zu schwach um richtige Bindungen einzugehen. Dadurch entsteht ein Netz von Molekülen, die stark zusammenhängen.“ (P0002) [Through the intramolecular polarization due to higher differences in electronegativities, an electron-poor pole is attracted to an electron-rich pole. This attraction is strong, yet too weak to form a real bond. Through this, a net of molecules results in which the molecules connect strongly.]	3.9		
partially correct definition (intermolecular force + free electron pair)	„Sie sind eine Wechselwirkung zwischen einem freien Elektronenpaar und einem H-Atom zweier benachbarter Moleküle“ (P0610) [they are the forces between a free pair of electrons and a H-atom of two neighboring molecules]	2.0	partially correct	11.1
partially correct definition (partial charges + free electron pair)	„Wasserstoffbrückenbindungen können zwischen stark positiv polarisierten H-Atomen und freien Elektronen an N, O, F (Cl) gebildet werden“ (P0737) [HB can be formed between strongly positively polarized H-atoms and free electrons of N, O, F (Cl)]	3.9		

partially correct definition (intermolecular force + partial charges)	„WBB sind intermolekulare Anziehungskräfte, die zwischen Wasserstoff und einem elektronegativen Pol stattfinden“ (P0592) [HB are intermolecular attraction forces that happen between hydrogen and an electronegative pole]	5.2		
correct definition, unusual terminology	„Ausbildung von schwachen Bindungen zwischen den partiell positiven und partiell negativen Regionen der Wassermoleküle“ (P0232) [formation of weak bonds between partially positive and partially negative regions of water molecules]	2.0	correct	3.3
correct definition, correct terminology	„H-Brücken entstehen zwischen einem H-Atom das an ein elektronegatives Atom (N,O,F) gebunden ist und einem freien Elektronenpaar eines anderen elektronegativen Atoms.“ (P0966) [HB appear between a hydrogen atom that is bonded to an electronegative atom (N, O, F) and a free pair of electrons of a different electronegative atom]	1.3		

* The deviance from 100% is a result due to rounding to the first decimal place.

11.1% of students can be considered having given a partially correct answer as they have left out only one of the three criteria necessary (intermolecular or intramolecular force, partially charged atoms, free electron pair). A total of 85.5% of students do not even partially define a hydrogen bond, when not writing anything at all (22.2%), or providing an insufficient definition by i) only stating one of the three criteria necessary (16.3%), ii) merely describing prerequisites for hydrogen bond formation such as differences in electronegativity, consequences of hydrogen bonds such as increased stability of the compound or a combination of both (14.4%) or iii) describing the prerequisites and/or consequences along with providing one of the above-mentioned criteria (9.1%). Additionally, 23.5% provide a strictly incorrect definition: About a third of these incorrect answers – this being the largest subgroup in this category – wrongly describe hydrogen bonds as forces between two hydrogen atoms, such as: *„intermolekulare Bindungen zwischen H-Atomen polarer Moleküle [intermolecular forces between H-atoms of polar molecules]“* (P0194). Students giving answers of this particular type are on average in their 4th chemistry semester (range of chemistry semester: 2 – 8, Ø4.4).

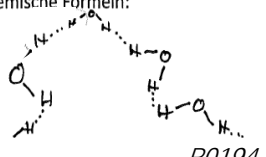
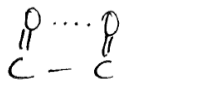
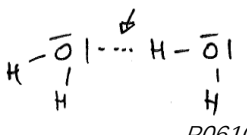
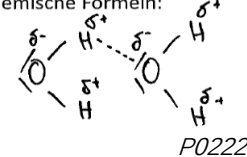
Table 22: Answer patterns and coding scheme for item R11_1, drawing of hydrogen bond

Answer pattern	%		%*
no drawing	19.0	incorrect	53.2
incorrect drawing (incl. hydrogen bond between two H-atoms; wrong or incomplete molecules)	27.5		
insufficient drawing (correct molecule that can do hydrogen bond, no free pair of electrons depicted)	3.3		
insufficient drawing (correct molecules that can form hydrogen bond and correct orientation, with free pair of electrons depicted)	.7		
insufficient drawing (correct molecules that can form hydrogen bond and correct orientation + partial charges, no free pair of electrons depicted)	.7		
insufficient drawing (correct molecules that can form hydrogen bond and correct orientation + partial charges, with free pair of electrons depicted)	2.0		
partially correct drawing (correct molecules that can form hydrogen bond and correct orientation + depiction of hydrogen bond, no free pair of electrons depicted)	10.5	partially correct	34.7
partially correct drawing (correct molecules that can form hydrogen bond and correct orientation + depiction of hydrogen bond, with free pair of electrons depicted)	20.9		
partially correct drawing (correct molecules that can form hydrogen bond and correct orientation + partial charges + depiction of hydrogen bond, no free pair of electrons depicted)	3.3		
correct drawing (correct molecules that can form hydrogen bond and correct orientation + partial charges + depiction of hydrogen bond, with free pair of electrons depicted)	12.4	correct	12.4

* The deviance from 100% is a result due to rounding to the first decimal place.

Considering the drawing of a hydrogen bond (i.e. R11_1 drawing), similar results can be obtained (cf. Table 22): 19.0% of students have not even attempted to draw a hydrogen bond with self-chosen molecules, 27.5% provide an incorrect drawing of either the molecules involved or the hydrogen bond itself (cf. P0194 in Table 23), and 6.7% give an insufficient depiction of a hydrogen bond either just providing molecules oriented toward other molecules in a certain way and possibly including partial charges but not indicating the existence of the actual hydrogen bond itself. 34.7% of students provide a partially correct drawing, i.e. explicitly depicting the actual hydrogen bond with or without the free pair of electrons yet without partial charges or drawing the hydrogen bond with partial charges for the relevant atoms but not depicting the free pair of electrons (cf. P0610 in Table 23). Only 12.4% of all students provide a complete and correct drawing of a hydrogen bond, including all three above-mentioned aspects (cf. P0222 in Table 23).

Table 23: Exemplary drawings of a hydrogen bond

Incorrect depictions		Partially correct depiction	Completely correct depiction
Chemische Formeln:  <i>P0194</i>	Chemische Formeln:  <i>P0063</i>	Chemische Formeln:  <i>P0610</i>	Chemische Formeln:  <i>P0222</i>

Again, the most interesting category is that of incorrect drawings. 12.1% of students who have provided a drawing, illustrate a hydrogen bond as an intermolecular force between two hydrogen atoms of neighboring molecules (cf. P0194 in Table 23) or two oxygen atoms of neighboring or the same molecule (cf. P0063 in Table 23). These students are on average in their 5th chemistry semester (range: 2 – 10, \bar{O} 4.7). An additional 3.2% of students, having attempted to draw a hydrogen bond, use hydrogen atoms of methyl groups to illustrate the partially positive hydrogen atom in a hydrogen bond.

5.5.3 CK – resulting instrument

In order to analyze the instrument as a whole, each test booklet was looked at individually at first.

Table 24: Discriminatory power per CK item, booklet B

item	discriminatory power
R27_3u4	-0.088
E32	-0.073
R24_2	0.016
E35	0.018
R3	0.019
E31	0.120
E26	0.152
R24_1	0.185
E27_3	0.197
E27_1	0.209
R2	0.217
R9	0.219
R21_1	0.219
R32	0.227
E7_2	0.251
E7_1	0.258
E4	0.271
E16	0.275
E22	0.279
E24	0.285
R18	0.299
E5	0.303
R16	0.344
E12	0.351
R10	0.373
R14	0.373
R13	0.378
R21_3	0.379
R22	0.395
E27_2	0.424
E1	0.428
R11_2	0.435
R21_2	0.435
R24_3	0.436
E18	0.444
R23	0.446
E17	0.459
R11_1	0.499
E30	0.518
R21_4	0.540
R27_1u2	0.547
R28	0.564

It was of interest to find out if an item was suitable to reveal differences in pre-service teachers' abilities. Theoretically, items with an item-total-correlation, i.e. the discriminatory power, of less than .2 should have been eliminated (for test booklet B these items would be E26, E27_3, E31, E32, E35, R3, R24_1, R24_2, R27_3u4, cf. Table 24 red items). According to the process described in the chapter on item difficulty (chapter 5.5.2), content-related considerations led to the retention of some items even though they should have been excluded from a test theory point of view.

The statistical values obtained for the three test booklets can be seen in Table 25. For each item set, the Cronbach's α value, as a measure of the internal consistency within a student's answers to the items of a test, shows acceptable ($\alpha=.71$) to good ($\alpha=.85$ and $\alpha=.88$) internal consistency.

Even though all three item sets emerged showing satisfactory statistical values, only item sets A and B were considered for further analysis as both subsets had a considerable number of items remaining (32 and 30 items, respectively) while many of the items in test booklet C had to be discarded. Thus two reliable sets of items resulted: those in items A and B.

Table 25: Statistical values of the three CK scales

	Item Set A	Item Set B	Item Set C
Number of items	32	30	13
Number of students answering item set	79	73	68
Reliability (Cronbach's α)	.88	.85	.71
Discriminatory power; average and span	.40 (.24 - .58)	.37 (.21 - .55)	.33 (.21 - .58)
Item difficulty; average and span	.52 (.22 - .78)	.43 (.08 - .79)	.48 (.24 - .76)

Due to the mentioned overlap of 25 items, both reliable item sets A and B were considered jointly for the following correlations and the chapter on factors influencing item difficulty. The resulting structure-property-instrument for content knowledge correlates significantly with the cognitive ability test ($r = .30, p < .05$), with the number of the students' university semester in chemistry ($r = .30, p < .05$) and with the students' overall grade on their school leaving certificate ($r = -.30, p < .05$; the negative correlation resulting from an inverse scaling of the grades with 1 being the best and 4 being the worst passing grade). All three correlations have a medium effect according to COHEN³³ (1992). Furthermore, this instrument on SPR also correlates significantly with the test instrument developed for assessing the students' understanding of chemical representations (cf. TASKIN, 2014) with $r = .59, p < .005$. According to COHEN (1992), this correlation has a large effect.

5.5.4 CK – factors influencing item difficulty

The item sets A and B were considered jointly not only for correlation analysis with other measures but also for an analysis regarding possible effects of the parameters described in the item design.

³³ COHEN (1992) presents the following cutoff values for effects sizes for r : .10/.30/.50 (small/medium/large) and for ANOVA: .10/.25/.40 (small/medium/large)

With regard to the levels of complexity, an ANOVA showed significant differences between the three levels ($F(2, 79) = 3.486$; $p < .05$; $\omega^{34} = .24$) and a post hoc contrast according to Scheffé indicated that the items at the level of causal relations were significantly more difficult than items on a factual level ($p < .05$).

Concerning the number of perspectives related to the chemical tetrahedron, an ANOVA-test ($F(2, 79) = 8.861$; $p < .05$; $\omega = .40$) was conducted and followed by a post hoc contrast according to Scheffé which demonstrated that items involving a combination of either two or three perspectives were significantly more difficult to solve correctly than those addressing only one perspective ($p < .005$ and $p < .05$, respectively).

With regard to the educational allocation, the two groups of items were compared and a t -test revealed that items with university relevance were significantly more difficult than school-relevant ones ($t(80) = -4.52$, $p < .05$, $r = .45$).

Regarding the item format, an ANOVA also estimated significant differences between the formats ($F(2, 79) = 7.039$; $p < .05$; $\omega = .35$) with a post hoc contrast according to Scheffé to show that open items were more difficult to solve than closed ones ($p < .005$).

Looking at effect sizes of these factors, complexity has a small effect on test scores, while number of perspectives, educational allocation, and item format each have a medium effect.

As the test was not designed to contain enough items to systematically fill all possible cells in a matrix of the above mentioned factors, it cannot be stated at this point whether there is a certain combination of these factors that makes items easy or difficult to answer. When looking at all items focusing on the content aspect of hydrogen bonding³⁵, for example, one can see that the items do vary in their difficulty, as can be seen in Figure 33.

³⁴ ω being "an unbiased estimate of r " (Field, 2009, p.390) to be used when calculating effect sizes for an ANOVA

³⁵ From their position in the test instrument it was not apparent for students that these items belonged together in any way based on their joint topic of hydrogen bonding.

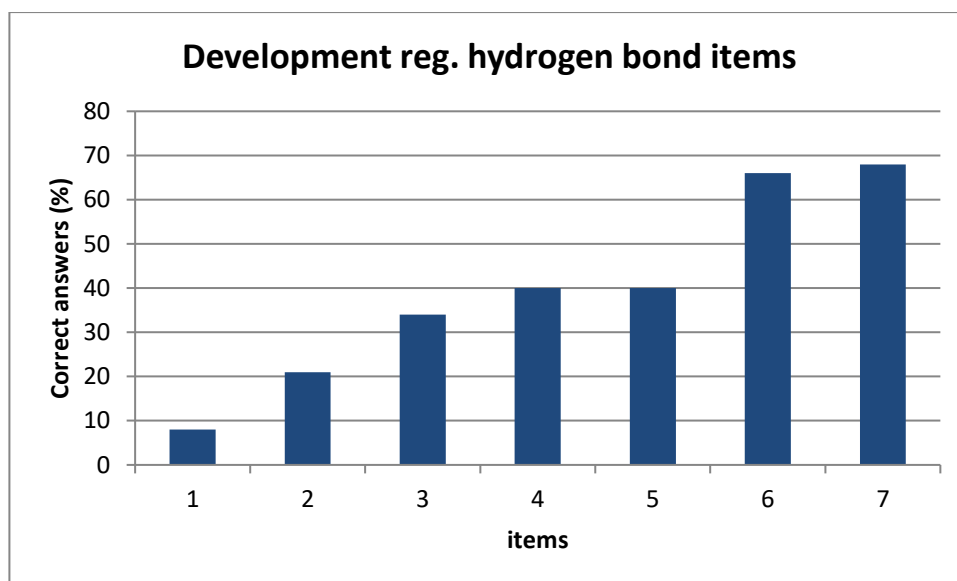


Figure 33: Variety of items addressing hydrogen bonds

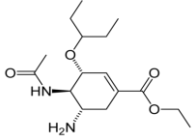
While qualitative results of the answers to some of the items in Figure 33 have been reported in chapter 5.5.2, a closer look at possible aspects in the items themselves might provide first content clues that might have led to the differences in difficulty. Table 26 shows these seven items in detail: Some of these items either contain or ask for an example of the generic term “compound” (items no. 2 and 7), while other items provide the names of a specific compound such as “diethylether” or “hydrazine” (items no. 6 and 1, respectively). Some items contain structural formulae (items no. 2 and 6), while others ask the students to draw structural formulae (items no. 3 and 5). One item contains a sum formula of two specific compounds (item no. 1). While some items include the cue that hydrogen bonds are the aspect to be considered (items no. 3, 5 and 7), the other items do not address this topic explicitly but rather require students to use the concept in their answers (items no. 1, 2, 4 and 6).

Neither form of specificity (general vs. specific compound) nor inclusion of or asking for chemical formulae nor the explicitness of the hydrogen bonding concept seem to have a sole effect on making an item more or less difficult.

The categorization regarding the items' complexity, however, suggests a potential trend as the items with a lower level of complexity seem to be generally easier than those with a higher level of complexity (item no. 6 needs to be neglected

at this point). A similar trend can be seen when looking at the items' number of perspectives: addressing fewer perspectives does seem to make items easier than addressing more perspectives. Students' familiarity with a compound and thus its categorization into educational allocation also seems to have an effect since water, alcohols and ethers are more likely to be covered in detail in school while nitrogen instead of oxygen forming hydrogen bonds is less likely to be covered at length in school and hydrazine is likely to only be addressed at university level.

Table 26: Characterization of items regarding hydrogen bonds

No.	Item wording	Complexity	Perspectives	Educational allocation										
1 [E22]	Pure hydrazine has a boiling point of 113°C, the corresponding phosphorous substance (P ₂ H ₄) boils (under decomposition) at 56°C. Explain this surprising decrease in boiling points from the lighter N ₂ H ₄ to the heavier P ₂ H ₄ .	causal relations	three (macro + submicro + symbolic)	university knowledge										
2 [R9]	Please explain if the compound with the given structure would be able to get into the groundwater in a noteworthy amount if applied to a lawn. 	causal relations	three (macro + submicro + symbolic)	university knowledge										
3 [R11_1]	By drawing appropriate chemical structures of hydrogen bonds, please explain in a chemically correct way what a "hydrogen bond" is.	facts	two (submicro + symbolic)	school knowledge										
4 [E4]	Various physical processes will lead to a change in the volume of water. Please indicate which one of the following physical processes states the corresponding change in the volume of water correctly: <table border="0" style="width: 100%; border-collapse: collapse;"> <tr> <td style="text-align: center; width: 50%;"><u>physical process</u></td> <td style="text-align: center; width: 50%;"><u>change in the volume of water</u></td> </tr> <tr> <td>○ warming from 0°C to 4°C</td> <td>increase in volume</td> </tr> <tr> <td>○ warming from 4°C to 12°C</td> <td>decrease in volume</td> </tr> <tr> <td>○ freezing of water</td> <td>increase in volume</td> </tr> <tr> <td>○ melting of ice</td> <td>increase in volume</td> </tr> </table>	<u>physical process</u>	<u>change in the volume of water</u>	○ warming from 0°C to 4°C	increase in volume	○ warming from 4°C to 12°C	decrease in volume	○ freezing of water	increase in volume	○ melting of ice	increase in volume	description of processes/procedures	one (macro)	school knowledge
<u>physical process</u>	<u>change in the volume of water</u>													
○ warming from 0°C to 4°C	increase in volume													
○ warming from 4°C to 12°C	decrease in volume													
○ freezing of water	increase in volume													
○ melting of ice	increase in volume													
5 [E17]	Decide if ether or alcohol with identical number of carbon atoms will form hydrogen bonds amongst its own kind more readily. Use chemical structures to demonstrate your idea.	description of processes/procedures	two (submicro + symbolic)	school knowledge										
6 [E5]	Explain which of these substances is better soluble in water (at a pH-value of 7). <table border="0" style="width: 100%; border-collapse: collapse;"> <tr> <td style="text-align: center; width: 50%;"> $\begin{array}{cccc} \text{H} & \text{H} & \text{H} & \text{H} \\ & & & \\ \text{H}-\text{C}-\text{C}-\text{O}-\text{C}-\text{C}-\text{H} \\ & & & \\ \text{H} & \text{H} & \text{H} & \text{H} \end{array}$ Diethylether </td> <td style="text-align: center; width: 50%;"> $\begin{array}{cccc} \text{H} & \text{H} & \text{H} & \text{H} \\ & & & \\ \text{H}-\text{C}-\text{C}-\text{C}-\text{C}-\text{H} \\ & & & \\ \text{H} & \text{OH} & \text{H} & \text{H} \end{array}$ Butan-2-ol </td> </tr> </table>	$\begin{array}{cccc} \text{H} & \text{H} & \text{H} & \text{H} \\ & & & \\ \text{H}-\text{C}-\text{C}-\text{O}-\text{C}-\text{C}-\text{H} \\ & & & \\ \text{H} & \text{H} & \text{H} & \text{H} \end{array}$ Diethylether	$\begin{array}{cccc} \text{H} & \text{H} & \text{H} & \text{H} \\ & & & \\ \text{H}-\text{C}-\text{C}-\text{C}-\text{C}-\text{H} \\ & & & \\ \text{H} & \text{OH} & \text{H} & \text{H} \end{array}$ Butan-2-ol	causal relations	three (macro + submicro + symbolic)	school knowledge								
$\begin{array}{cccc} \text{H} & \text{H} & \text{H} & \text{H} \\ & & & \\ \text{H}-\text{C}-\text{C}-\text{O}-\text{C}-\text{C}-\text{H} \\ & & & \\ \text{H} & \text{H} & \text{H} & \text{H} \end{array}$ Diethylether	$\begin{array}{cccc} \text{H} & \text{H} & \text{H} & \text{H} \\ & & & \\ \text{H}-\text{C}-\text{C}-\text{C}-\text{C}-\text{H} \\ & & & \\ \text{H} & \text{OH} & \text{H} & \text{H} \end{array}$ Butan-2-ol													
7 [R11_2]	Please name a compound from nature where hydrogen bonds play an important role.	facts	one (submicro)	school knowledge										

difficult



easy

5.5.5. PCK – individual item analysis

A total of 42 pedagogical content knowledge items regarding SPR went into study I. As described in chapter 5.2, test booklet A did not include any PCK items for structure-property-relations. As with the CK items, the first step is to look at how difficult or easy the items were for the students. In order to do this, Figures 34 to 35 present the statistical difficulty of each item in the study, grouped into the respective test booklets. Again, a value close to 1 means that the item is easy (has a high likelihood of being solved correctly) and a value close to 0 points out that the item is difficult.

From a test theory point of view, all items with difficulty values below .2 are discarded as they are too difficult and items with a difficulty value above .8 are excluded from further analysis as well as they are too easy for the students. For test booklet B³⁶, this would mean that item I14 should be eliminated. For test booklet C, items I10, I11, and S11 would be excluded due to this criterion. The majority of items fulfill this criterion and therefore remain in the item sets at this point.

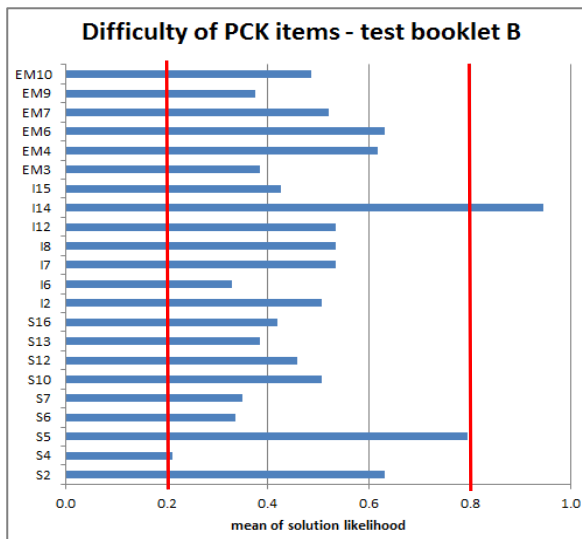


Figure 34: Difficulty of PCK items in test booklet B

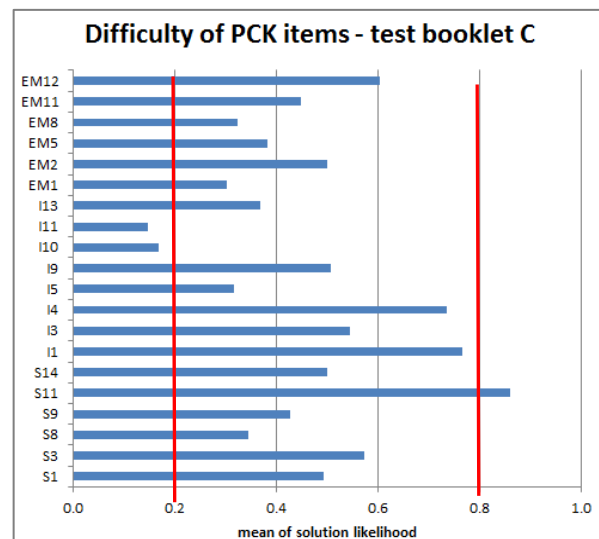


Figure 35: Difficulty of PCK items in test booklet C

While Figures 34 to 35 present the statistical difficulty for each item, the question on how students score in the resulting instrument will be looked at in chapter 5.5.7.

³⁶ No PCK items on SPR had been included in test booklet A.

QUALITATIVE ANALYSIS OF SELECTED PCK ITEMS

Similar to the analysis in the CK section, pre-service teachers' answers to a small number of PCK items shall be presented at this point. The reason for the PCK item selection is twofold: i) one item from each sub-dimension (i.e. student cognition, instructional strategies, or experiments and models) shall be presented and ii) only those items shall be presented in detail here that have already been publicized before, as the test instrument is still being used.

Pre-service teachers' answers showed a variety of answer patterns, which were classified and then categorized as correct, partially correct and incorrect. As the differentiation into correct and incorrect is not as clear as with CK items, the answers to PCK items were assessed according to their appropriateness and reasonability. In all tables providing exemplary student answers, pre-service teachers' original German statements are provided along with a translation that remains as close to the original as possible.

In Item I6, for example, pre-service teachers are asked to describe an experimental setup to show a student that his/her thinking is incorrect. The example given is that the student tells the teacher that diamond is everlasting and because of its properties cannot possibly be made of only carbon just as graphite is. The item requires the participants in the study to provide a hypothetical teaching action that specifically addresses a student's misconception. In order to answer this item with full credit, pre-service teachers need to state two things: i) describe an experiment showing that diamond is not everlasting, such as burning it, and ii) refer to a method of analytical detection of a substance that allows deductions regarding the composition of diamond. A second possibility for a full-credit score is describing an experiment showing that graphite can be directly turned into diamond. A mere reference to an experiment showing that diamond is not everlasting without mentioning the method of analytical detection of carbon dioxide leads only to a partial credit coding for this item: Simply stating the reaction, i.e. tapping into one's content knowledge, without the detection and subsequent deduction part is not

sufficient as the student's confusion, as presented in the item text, is that graphite and diamond should consist of the same atoms. Regarding pre-service teachers' answers to this item, it is irrelevant whether participants describe an experiment that could be performed in an ideal setting or whether they state that due to a realistic assessment of a financial situation, only a video presenting that experiment could be shown. Table 27 shows pre-service teachers' answers and the corresponding classification.

27.4% of pre-service teachers answer this item correctly either by providing an experiment to have diamond react with oxygen and a subsequent detection method of carbon dioxide or by referring to a synthesis of diamond from graphite. 15.1% of answers fall into the partially-correct category as they only provide an experiment in which diamond is burned. Apart from 38.4% of pre-service teachers not having answered this item at all, 19% of participants' answers are incorrect. This is either due to content errors such as the notion that diamond cannot be changed by any chemical process or can be changed by processes not actually able to alter diamond, or due to a lack of understanding the underlying problem.

Item S4 asks pre-service teachers to provide three examples in which incorrect student ideas can be traced back to the direct transfer of observable properties of a substance to the particle level. For an appropriate answer, it was necessary to name a macroscopic property and state a corresponding reference to the submicroscopic level which this property could be wrongfully assigned to. Naming only a macroscopic property, only a reference to the submicroscopic level, or a non-corresponding pair, resulted in an incorrect answer. Table 28 shows some examples of coding of appropriate and inappropriate answers for this item.

Table 27: Answer patterns and coding scheme for item I6

Answer pattern	Exemplary student answer	%		%*
no answer	--	38.4		
not situation-specific	„Schwefel kann erhitzt und auf eine Wasseroberfläche gegossen werden. Gleicher Stoff → unterschiedliches Aussehen und Struktur“ (P0508) [Sulfur can be heated and poured onto water surface. Same substance → different looks and structure]	2.7	incorrect	57.4
incorrect concept	„Graphit geht kaputt → Diamant nicht“ (P0593) [graphite can be destroyed → diamond cannot]	6.8		
incorrect experimental conditions	„Diamant mit Säure lösen“ (P0414) [Dissolve diamond in an acid]	6.8		
no experiment, simply paraphrasing	„Diamant ist optisch ganz anders als Graphit. Die Vorstellung dort sind die gleichen Teilchen fällt schwer. Schlussfolgerung der Schüler: Diamant reagiert anders als Graphit“ (P0301) [Diamond is very different from graphite visually. The conception that there are the same particles is difficult. Students conclude: Diamond reacts differently from graphite.]	2.7		
diamond not everlasting, can be burned	„Wir verbrennen einen (sehr kleinen) Diamanten unter Sauerstoffatmosphäre und betrachten das Ergebnis.“ (P1201) [We burn a (very small) diamond under oxygen conditions and examine the result.]	15.1	partially correct	15.1
diamond not everlasting and carbon dioxide can be detected after reaction	„Verbrennung von Diamant → CO ₂ entsteht, Gasnachweis (Kalkwasser)“ (P0565) [burning of diamond → CO ₂ forms, detection of gas (limewater)]	23.3	correct	27.4
diamond can be made from graphite	„Herstellung eines Diamanten aus Graphit (teuer aber effektiv)“ (P0139) [Making a diamond from graphite (expensive but effective)]	4.1		

* The deviance from 100% is a result due to rounding to the first decimal place.

Table 28: Coding of answers for item S4

coding	Exemplary student answers
inappropriate answers	„Molmasse → Umrechnung“ (P0584) [molecular weight → conversion] „Unterschied Schmelzpunkt/Siedepunkt/Gefrierpunkt“ (P0103) [Difference in melting point/boiling point/freezing point] „Spannungsreihe → edel/unedel → mehr oder weniger ‚Wert‘ sein“ (P1201) [electrochemical series → noble/not noble → being ‚worth‘ more or less] „Leitfähigkeit“ (P0550) [conductivity]
appropriate answers	„das gelbe Schwefelatom“ (P0333) [the yellow sulphur atom] „Wasserteilchen beginnen bei 100°C zu sieden“ (P0508) [water particles start boiling at 100°C] „Löslichkeit: Beim Lösen von Feststoffen in Flüssigkeiten lösen sich die Feststoff-Atome auf“ (P0565) [solubility: upon solving a solid in a liquid, the atoms of the solid dissolve] „ein leichter Stoff hat leichte Atome“ (P4400) [a light substance has light atoms]

Through this item, pre-service teachers' extent of anticipated student thinking was evaluated as only the naming of three such corresponding pairs resulted in full

credit while naming one to two of these pairs only resulted in partial credit. For an overview of the final classification regarding this item, see Table 29.

Table 29: Distribution of answers for item S4

answers	%	classification	%*
no answer	30.1	incorrect	67.1
no correct pair	37.0		
one to two appropriate pairs	23.3	partially correct	23.3
three appropriate pairs	9.6	correct	9.6

* The deviance from 100% is a result due to rounding to the first decimal place.

From Table 29 it can be seen that 30.1% of the participants provide no answer to this item while 37.0% of pre-service teachers does not give a single appropriate macroscopic property that students might wrongfully attribute to the submicroscopic structure. While 23.3% give one or two appropriate pairs, with 9.6% only a small fraction of pre-service teachers knows of three appropriate pairs.

An exemplary item from the third sub-dimension of PCK is EM10, which asks pre-service teachers to state three experiments with a clear every-day life context that could be used to introduce students to the topic of acids and bases. As with item S4, this item required participants to demonstrate that their knowledge regarding this aspect is varied – a completely correct categorization necessitated three appropriate experiments. In order to be rated appropriate, each answer had to include three aspects: i) naming an experiment using an acid or a base, ii) referring to an every-day context, and iii) possibility to be reasonably used as introductory experiment. An answer was inappropriate if it consisted of only the name of a compound, an experiment that could not be explained until upper secondary, or an experiment not fitting into the topic of acids and bases, being potentially dangerous or with an unclear intention (examples can be found in Table 30). For an overview of the final classification regarding this item, see Table 31.

Table 30: Coding of answers for item EM10

coding	Exemplary student answers
inappropriate answers	<p>„Essigessenz“ (P1299) [vinegar essence]</p> <p>„Titration von Cola“ (P0568) [titration of cola]</p> <p>„Gehalt an ASS bestimmen“ (P0973) [investigate the content of ASS]</p> <p>„Geschmacksprobe von Zitrone und Seife → Gegenteil von sauer ist nicht süß“ (P0234) [Testing the taste of lemon and soap → the opposite of 'acidic' is not sweet*]</p>
appropriate answers	<p>„Essig löst Eierschalen auf“ (P0508) [vinegar dissolves egg shells]</p> <p>„das Auflösen von Kalk mit Essig (Badezimmer, Wasserkocher)“ (P1293) [dissolving limescale with vinegar (bathroom, kettle)]</p> <p>„Sinn von Zitronensaft bei sehr grätigen Fischen“ (P0194) [purpose of lemon juice for bony fish]</p> <p>„Reinigen von Marmorböden (Säure + Marmor)“ (P1250) [cleaning marble floors (acid + marble)]</p> <p>„Farbänderung von Rotkohl/Blaukraut bei Zugabe von Zitronensaft“ (P4378) [change in color with red cabbage** when adding lemon juice]</p>

* the German term *sauer* can denote both 'acidic' and 'sour' in English

** the German terms *Rotkohl* and *Blaukraut* are local variations of 'red cabbage', already implying the difference in color due to differences in preparation: while in northern Germany the term *Rotkohl* is used and the cabbage is prepared with pieces of apples, in Bavaria the term *Blaukraut* is used and the cabbage is prepared with baking soda

Table 31: Distribution of answers for item EM10

answers	%	classification	%*
no answer	6.8	incorrect	16.4
no appropriate answer	9.6		
one to two appropriate answers	69.9	partially correct	69.9
three appropriate answers	13.7	correct	13.7

* The deviance from 100% is a result due to rounding to the first decimal place.

The distribution of pre-service teachers' answers shows that with 69.9% the vast majority can provide one or two appropriate answers to this item. 16.4% of participants' answers fall in the incorrect category while 13.7% name three appropriate answers thus answering this item correctly.

5.5.6 PCK – resulting instruments

As with the CK items, PCK items with an item-total-correlation of less than .2 and with difficulty values below .2 or above .8 are excluded. The statistical values obtained for the resulting test booklets can be seen in Table 32. For each item set, the Cronbach's α value shows acceptable internal consistency ($\alpha_{\text{test booklet B}} = .79$ and α_{test}

booklet $c=.74$). The value for item difficulty gives the proportion of students who have answered the item correctly. A low item difficulty value implies that many students have answered the item correctly, thus making it an easy item. The average item difficulty is close to .5 for each item set. The discriminatory power is almost identical with a value of .38 and .39. Overall, both item sets show satisfactory reliability.

Table 32: Statistical values of the two PCK scales

	Item Set B	Item Set C
Number of items	18	11
Number of students answering item set	73	68
Reliability (Cronbach's α)	.79	.74
Discriminatory power (average and span)	.38 (.21 - .58)	.39 (.24 - .58)
Item difficulty (average and span)	.48 (.22 - .79)	.47 (.17 - .87)

5.5.7 Performance on tests

Both for CK and PCK, test booklet B shows satisfactory statistical values. Since in study II the surrounding conditions (time frame, administration along with those tests on chemical representations and pedagogy, sample composition) are expected to be very similar to those of this study and since for test booklet B a correlation between the CK and PCK measures could be shown to be $r = .51$, $p < .01$, presenting a large effect according to COHEN (1992), it was decided to use exclusively this test booklet for the forthcoming analyses and study II. In order to evaluate how well students do on both CK and PCK tests, the test scores are divided into certain intervals (answering 0-20%, 20-40%, 40-60%, 60-80%, 80-100% of the overall items correctly). Figure 36 shows the percentage of students whose test scores fall into those intervals. While for PCK, the largest group of students answer 40-60% of the test correctly, the majority of students answer only 20-40% of the CK test correctly.

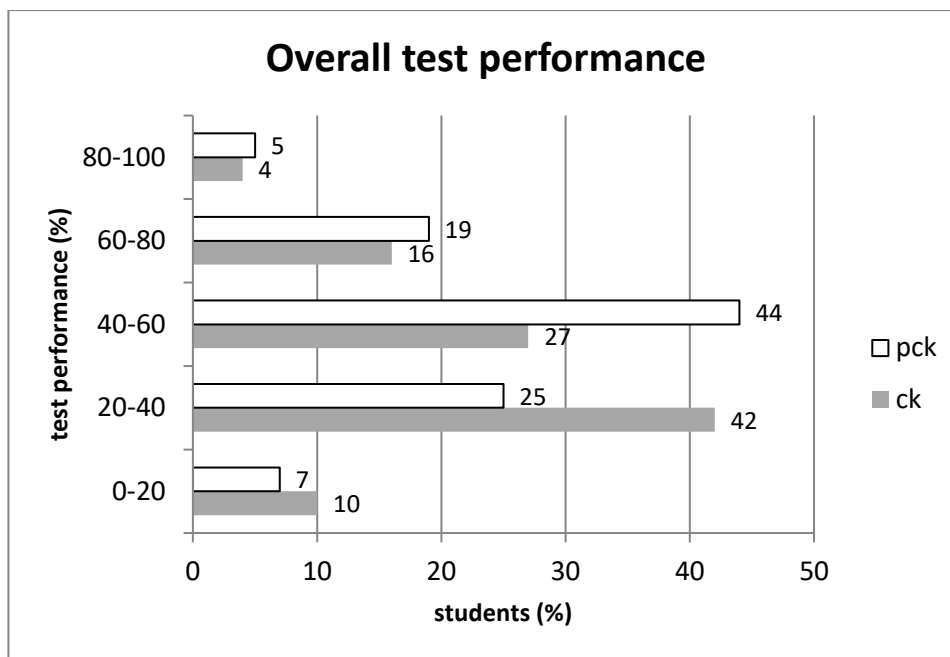


Figure 36: Overview of overall test performance

When looking more closely at test scores based on the year in students' studies, one can see the following trends: i) up to year four, students who have spent more time at university have higher PCK test scores than those in earlier years, while for year five, this trend cannot be observed; and ii) up to year three, students who have spent more time at university have higher CK test scores, while for year four and five, no increase is observed (Figure 37).

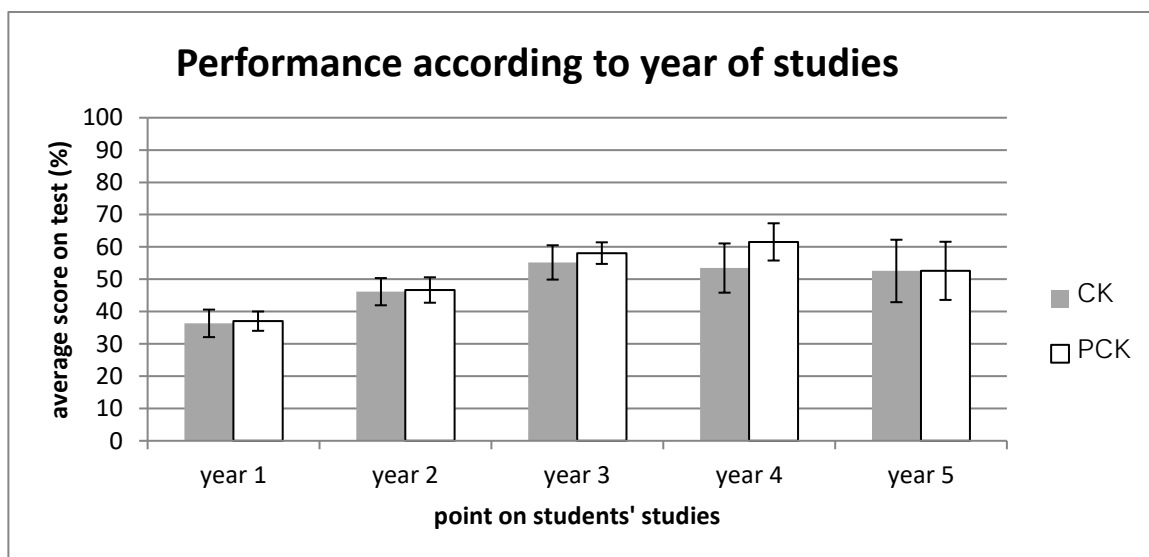


Figure 37: Test performance according to year of studies (test booklet B)

As one goal of this study has been to shed first light into a possible development, only a group comparison by years of study will be calculated here. A more in-depth investigation using regression analysis considering, among other factors, the participants' year of chemistry studies will be done with the data obtained in study II (cf. chapter 6). Comparing the means of these five groups, an ANOVA reveals that for success in the CK test, the year of students' studies does not have a significant effect ($F(4, 63) = 2.41, p = .058$). As the significance value is barely above the $p = .05$ cutoff and a priori hypotheses existed that students in higher semesters should outperform those in lower semesters, planned contrasts are calculated regardless of the overall non-significant effect to find out if the groups differ to some extent. These planned contrasts reveal that being in a year beyond year one significantly increases the result on the CK test ($t(63) = 2.78, p < .05$ (1-tailed), $r_{\text{contrast1}} = .33$). For success in the PCK test, an ANOVA shows the year of students' studies to have a significant medium overall effect ($F(4, 63) = 6.25, p < 0.05, \omega = .48$). Planned contrasts reveal that being in a year beyond year one significantly increases the result on the PCK test showing a medium effect ($t(63) = 4.19, p < .05$ (1-tailed), $r_{\text{contrast1}} = .47$) and that being in a year beyond year two significantly increases the result on the PCK test showing a small effect ($t(63) = 2.17, p < .05$ (1-tailed), $r_{\text{contrast2}} = .26$).

5.6 Summary and discussion of findings

The statistical analysis of the data shows that both research goals regarding reliable test instruments (G1 and G4) could be reached: For both CK and PCK, two subsets of items each make up test instruments with a sufficient internal consistency (as seen in chapters 5.5.3 and 5.5.6, respectively), thus yielding reliable measures. For the validity of the instruments, several considerations have been taken into account to assure that the tests do actually measure what has been intended: CK items have been developed based on secondary curricula and research in the field (cf. chapter 2.3), on school and university text books, lectures, and exams (cf. chapter 4.1.1) as well as on reported student problems with the SPR concept (cf. chapter 2.3.2.3). German university lecturers and international researchers testified to the items' relevant content (cf. chapter 4.3). Regarding PCK, similar steps have been taken: Items have been formulated with other literature on PCK models in mind and a literature review of misconceptions and problems with PCK as well as concrete examples of possible problems while teaching from video analysis or oral exams in teaching modules were also taken into account (cf. chapter 4.2). The items were also assessed and rated as relevant PCK content by lecturers and international researchers; a non-representative survey with chemistry doctoral students also emphasizes the PCK nature of most items (cf. chapter 4.1.1). Based on these considerations, both tests can be considered to have content validity.

When considering the correlations with other measures, the CK instrument significantly correlates with the cognitive ability test, GPA, and the number of the students' university year (cf. chapter 5.5.3) as expected: Students with larger cognitive abilities, a better GPA or more years at university reach higher scores on the tests. Similarly, the instrument correlates significantly with the CK instrument on chemical representations as developed by TASKIN (2014), illustrating that students who do better on the chemical representations test also show higher test scores in the SPR test. As these two areas are closely related (cf. chapter 3), a large correlation makes sense. The significant correlation between PCK instrument and CK test has a large effect – a finding that has also been expected as both tests have been based on the SPR concept. All of these findings emphasize convergent validity of the instruments.

The exemplary qualitative analysis of students' responses to CK items supports the validity of the research goal (G1): Misconceptions that have also been reported by other researchers and presented in chapter 2.3.2.3 could be identified in this study. Just as BHATTACHARYYA AND BODNER (2005) have pointed out, pre-service teachers in this study do also seem to have problems with grasping (larger) functional units, even at a more rudimentary level than analyzed by the researchers: when looking for possible hydrogen atoms that can form hydrogen bonds, pre-service teachers wrongly select hydrogen atoms of methyl groups even though these are not partially positively polarized (cf. item E17 or R11_1 drawing). This thinking also ignores the issue of bond polarity, another difficulty pointed out by PETERSON ET AL. (1989) or TAAGEPERA AND NOORI (2000). The latter also provide results on students' error of bond breaking which appears in pre-service teachers' answers to items in this study as well: GARNETT ET AL. (1995) or TAAGEPERA AND NOORI (2000), for example, point out that students often confuse distinguishing between the breaking of hydrogen bonds necessary for changes in aggregation states to occur and the breaking of covalent bonds in chemical reactions – the same can be seen in answers to items E1 (state changes) and E5 (solubility), for example. In addition, many pre-service teachers in this study give a strikingly wrong definition of a hydrogen bond stating that this type of bond occurs between two hydrogen atoms (of different molecules) or incorrectly draw a hydrogen bond as a bond between two hydrogen atoms of different molecules (item R11_1). This shows that these pre-service teachers have not understood the submicroscopic description or makeup of hydrogen bonds at all. Referring to the understanding of hydrogen bonds mentioned by LEVY NAHUM ET AL. (2008), pre-service teachers in this study also have problems with this concept. Similar to findings by RUSHTON ET AL. (2008), participants in this study also seem to consider steric hindrance as a key factor even when it is not (items E5 & E17). By invoking classic key phrases of organic chemistry such as polarity or steric hindrance when arguing certain reactivity or property (items E1, E5 or E17), many participants in this study exhibit pseudo-conceptual thinking (cf. VINNER, 1997); stating these phrases is not necessarily wrong but shows a simplified thinking and incomplete understanding as these answers forego the complete logical reasoning process. This is also reported by MAEYER AND TALANQUER (2010), who state that students often base their answers on shortcut reasoning procedures, which unfortunately often leads to

incorrect results. This can be seen by participants in this study as well when pre-service teachers refer to polarity of molecules instead of going through the whole reasoning process from polar bonds to polar parts of molecules to intermolecular forces (i.e. from the submicroscopic to the meso level) as a consequence of the orientation of molecules with polar and non-polar parts. Participants also fail to properly explain macroscopic properties such as solubility as they do not explicitly refer to representational and submicroscopic makeup of compounds: not realizing that ether molecules have a free pair of electrons and can therefore form hydrogen bonds with water molecules (item E5) or not recognizing that an illustration representing intermolecular bonds requires more than one molecule (items E17, R11_1 drawing). These findings show learners' difficulties in repeatedly and explicitly integrating all perspectives from the chemical tetrahedron and in relating the different chemical perspectives (cf. MAHAFFY, 2006; GILBERT & TREAGUST, 2009).

With regard to research goal G2, some factors could be identified that have an influence on item difficulty: the level of complexity, the number of perspectives according to the chemical tetrahedron as well as the educational allocation. These findings are in line with research on secondary chemistry: BERNHOLT AND PARCHEMANN (2011) had shown that test items for secondary students ordered according to the Model of Hierarchical Complexity for Chemistry increase in difficulty according to an increase in complexity level. Through this study within the *KiL*-project, a transfer of those findings to tertiary students is possible. Regarding the number of perspectives, many research studies have provided insight into student difficulties with the symbolic aspects of chemistry (see among others KEIG & RUBBA, 1993 or YARROCH, 1985) and the submicroscopic aspects of chemistry (see among others KIND, 2004 or TABER, 2002). If even a single perspective can cause considerable problems for students who are attempting to understand chemistry, it is reasonable to conclude that a combination of those perspectives increases the difficulty of the items significantly, which could be shown through this instrument. Through these findings, the test items seem capable of assessing the concept of SPR in an appropriate way. Having incorporated items that can be answered by an in-depth understanding of school knowledge appears to make sense as this not only forms the basis for everything taught at university in regard to SPR (sensible from a concept

development view), but also enables students in their first year at university to answer a certain number of test items without further university knowledge (sensible from a test design view). This not only increases those students' motivation to fill in the instrument, it also makes it more practicable to use the instrument at different stages of students' university education, either for assessing what students know when entering university or to see how much they might have improved through certain courses or stages.

The exemplary qualitative analysis of students' responses to PCK items supports the validity of the research goal (G4): both a lack in content knowledge as well as a lack in knowledge specific to teaching can stand in the way of answering these items successfully. In the exemplary item with a strong reference to CK aspects (item I6), one can see missing content knowledge as one dominant factor preventing correct answers in PCK items. Some pre-service teachers, for example, do not know how they could hypothetically³⁷ show that diamond can be made to react to carbon dioxide. While the majority of pre-service teachers' answers to this item show that they do know that diamond can indeed be burned, slightly over fifty percent of these participants provide the subsequent part necessary for a fully appropriate answer, i.e. describing the analysis of the product which allows a student to conclude that, as only carbon could have reacted with oxygen in the burning process to form carbon dioxide, which can then be detected, diamond must be made of carbon. The other answers lack this second step – this omission being a truly PCK-specific criterion as it shows future teachers' knowledge (or more specifically lack thereof) of how to ideally react in a teaching situation in order to reduce potential pitfalls or anticipate students' problems. When asked to anticipate student examples relating to one of the most common misconceptions described in chapter 2.3.2.3, i.e. wrongfully assigning macroscopic properties to individual atoms (item S4), only a third of the pre-service teachers provide at least one appropriate answer. As knowing and anticipating what is on students' minds is a central sub-dimension of pedagogical content knowledge (cf. chapter 4.2.1), and as pre-service teachers are likely to encounter this type of problem in their future career up to the middle of secondary education, it has been expected that participants in this study would be able to name

³⁷ While financial conditions do not usually permit this experiment in school, reaction conditions would allow for it.

at least one such example. One might speculate that participants in this study either hold that misconception themselves or have not started thinking about possible student answers to a large extent during their university education. When it comes to the exemplary item dealing with experiments and models (item EM10), answers show that pre-service teachers are, on average, more knowledgeable regarding the utilization of possible experiments introducing the topic of acids and bases than they are regarding the students' misconceptions item (item S4). As both items require participants to name three examples of a certain PCK aspect, the item format should not be a decisive differentiation. It is more likely that different learning opportunities arise during pre-service teachers' education and that there is a difference in accessibility of the knowledge aspects: While they already know introductory experiments from their own school experience, from observing chemistry lessons at school and from being exposed themselves to experiments during their CK modules at university, observing students' possible thinking is not as easy as it is not directly accessible. When observing teaching situations during their tertiary education, pre-service teachers have to actively and deliberately set out to make sense of student answers, reflecting upon underlying reasons or pondering which concepts the students might hold. From this short discussion of exemplary PCK items, no general conclusion can be made yet. It shows, however, that different foci can already be pointed out with these three items as possible differentiating factors: the relation of CK and PCK aspects in a PCK item, the availability of learning opportunities in tertiary education, and the degree of explicitness of certain PCK aspects so they can be integrated into the present knowledge base.

Regarding the research goal of assessing student performance on both tests (G3 and G5), it was found that students' scores on both tests were at or below 50% of the possible score: While almost half of the students, and with this the largest subgroup, answered an average 50% of the PCK test correctly, a similarly large group of students only scored 20-40% on the CK test. This might look as if in particular the CK test was too difficult. Since the test instrument is to be used as a measure to assess CK and its growth across the complete duration of tertiary education, however, and the test population in this study are on average only in their 5th semester, the instruments offer some room for assessing further knowledge improvement. A

comparable result has been obtained by TASKIN (2014, p.142f.): students averaged 58% of correct answers in the CK instrument and 61% in the PCK instrument on chemical representations correctly. As expected, pre-service teachers in their first year of tertiary education have a significantly lower result on both CK and PCK test than students in higher years; for pre-service chemistry teachers in further years the CK results do not differ significantly while students in the second year also have significantly lower results on the PCK test than more advanced students. This finding can be explained through the general study structure in tertiary education. In their first year, pre-service teachers are introduced to general chemical ideas and take first courses in the different sub-disciplines. Further courses in the sub-disciplines follow throughout the remainder of the degree program in which that first-year knowledge is applied and built on. During these various application and in-depth instruction phases many specialized knowledge facts and some further aspects of the SPR-concept are likely to be covered, so that it is not surprising that the CK results in a SPR concept test do not improve considerably after year one. For PCK a somewhat similar explanation comes into play. Introductory PCK courses in tertiary education start somewhere between year one and two, so that this can be considered the basis which further courses build on. The fact that beginning pre-service teachers already score relatively high on the PCK test might be explained through a possible reflection and integration of the pedagogical content knowledge learnt at university with their knowledge and experiences gained at school (cf. Figure 1).

After the successful creation of valid and reliable test instruments assessing pre-service teachers' content and pedagogical content knowledge regarding the concept of structure-property-relations, the subsequent study II shall also supply information about the correlation of the CK and PCK scales. As the *KiL*-project's intention was to also investigate how the three areas of professional knowledge are related and whether certain conditions favor the acquisition of professional knowledge, data from study II will hopefully also shed more light on the conceptualization of pre-service teachers' professional knowledge. This might turn out to be especially beneficial as, in a second step, the SPR instruments will be considered jointly with the instruments on chemical representations developed by

TASKIN (2014) as well as the instrument on pedagogical knowledge developed by other members of the *KiL*-project.

While in study I, each university was visited only for one day, in study II, each university will be visited on two consecutive days giving participants the opportunity to fill in tests in two of the *KiL* project's subjects if they studied two of the sciences and/or math.

As the subsequent study II will also be a cross-sectional study with an overall larger sample, the development within the university stages shall be investigated further. It shall be interesting to see whether other variables such as the student's background also play a role. In order to investigate this, other aspects and instruments will complement study II to see whether certain learning opportunities or psychological constructs such as self-concept or interest have an influence on answering CK and PCK items.

6 STUDY II: STRUCTURE OF PROFESSIONAL KNOWLEDGE

6.1 Research objective

In contrast to study I which has been intended to test the items and get information on narrowing the item pool, a subsequent second study shall assess the psychometric quality of the test instruments and identify correlations with other instruments to provide insight into the chemistry knowledge base of pre-service teachers. Moreover, an additional goal is to get an idea of the development of this chemistry knowledge over the course of pre-service teachers' tertiary education.

The main intention of this study is, therefore, to shed more light on professional knowledge, thus assessing its dimensionality and correlations with other measures for validity purposes. A further purpose of this study is to get first insights into how knowledge develops by comparing pre-service teachers of different years of study.

Therefore, the resulting overall research objective guiding this second study is the following:

- O_{study II}:** To describe the overall structure of pre-service chemistry students' professional knowledge and to find aspects that influence pre-service teachers' chemistry content knowledge and chemistry pedagogical content knowledge.

This overall research objective can be further divided into more detailed questions first centering solely on the SPR-part (Q1-Q3) and then also including data and results from the adjacent thesis on chemical representations (cf. TASKIN, 2014, Q4-Q5) as well as pedagogical knowledge (Q5):

- Q1:** What structure of the subject-specific knowledge dimensions regarding SPR can be found for pre-service chemistry teachers?

Hypothesis: As other studies have found the traditional structure of the two subject-specific dimensions CK and PCK to be distinct dimensions in their models (cf. *TEDS-M*, *COACTIV*, *ProwiN*), a separation between CK and

PCK dimensions regarding SPR is also anticipated. In the above-mentioned studies, both dimension show small to medium correlations. As both dimensions have been designed with an identical conceptual focus, it is expected to also find a medium correlation between the CK and PCK dimension for SPR.

Q2: Which factors explain variance in test outcome?

Hypothesis: Person variables, school variables as well as university variables are expected to contribute significantly in explaining variance in test outcomes both for CK and PCK. As being male has been found to be a predictor for success by TASKIN (2014), it is expected to also have an effect in this study. Students with a better GPA and a higher semester in their tertiary program had better results in study I, therefore comparable results are expected in this study. As having taken chemistry courses at upper secondary, being enrolled in two STEM subjects for teaching and having passed exams instead of just listening to lectures at university provide more chemistry learning opportunities for both CK and PCK, these factors should also have an impact on test outcome. The teacher track is also expected to have an effect as in *COACTIV* math teachers for the Gymnasium showed better results in CK and PCK than their peers for other school types (KRAUSS ET AL., 2008). As the first group has more learning opportunities for content knowledge aspects in chemistry as well, a similar finding is expected in this study.

Q3: What differences in subject-specific knowledge dimensions regarding SPR can be found for pre-service teachers in different stages of their tertiary education?

Hypothesis: As has already been suggested by results from study I, the test outcome is expected to increase with students' higher semester at university. While this is anticipated for the CK and PCK dimension, CK is expected to increase more steeply during the beginning of the studies and taper off toward the end, and PCK should start to increase at a later point and develop toward the end of the studies. This trend is expected as it

mirrors the focus of university education providing more modules on CK in the beginning and a shift to PCK towards the end.

- Q4:** How are the instruments on SPR and chemical representations (CRep) connected?

Hypothesis: As CRep and SPR are conceptually closely related (cf. chapters 3 and 4.1), a large correlation between the CRep and SPR dimensions is expected to be found.

- Q5:** What is the overall structure of pre-service chemistry teachers' professional knowledge?

Hypothesis: As it is expected that CRep and SPR measures will correlate to a large extent (cf. hypothesis to Q4), the chemistry professional knowledge should mirror the traditional structure of three dimensions (CK, PCK, PK, cf. KRAUSS ET AL., 2013; BLÖMEKE ET AL., 2010; and KIRSCHNER, 2013) when considering CRep and SPR jointly.

6.2 Instruments

As only one test booklet was applied during study II, a sub-set of items from study I had to be selected in accordance with statistical analyses of that data (cf. chapter 5.5) and with content-related considerations: the remaining instrument was to still contain a variety of items regarding sub-disciplines, complexity, educational allocation, and regarding a development of SPR aspects. Again, an agreement with the colleague working on chemical representations had to be reached so that items would not be a necessary prerequisite of one another.

From the items trialed in study I regarding SPR, 20 items for CK and 14 items for PCK were chosen to be administered in study II. The majority of the CK and PCK items were taken from test booklet B of study I.

As with study I, other instruments were administered along with tests measuring pre-service teachers' CK and PCK regarding SPR. Measures to assess students' CK and PCK regarding chemical representations were again included in the test booklet. Also, students were again given an instrument on pedagogical knowledge (cf. chapter 3) to compliment the dimensions of professional knowledge. Instead of a cognitive ability test, only the student's GPA was asked this time. It was, however, accompanied by the student's last chemistry grade in upper secondary. Furthermore, participants answered questions regarding their chemistry education at secondary level so that it would be possible to compare groups of students with different chemistry education background. As the test instruments have been developed to assess a thorough understanding of the SPR concept both for the content and pedagogical content knowledge, more intense learning opportunities in school were expected to correlate with a higher test outcome to a larger extent. Participants were asked to provide their gender, age and semester at university, their second subject, their program of study (Bachelor vs. Master vs. Staatsexamen)³⁸ as well as the anticipated level at which they would be teaching after finishing their education, i.e. their teacher track (cf. chapter 2.1.1). The latter two variables (program of study and teacher track) have accounted for a large difference in teachers' content or pedagogical content knowledge as found in other studies, such as the *COACTIV*-study and *ProwiN*-study (KRAUSS ET AL., 2008, and WALPUSKI ET AL., 2012, respectively) and it could therefore reasonably be expected to find similar results in this study.

According to the *COACTIV*-model of professional competence (BAUMERT & KUNTER, 2013, cf. Figure 8), further aspects besides professional knowledge influence teachers' competence, i.e. the way in which a teacher prepares for, acts in, and reflects on a specific learning situation in a particular class: beliefs/values/goals, motivational orientations, and self-regulation.

³⁸ At the time of this study, pre-service teachers could study under the traditional teaching program, the Staatsexamen, which prescribed to a much lesser degree the exact courses one had to take as well as their order, or they could study under the newer Bachelor/Master-model.

Assuming that this model can also be used for pre-service teachers and their prospective performance as teachers, the following validated and published instruments were administered to the participants alongside the CK and PCK test to provide some insights into the validity of the subject-specific measures:

a) Domain-specific interest

Six items of the Study Interest Questionnaire (*SIQ*; SCHIEFELE, KRAPP, WILD & WINTELER, 1993) were used to assess participants' interest in the chemistry content part, in the teaching of chemistry part, and in the general pedagogy part of their studies with items such as "Busying myself with the contents and issues of this domain is one of my favorite activities". Participants were requested to answer the items separately for each part of their studies on a 4-point Likert scale (from 1: *does not apply at all* to 4: *fully applies*). The SIQ has been found to correlate significantly with an achievement measure obtained through participants' average university grades, having a medium effect ($r = .33$; SCHIEFELE ET AL., 1993, p.346).

b) Domain-specific self-concept

To obtain participants' self-assessment regarding their competencies, two subscales of the Berlin Evaluation Instrument for self-evaluated student competencies (acronym: *BEvaKomp*; BRAUN, GUSY, LEIDNER & HANNOVER, 2008) were used: content competence and methodological competence. Participants were asked to answer the items for the chemistry content part, the teaching of chemistry part and for the pedagogical part of their studies on a 4-point Likert scale (from 1: *does not apply at all* to 4: *fully applies*). An item in this scale is worded like this, for example "I can work out similarities and discrepancies (e.g. those between different models) within the domain". The authors themselves state that the way the content competence scale has been conceived, it relates closely to the way academic self-concept³⁹ is measured (BRAUN ET AL., 2008, p.38). As WEIBENO AND ECK report in their study on performance in politics, self-concept and knowledge correlate to an average of $r = .25$ (WEIBENO & ECK, 2013, p.59).

³⁹ For an overview regarding academic self-concept, see MÖLLER AND KÖLLER (2004), for example.

c) Self-efficacy beliefs about experiments in the classroom

In order to assess participants' self-efficacy regarding successfully planning and executing experiments in the classroom, 14 items of an instrument originally developed for the physics classroom by MEINHARDT, RABE AND KREY (2016) were adapted to chemistry. Pre-service chemistry teachers were asked to answer on a 6-point Likert scale (from 1: *does not apply at all* to 6: *fully applies*) items such as this one: "Even time pressure during the preparation phase does not stop me from developing an experiment that fits my teaching goals". In a meta-analysis of studies by MULTON, BROWN & LENT (1991), for example, self-efficacy was reported to correlate with academic performance for college students ($r = .35$).

d) Self-efficacy beliefs about teacher activities

Four items were used to assess participants' self-efficacy beliefs about their own prospective activities as a teacher in the classroom (adapted from the Efficacy for instructional strategies factor by TSCHANNEN-MORAN & WOOLFOLK HOY, 2001). Participants were requested to answer items such as "How convinced are you that in your chemistry class, you could find an alternative explanation or a different example if students do not understand something?" on a 9-point Likert scale (from 1: *not convinced at all* to 9: *totally convinced*), providing an insight into their own teaching convictions.

e) Beliefs about Nature of Science

In order to measure participants' understanding about how scientific knowledge develops (i.e. Nature of Science), 23 items of the 'Student Understanding of Science and Scientific Inquiry (SUSSI)' (LIANG ET AL., 2006) were given to pre-service teachers. Items were to be assessed on a 5-point Likert scale (from 1: *strongly disagree* to 5: *strongly agree*) and looked like this: "Scientists do not use their imagination and creativity because these conflict with their logical reasoning" (LIANG ET AL., 2006, p.29).

The instruments described above were included in all test booklets in this study and are expected to correlate with the SPR knowledge dimensions according to Table 33:

The highest correlations are expected with the respective measures, i.e. the CK dimension with all the chemical content-related measures and the PCK dimension with those measure assessing interest, self-concept or beliefs regarding the teaching of chemistry. The correlations between the CK dimension and measures assessing some sort of chemistry teaching (and vice versa) are expected to correlate less strongly. Measures evaluating students' interest and beliefs in general pedagogy are expected to correlate to a lesser extent with the PCK and to an even lesser extent, if at all, with the CK dimension.

Table 33: Additional measures in study II and expected correlations with the SPR knowledge dimensions

Instrument	Expected correlations with...	
	CK _{SPR}	PCK _{SPR}
Domain-specific INTEREST (<i>SIQ</i>)		
Chemical content	high	medium
Teaching chemistry	medium	high
Pedagogy	none	small
Domain-specific SELF-CONCEPT (<i>BeVaKomp</i>)		
Chemical content	high	medium
Teaching chemistry	medium	high
Pedagogy	none	small
Domain-specific BELIEFS		
Self-efficacy beliefs about experiments (RABE)	high	high
Self-efficacy beliefs about teacher activities (TSCHANNEN-MORAN & WOOLFOLK HOY)	medium	high
Nature of Science (<i>SUSS</i>)	medium	medium

6.3 Data preparation & statistical measures

After a test administration that paralleled that of study I a year earlier, students' answers were again digitalized and coded according to the coding schemes already developed for study I. As before, missing data were assigned a false coding.

For the description of the sample, classical test theory was applied and the program *SPSS 23.0* was used. For all other analyses, however, item response theory was chosen. A statistical analysis of item fit parameters was calculated with the program *ConQuest*. In particular, the weighted mean square and corresponding t values of the included items are of interest: weighted mean-squares (*WMNSQ*) express the discrepancy between the observed data and the model and should lie within a range of 0.70 to 1.30 (WRIGHT & LINACRE, 1994) while corresponding t values, suggested to be within a range between -2 and 2, state whether the deviation of the *WMNSQ* values from the expected values is significant. In addition, Rasch analysis provides reliability values for the dimensions in question.

EAP/PV (Expected A Posteriori/Plausible Value), generated by the *ConQuest* software, is used as a reliability estimate for the CK and PCK dimensions. It provides information on "how reliably the items can be used to distinguish students' underlying abilities" (FULMER, CHU, TREAGUST & NEUMANN, 2015, p.8). It is obtained as "explained variance according to the estimated model divided by total person variance" (DRANEY & WILSON, 2008, p.425) and its value is comparable to Cronbach's α (ROST, 2004), yet a value in the range of 0.64 to 0.72 is sometimes reported as moderate to good reliability (FULMER ET AL., 2015).

Furthermore, Wright Maps are used which show the difficulty for each item on the right (denoted by a number) and the ability of a person for each dimension investigated on the left (denoted by an 'x'). The more difficult an item, the further toward the top it is. Similarly, the more able a person, the further toward the top that person is. If person ability and item difficulty are on the same level, than that person has a 50% probability of solving that item correctly. For an item below the level of a person's ability, the probability for the person to solve that item correctly is larger than 50%. One mean (item difficulty or person ability) is set to zero and the other mean can be calculated. The smaller the difference in means between item parameter and

person parameter, the better the match between items and sample population, or the more appropriate the test is for the sample (WU, ADAMS, WILSON & HALDANE, 2007). Wright Maps thus provide insights into the range of item difficulties, whether the person ability is distributed symmetrically, or whether the data are skewed (MILES & SHEVLIN, 2001).

When considering the dimensionality of models, two information-based criteria are taken into account: i) the AIC (Akaike's Information Criterion, cf. AKAIKE, 1981) and ii) the BIC (Bayes' Information Criterion, cf. WILSON, DE BOECK & CARSTENSEN, 2008). While the AIC yields more sound results in short tests that have many response patterns, the BIC is to be preferred for long tests with few response patterns (ROST, 2004). For both information-based criteria it follows that the lower the coefficient, the better the model fits the data (SCHERMELLEH-ENGEL, MOOSBRUGGER & MÜLLER, 2003; WILSON ET AL., 2008).

In order to elicit factors influencing students' performance on the tests, a stepwise multivariate regression will be calculated.

6.4 Results

6.4.1 Description of sample

In this study conducted in 2013, 322 students from twelve tertiary education institutions participated. Of those, 11.5% had taken no chemistry at upper secondary level, 44.4% had taken courses at a basic and 43.8% at an advanced level. Regardless of students' level of chemistry courses at upper secondary, the average participant's GPA was 2.2 ($SD = 0.5$) with 1.0 being the best and 4.0 being the worst passing grade. At the point of data collection, participants were also asked to provide their last chemistry grade from upper secondary, which was on average 1.8 ($SD = 0.8$). Participants were on average in their 5th chemistry semester ($SD = 2.5$) and the majority was enrolled in a pre-service teaching program following the Staatsexamen model (53.1%, and 46.0% being then enrolled in a Bachelor/Master-program). With 86.0%, most of the participants were striving for a degree that would later enable them to teach at the highest school type in Germany, i.e. the Gymnasium. 37.6% of the participants had not yet planned and taught a single chemistry lesson themselves, 17.4% had gained very little experience (one to two lessons), 24.5% could be considered somewhat experienced (having taught three to six lessons) and 17.4% had gained quite a bit of practical experience (seven to fifteen lessons) when it comes to having taught chemistry. Regarding the second subject, 60.6% were educated in at least one other science (biology, physics, geography), 23.4% had chosen math or informatics, and 15.8% had taken up a subject other than science or math.

The students were on average 23.3 years old ($SD = 3.0$) and the majority of them was female (62.7%).

Table 34: Comparison of sample of study I and II

variables		Study I	Study II
number of participants		220	322
average age		23.1	23.3
average semester		5.2	5.5
average grade in school leaving certificate		2.3	2.2
gender	female	60.5	62.7
	male	39.5	37.3
level of chemistry at upper secondary	none	7.7	11.5
	basic	43.6	44.4
	advanced	44.1	43.8
degree	BA/MA	46.3	46.0
	Staatsexamen	53.6	53.1
future teaching at school type	non-academic track	26.8	14.0
	academic track	73.2	86.0
second subject STEM	no	19.5	15.8
	yes	80.5	83.5

Comparing the sample in study II to that of study I conducted a year earlier (cf. Table 34), both are quite similar in gender, age, distribution regarding chemistry at upper secondary, GPA, and their enrolled teaching program. The samples differ slightly regarding their teaching aspiration at their anticipated school type and their second subject: A higher percentage of participants of study II is planning on going into teaching at the Gymnasium and is doing so with a science as a second subject.

6.4.2 Descriptive data analysis

When investigating the structure of pre-service teachers' SPR knowledge, the resulting instruments need to be looked at first regarding item fit parameters and reliability values. Then, the question of dimensionality needs to be answered, and if more than one dimension exists, the corresponding correlation has to be calculated. As a third step, a closer analysis of item and person parameters through a Wright Map provides further insights into the dimension(s).

Considering item fit parameters, the resulting *WMNSQ* parameters show appropriate values for all items and the *t* values are within the suggested range only

after excluding three items (i.e. FW36, FDW14, and FDW24) from the model. For the remaining 19 CK and 12 PCK items in the resulting model, the *WMNSQ* values are in the range of .87 to 1.15 and the corresponding *t* values in the range of -1.6 to 2.0. Regarding the reliability of both resulting measures, the EAP/PV (Expected A Posteriori/Plausible Value) values are sufficient with $EAP/PV_{CK} = .79$ and $EAP/PV_{PCK} = .74$.

Table 35: Model fit analysis, SPR dimensions

	AIC	BIC
1 dimension (SPR knowledge)	15050.50	15273.20
2 dimensions (SPR _{CK} vs. SPR _{PCK})	14979.96	15210.21

Note. AIC = Akaike information criterion; BIC = Bayesian Information Criterion.

A dimensional analysis (Table 35) and subsequent χ^2 test confirms that the two-dimensional model (CK and PCK as separate dimensions) fits significantly better than the one-dimensional model ($\chi^2 = 74.54$, $p < .001$), which assumes all SPR items to fall into one dimension. The latent correlation between both dimensions is $\psi = .72$.

When analyzing the relation between the difficulty of items and the ability of students, a Wright Map is generated, directly comparing those two parameters (Figure 38⁴⁰). In this Wright Map, person abilities are mapped for both CK and PCK test instrument and each 'x' represents 2 cases⁴¹.

The Wright Map shows that the person abilities are distributed rather symmetrically, and that the corresponding values for item difficulties are distributed along the scale with a similar range as the person abilities.

When looking at the mean of item parameters and standard deviation (cf. Table 36), it becomes apparent that there is a high overlap

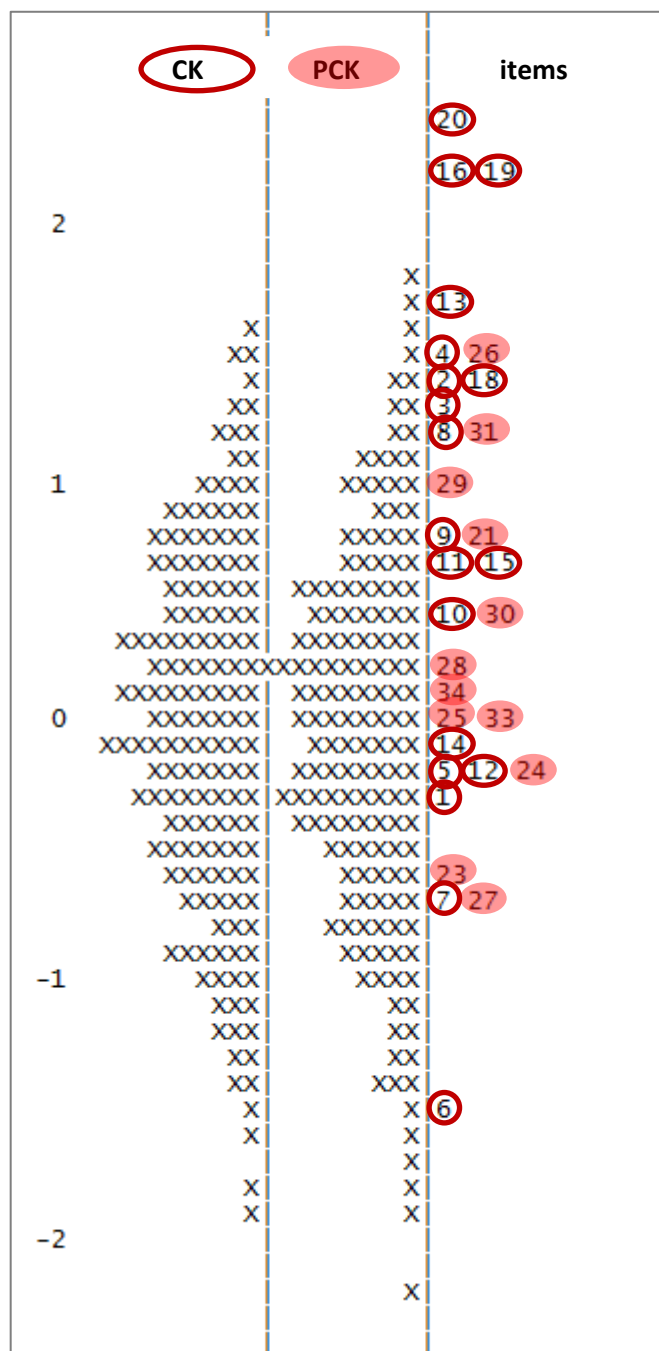


Figure 38: Wright Map of CK and PCK

⁴⁰ The color coding in this Wright Map has been chosen so that it is consistent with the Wright Map in chapter 6.4.5 and resulting professional knowledge structures: SPR items are in red and items on chemical representations will be in blue, CK items are circled while the PCK items have a corresponding colored solid filling.

⁴¹ For an overview which item number in the Wright Map corresponds to which item in the test, see Annex III.

between the distributions of the item and person parameters and a difference of up to 0.78 logits exists between their means⁴². It can therefore be concluded that the items are marginally too difficult for the population, but both tests can still be considered appropriate for the sample.

Table 36: Overview of means for both SPR dimensions

	Mean item difficulty	<i>SD</i>
CK	.78	1.03
PCK	.33	.65

6.4.3 Factors explaining variance in test outcome

In addition to the subject-specific knowledge tests, participants were asked to provide other data as well (cf. chapter 6.2). Some of these variables can be grouped into three main categories (Figure 39, the color coding will be used in the presentation of results as well): a personal variable, school variables, i.e. variables assessing students' secondary education, and university variables, i.e. variables assessing students' tertiary education.

Considering the effects of the variables individually might already show first insights into their possible influence as factors when it comes to looking at a multivariate regression later. For the following analyses, person abilities which have been obtained through Rasch analysis are used.

⁴² The mean of person parameters is set to zero.

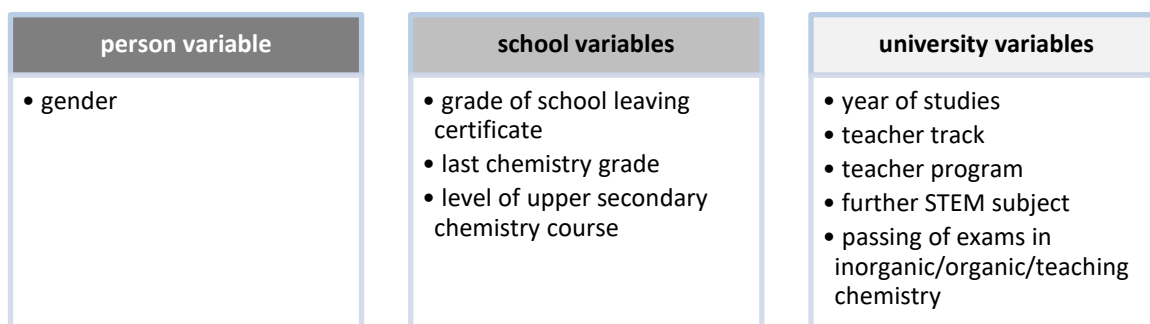


Figure 39: Overview of different variables that might have effect on test performance

Person variable

An independent t -test comparing the means regarding **gender** shows that for the CK instrument, females ($M = -.19$, $SE = 0.07$) have a significantly lower ($t(320) = -4.20$, $p < .001$) test outcome than males ($M = .24$, $SE = 0.07$). For the PCK instrument, females ($M = -.20$, $SE = 0.07$) also show a significantly lower outcome ($t(320) = -3.88$, $p < .001$) than males ($M = .22$, $SE = 0.08$). For both instruments, the effect ($r_{CK} = .23$ and $r_{PCK} = .21$) is small.

School variables

A correlation analysis shows significant correlations between the person parameters and the **grade of the school leaving certificate** (Pearson, bivariate: $r_{CK} = -.45$, $p < .01$, $r_{PCK} = -.29$, $p < .01$) as well as the students' **last chemistry grade at the upper secondary level** (Pearson, bivariate: $r_{CK} = -.36$, $p < .01$, $r_{PCK} = -.23$, $p < .01$). Students with a better grade in chemistry or school leaving certificate perform better on both test instruments. For both variables, the CK instruments correlates to a higher degree.

When comparing the test outcome for students with different **levels of chemistry at upper secondary**, the following charts (Figure 40 and 41 for CK and PCK, respectively) show the means of the three groups.

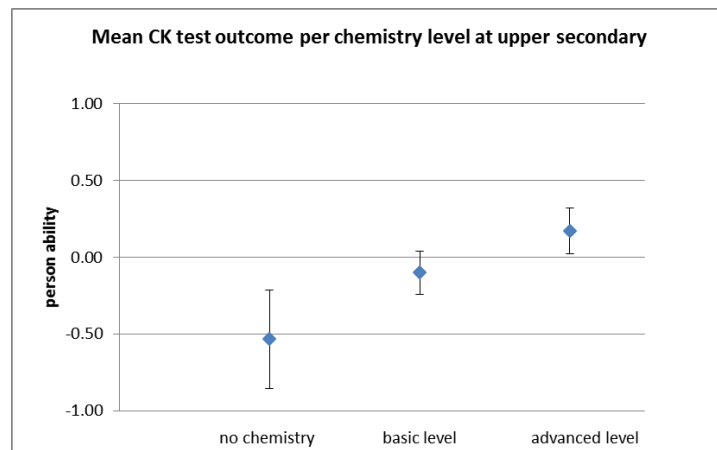


Figure 40: Mean (span is +/- 2 SD) CK test outcome per chemistry level at upper secondary

An ANOVA shows that having studied chemistry at upper secondary level has a significant effect on students' success in the CK test, $F(2, 318) = 10.15$, $p < .001$, $\omega = .23$. Planned contrasts reveal that those students who had taken chemistry at upper secondary show significantly better results in the CK test than those who had not taken chemistry at upper secondary, $t(318) = 3.68$, $p < .05$, $r_{contrast1} = .20$, and that those having taken chemistry at an advanced level show significantly better results than students who had studied chemistry at a basic level, $t(318) = 2.60$, $p < .05$, $r_{contrast2} = .14$.

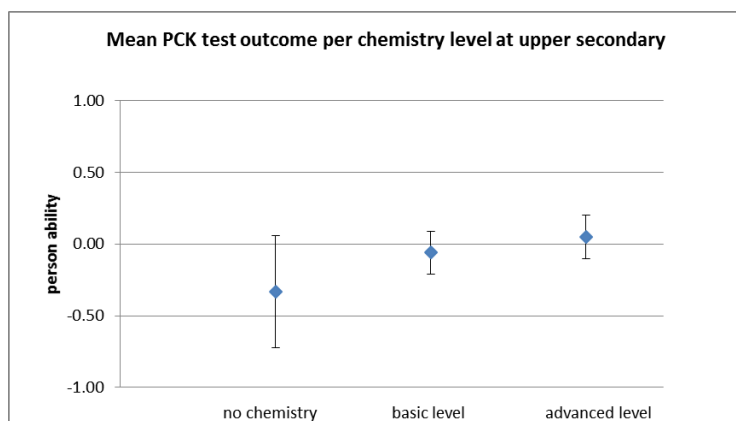


Figure 41: Mean (span is +/- 2 SD) PCK test outcome per chemistry level at upper secondary

Regarding students' success in the PCK test, having studied chemistry at upper secondary level does not have a significant effect, $F(2, 318) = 2.45, p = .088$. However, a priori formulated hypotheses state that students having taken chemistry at upper chemistry and thus having had more unintentional learning opportunities (cf. Figure 1) for student mistakes and difficulties, should show better test outcomes than those without any chemistry courses at upper secondary. That is why planned contrasts are calculated regardless the overall non-significant effect. These reveal that those students who had studied chemistry at upper secondary show significantly better results in the PCK test than students not having taken chemistry at upper secondary, $t(318) = 1.99, p < .05, r_{contrast1} = .11$. However, results in the PCK test do not differ significantly between students who had taken chemistry at an advanced or basic level, $t(318) = 0.98, p = .326$.

University variables

A correlation analysis shows significant correlations between the person parameters and the **number of years** of students' chemistry education program (Pearson, bivariate: $r_{CK} = .20, p < .01, r_{PCK} = .27, p < .01$). Students who have been studying at university for longer periods of time outperform their colleagues in both CK and PCK test who have attended university for fewer semesters.

An independent t -test comparing the means regarding **teacher track**⁴³ shows that for the CK instrument, students aiming for a future career in the non-academic track ($M = -.45, SE = 0.13$) have a significantly lower ($t(320) = -3.37, p < .05$) test outcome than students aiming for a career in the academic track ($M = .04, SE = 0.05$). For the PCK instrument, students aiming for a future career in the non-academic track ($M = -.46, SE = 0.14$) also show a significantly lower outcome ($t(320) = -3.24, p < .05$) than students aiming for a career in the academic track ($M = .02, SE = 0.06$). For both instruments, the effect is small (for both: $r = .18$).

⁴³ A future career in the academic track means a teacher will later work in the Gymnasium-type schools while a teacher pursuing a degree in the non-academic track will later work at school types that do not offer upper secondary education to students.

An independent *t*-test comparing the means regarding the students' **second (or third) subject** shows that for the CK instrument, students with a second subject other than STEM ($M = -.26$, $SE = .13$) do not have a significantly different ($t(318) = -1.93$, $p = .054$) test outcome than students with their second subject from the STEM disciplines ($M = .01$, $SE = .05$). For the PCK instrument, however, students with a second subject other than STEM ($M = -.30$, $SE = .16$) show a significantly lower outcome ($t(318) = -2.06$, $p < .05$) than students with their second subject from the STEM disciplines ($M = .00$, $SE = .06$). With $r = .11$, the effect is small.

An ANOVA shows that **having attended classes and passed inorganic chemistry exams** has a significant effect on students' success in the CK test ($F(3, 318) = 5.37$, $p < .05$, $\omega = .21$). Planned contrast reveal that neither attending classes ($t(318) = 1.34$, $p = .181$) nor attempting to take exams but possibly failing ($t(318) = 1.07$, $p = .287$) have a significant effect. Rather, compared to the other students, only those who had passed one or more exams in inorganic chemistry ($t(318) = 2.04$, $p < .05$, $r_{contrast3} = .11$) show a significantly higher outcome on the CK test. Even though regarding PCK, the ANOVA states that having attended classes and passed inorganic chemistry exams has a significant effect on students' success in the PCK test as well ($F(3, 318) = 2.65$, $p < .05$, $\omega = .13$), planned contrast reveal no significant differences between the groups.

An ANOVA shows that **having attended classes and passed organic chemistry exams** has also a significant effect on students' success in the CK test ($F(3, 318) = 13.25$, $p < .05$, $\omega = .32$). Planned contrast reveal that neither attending classes ($t(318) = 0.92$, $p = .361$) nor attempting to take exams but failing ($t(318) = 1.31$, $p = .192$) are significant. Rather, only students who had passed one or more exams in organic chemistry ($t(318) = 3.82$, $p < .05$, $r_{contrast3} = .21$) show a significantly higher outcome on the CK test. Regarding PCK, having attended classes and passed organic chemistry exams also has a significant effect on students' success ($F(3, 318) = 8.38$, $p < .05$, $\omega = .26$). Planned contrast reveal that simply attending classes does not have a significant effect ($t(318) = 0.22$, $p = .827$), while having attempted an exam but possibly failed significantly increases PCK test outcome compared to just having attended classes ($t(318) = 2.52$, $p < .05$, $r_{contrast2} = .14$). Compared to students having failed the exam(s), those having passed one or more exams in organic

chemistry ($t(318) = 3.09, p < .05, f_{contrast3} = .17$) show a significantly higher outcome on the PCK test.

An ANOVA shows that **having attended classes and passed exams on how to teach chemistry** has a significant effect on students' success in the CK test ($F(3, 318) = 3.77, p < .05, \omega = .17$) but planned contrast reveal no significant differences between the groups. Regarding PCK, having attended classes and passed exams on how to teach chemistry does also have a significant effect on students' success ($F(3, 318) = 4.87, p < .05, \omega = .19$). Planned contrasts reveal that neither attending classes ($t(318) = -0.76, p = .448$) nor attempting to take exams but possibly failing ($t(318) = -1.04, p = .297$) have a significant effect. Rather, only students who had passed one or more exams on how to teach chemistry ($t(318) = 2.17, p < .05, f_{contrast3} = .12$) show a significantly higher outcome on the PCK test.

According to the different operations performed, the results of the individual factors' impact are summarized in the following tables (Tables 37 to 39, non-significant results are faded out) before a regression analysis is conducted.

Table 37: Results for correlations of person parameters and certain variables

Correlations	CK _{SPR}		PCK _{SPR}	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
Grade of school leaving certificate	-0.45	**	-0.29	**
Last chemistry grade in school	-0.36	**	-0.23	**
Year of studies	.20	**	.27	**

Table 38: Results for t-tests of certain variables on person parameters

t-tests		CK _{SPR}			PCK _{SPR}		
		<i>Mean</i>	<i>p</i>	Effect size	<i>Mean</i>	<i>p</i>	Effect size
Gender	female	-0.19	***	.23	-0.20	***	.21
	male	.24			.22		
Teacher track	non-academic	-0.45	*	.18	-0.46	*	.18
	academic	.04			.02		
Further subject	no STEM	-0.26	n.s.		-0.30	*	.11
	STEM	.01			.00		

Table 39: Results for ANOVA of certain variables on person parameters

ANOVA	CK _{SPR}		PCK _{SPR}	
	<i>p</i>	Effect size	<i>p</i>	Effect size
Chemistry level, overall	***	.23	n.s.	
any chemistry better than none	**	.20	**	.11
advanced chemistry better than basic	**	.14	n.s.	
Inorganic chemistry, overall	*	.21	*	.13
attendance (and exams) better than no attendance	n.s.		n.s.	
exam taking (failing/passing) better than mere attendance	n.s.		n.s.	
exam passing better than exam failing	*	.11	n.s.	
Organic chemistry, overall	*	.32	*	.26
attendance (and exams) better than no attendance	n.s.		n.s.	
exam taking (failing/passing) better than mere attendance	n.s.		*	.14
exam passing better than exam failing	*	.21	*	.17
Chemistry teaching, overall	*	.17	*	.19
attendance (and exams) better than no attendance	n.s.		n.s.	
exam taking (failing/passing) better than mere attendance	n.s.		n.s.	
exam passing better than exam failing	n.s.		*	.12

Comparing the effect sizes for the variables, one can find several small effects and a few medium effects. For the CK instrument, the following single variables have a medium effect: i) GPA, ii) last chemistry grade in school and iii) having attended and passed organic chemistry exams at university. All other significant effects of the variables investigated are small. For the PCK instrument, on the other hand, all variables showing significant differences have a small effect on test outcome.

An analysis of the impact of single factors on CK and PCK test outcome does not provide any information on simultaneous influences of other covariates. When looking at the results regarding the participants' chemistry level in upper secondary education as presented above, one might, for example, wonder whether gender has an influence here. Table 40 shows the distribution of gender per chemistry level and a Chi-Square test according to Pearson ($\chi^2(2) = 4.33, p > .05$) reveals that the selection of chemistry level in upper secondary does not differ significantly based on the students' gender.

Table 40: Distribution of chemistry level in upper secondary education by gender and corresponding chi-square test

Chemistry level in upper secondary education * Gender Crosstabulation				
Numbers				
		Gender		total
		female	male	
Chemistry level in upper secondary education	none	28	9	37
	Basic level	92	51	143
	Advanced level	81	60	141
Total		201	120	321

Chi-Square-Tests			
	Value	df	significance (two-tier)
Pearson Chi-Square	4.486 ^a	2	.106
Likelihood-Ratio	4.638	2	.098
Linear-by-Linear Association	4.336	1	.037
Number of valid cases	321		

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 13.83.

Even though gender and chemistry course selection at upper secondary do not significantly influence each other, other covariates might still do that. Therefore it is necessary to look at the combination of possible factors.

Joint consideration of variables

In order to find out which combination of factors significantly contributes to an explanation of variance of both knowledge dimensions, a multivariate linear model was calculated for each dimension. In the stepwise calculation, the three models represent the person variable, the inclusion of university variables as that is the main focus of test construction and assessment, and the inclusion of school variables, respectively. While all variables that had shown a significant individual effect on CK or PCK outcome (cf. Tables 37 to 39) went into the calculations⁴⁴, only those remaining significant in the models are reported in Table 41.

⁴⁴ Non-dichotomous items were first dummy coded.

Table 41: Regression analysis regarding success in SPR-CK test

	Model 1			Model 2			Model 3		
	<i>B</i>	<i>SE B</i>	<i>β</i>	<i>B</i>	<i>SE B</i>	<i>β</i>	<i>B</i>	<i>SE B</i>	<i>β</i>
Constant	-0.18	0.07		-0.01	0.07		1.53	0.20	
Gender	0.41	0.11	.22***	0.35	0.11	.19**	0.35	0.10	.19***
OC exam(s) passed vs. visited but exam(s) failed				-0.75	0.24	-.18**	-0.60	0.21	-.15**
OC exam(s) passed vs. visited but no exam(s) taken				-0.64	0.18	-.21***	-0.47	0.16	-.15**
OC exam(s) passed vs. not visited				-0.68	0.18	-.21***	-0.48	0.16	-.15**
Grade in school leaving certificate							-0.48	0.11	-.27***
Last school grade in chemistry							-0.21	0.07	-.18**
Advanced level vs. basic level in upper secondary							-0.23	0.10	-.13*
Advanced level vs. no chemistry in upper secondary							-0.47	0.16	-.17**
<i>R</i> ²	.05			.15			.35		

Note:*** $p < .001$, ** $p < .005$, * $p < .05$.

Explaining 34.6% of the total variance of students' CK achievement, the combination of variables in model 3 contributes most to the prediction of students' performance (Table 41). It includes variables or factors of all three categories, namely: i) gender, ii) successfully passing the exam(s) in organic chemistry, iii) grade in school leaving certificate, iv) last school grade in chemistry, v) having studied chemistry in upper secondary.

As for CK, a stepwise multivariate regression was performed to find out which combination of factors could explain the most variance in the PCK test. Again, the three models represent the person variable, the inclusion of university variables as that is the main focus of test construction and assessment, and the inclusion of school variables, respectively.

Table 42: Regression analysis regarding success in SPR-PCK test

	Model 1			Model 2			Model 3		
	<i>B</i>	<i>SE B</i>	<i>β</i>	<i>B</i>	<i>SE B</i>	<i>β</i>	<i>B</i>	<i>SE B</i>	<i>β</i>
Constant	-0.20	0.07		-0.53	0.20		0.63	0.28	
Gender	0.41	0.11	.21***	0.38	0.10	.19***	0.40	0.10	.20***
Exam(s) on how to teach chemistry passed vs. visited but exam(s) failed				-0.28	0.13	-.12*	-0.30	0.12	-.13*
OC exam(s) passed vs. visited but no exam(s) taken				-0.53	0.18	-.16**	-0.45	0.17	-.14*
Teaching program (BA/MA vs. Staatsexamen)				-0.23	0.10	-.12*	-0.26	0.10	-.13*
Progress in studies				.27	.08	.19**	.26	.08	.19**
Grade in school leaving certificate							-0.52	0.09	-.28***
<i>R</i> ²	.04			.15			.23		

Note:*** $p < .001$, ** $p < .005$, * $p < .05$.

Explaining 23.0% of the total variance of students' PCK achievement, the combination of variables in model 3 contributes most to the prediction of students' performance (Table 42). It includes variables or factors of all three categories, namely: i) gender, ii) successfully passing the exam(s) in teaching chemistry, iii) successfully passing the exam(s) in organic chemistry, iv) teaching program, v) progress in studies, vi) grade in school leaving certificate.

6.4.4 Dimensional analysis of all chemistry items from the *KiL*-project: Chemical Representations (CRep) and SPR

As already suggested by the learning progression on molecular structure and properties by COOPER ET AL. (2012b), at this point it makes sense to combine the data obtained on pre-service teachers' knowledge regarding SPR with that regarding chemical representations (CRep). Therefore, for the remaining analyses, the data presented in this thesis so far shall be combined with the data obtained in and presented in the adjacent thesis on CRep (cf. TASKIN, 2014). In order to analyze

discriminant validity of all chemistry-related measures, a multidimensional Rasch analysis was applied to examine the empirical separability of these measures. Since the individual analyses on both SPR knowledge (cf. chapter 6.4.2) and CRep knowledge (TASKIN, 2014) showed CK and PCK as separate dimensions, a four-dimensional model was fitted first to the joined data. The four-dimensional model assumes (1) content knowledge regarding CRep, (2) content knowledge regarding SPR, (3) pedagogical content knowledge regarding CRep, and (4) pedagogical content knowledge regarding SPR. This model was compared to the corresponding fit values of a one-, and a two-dimensional model as alternative models. The one-dimensional model assumes a single latent trait behind teachers' responses whereas the two-dimensional model contains one dimension for content knowledge as well as one dimension for pedagogical content knowledge. Table 43 shows that the four-dimensional model fits the data best. As the four-dimensional model is significantly better than the two-dimensional model ($\chi^2 = 114.44$, $p < .001$), this suggests that CK_{CRep} , CK_{SPR} , PCK_{CRep} and PCK_{SPR} indeed represent different constructs.

Table 43: Model fit analysis, CRep and SPR dimensions

	AIC	BIC
1 dimension (chemistry knowledge)	31427.90	31861.97
2 dimensions (CK vs. PCK)	31242.39	31684.01
4 dimensions (CK_{CRep} vs. CK_{SPR} vs. PCK_{CRep} vs. PCK_{SPR})	31141.95	31609.99

Note. AIC = Akaike information criterion; BIC = Bayesian Information Criterion.

All four dimensions show sufficient reliability values (EAP/PV for $CK_{CRep} = .84$; for $CK_{SPR} = .87$; for $PCK_{CRep} = .75$; for $PCK_{SPR} = .76$). Even though the correlation between the CK dimensions is quite high ($\psi = .87$), this had been expected due to the theoretical design of these dimensions (one focusing only on various structures and the other focusing on the relationship between structures and observable properties). The PCK dimensions correlate to a lesser extent ($\psi = .67$).

In order to examine the relation between the difficulty of items and the ability of students, a Wright Map is generated (Figure 42, here each 'x' represents 3 cases⁴⁵).

⁴⁵ For an overview which item number in the Wright Map corresponds to which item in the test, see Annex III.

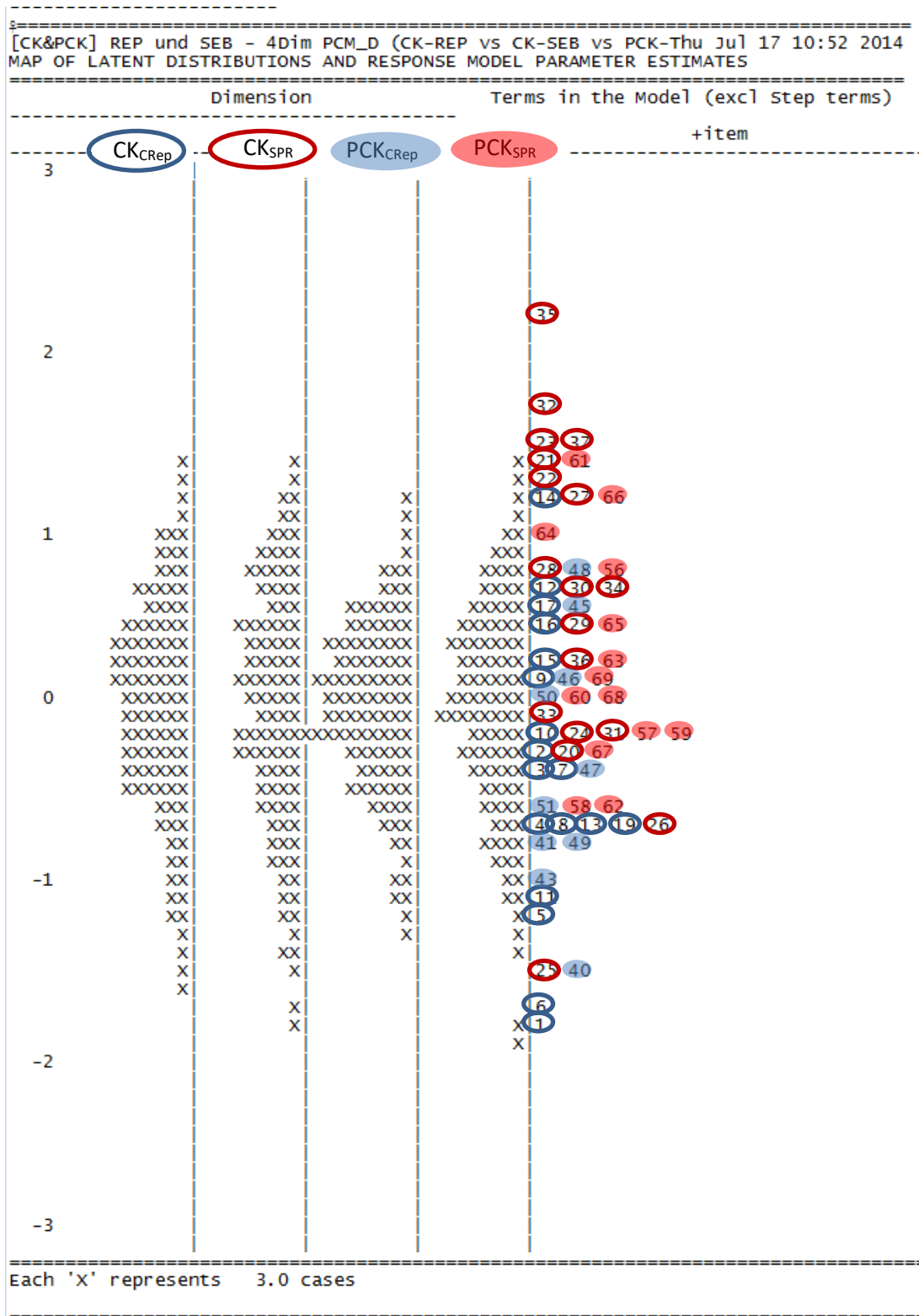


Figure 42: Wright Map of all four chemistry dimensions

The Wright Map shows that the person abilities for all four test instruments are distributed rather symmetrically, and that the corresponding values for item difficulties are distributed along the scale with a similar range as the person abilities.

When looking at the mean of item difficulty and standard deviation (cf. Table 44), it becomes apparent that there is a high overlap between the distributions of the item and person parameters in general. A difference of up to 0.88 logits exists between the respective means.

Table 44: Overview of means for all four dimensions

	Mean item difficulty	<i>SD</i>
CK _{CRep}	-.38	.79
CK _{SPR}	.88	1.05
PCK _{CRep}	-.37	.55
PCK _{SPR}	.31	.65

A Wright Map also allows insights into the range of item difficulties. Figure 42 shows that the CRep items (both CK and PCK) seem to be more present in the lower half of the map while the SPR items seem to be more frequent in the upper half of the map. This points to the fact that the CRep test instruments are easier to solve correctly for the sampled pre-service teachers than the SPR test instruments.

Figure 43 shows the mean test results for all four dimensions for the largest subgroup of the chemistry cohort, i.e. students striving for a degree that would later enable them to teach at academic-track schools in Germany, i.e. the Gymnasium.

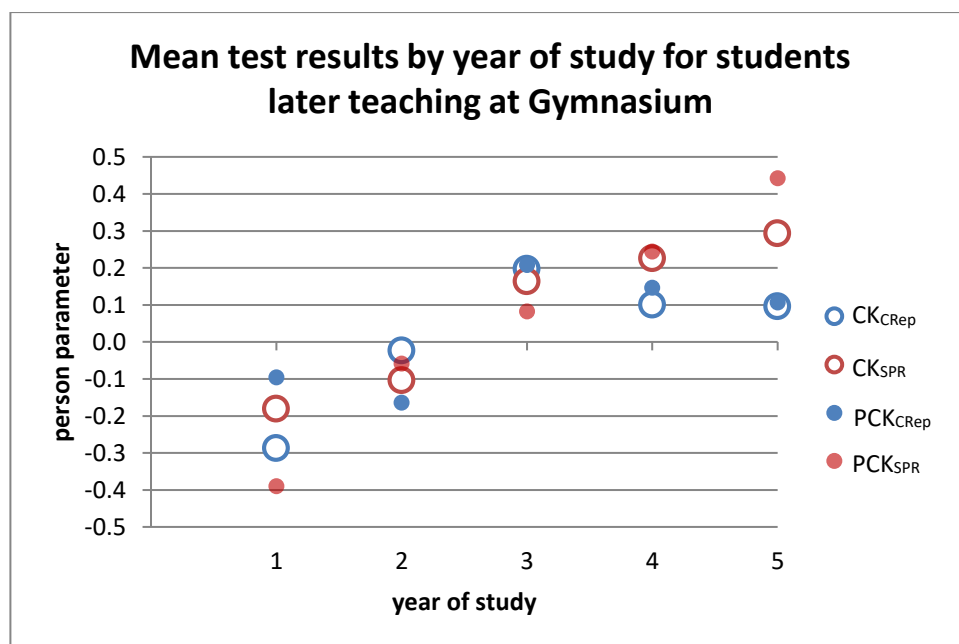


Figure 43: Test results regarding the students' year of their studies

Even though it needs to be noted that the data are derived from a cross-sectional study, Figure 43 shows that there seems to be a continuing increase through the students' university education in regard to both SPR-related dimensions, while there is stagnation or even decline for the CRep-related dimensions after year three. For the whole cohort, i.e. all participants, the correlations with the students' chemistry semester are small ($r(\text{CK}_{\text{CRep}}) = .22$, $r(\text{CK}_{\text{SPR}}) = .21$, $r(\text{PCK}_{\text{CRep}}) = .09$, $r(\text{PCK}_{\text{SPR}}) = .26$, with $p < .05$ for all but PCK_{CRep}), yet when judging the overall effect of this knowledge gain over the course of the university education (Figure 44), there is a significant increase in mean test outcome in three of the four dimensions (CK_{CRep} : $t(86) = -3.01$, $p < .05$; CK_{SPR} : $t(86) = -3.19$, $p < .05$; PCK_{CRep} : $t(86) = -.49$, $p > .05$; PCK_{SPR} : $t(86) = -4.81$, $p < .05$) from students who are at the beginning of their studies (semester one and two) to those at the end of their studies (semester nine and above). For the three significant increases, the effect size of these gains is medium to large ($d^{\text{A6}}(\text{CK}_{\text{CRep}}) = 0.68$; $d(\text{CK}_{\text{SPR}}) = 0.72$; and $d(\text{PCK}_{\text{SPR}}) = 1.08$).

⁴⁶ d being the effect size for a t -test, according to COHEN (1992) has the following cutoff values: .20/.50/.80 (small/medium/large)

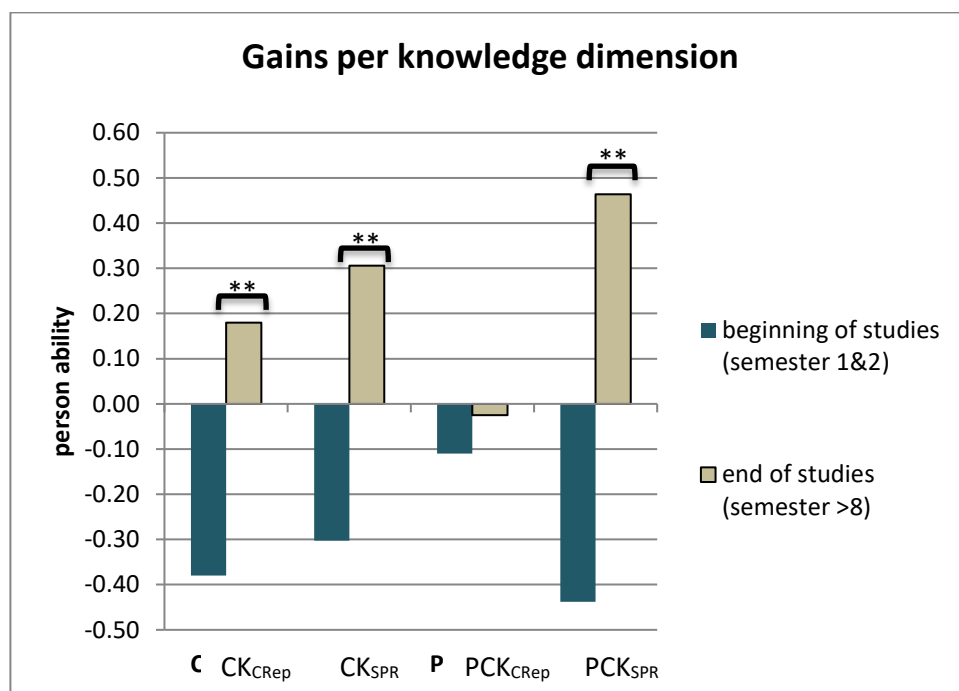


Figure 44: Knowledge development for each dimension from beginning to end of studies

Considering the overall structure of pre-service teachers' professional knowledge, i.e. combining the pedagogical knowledge also assessed in the *KiL*-project (cf. chapter 3) with the data presented above, the following model for chemistry pre-service teachers' professional knowledge evolves (Figure 45).

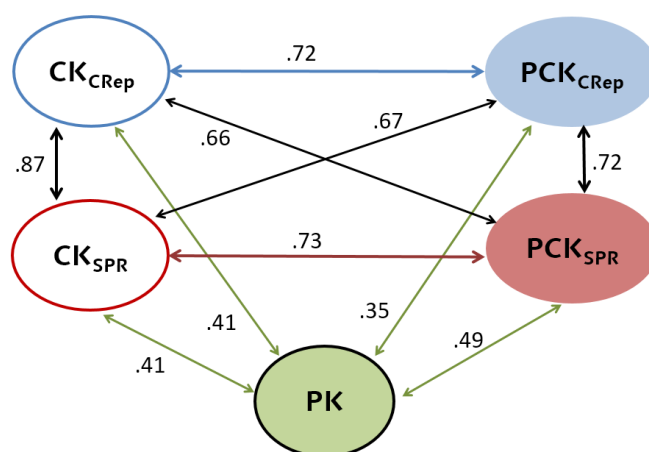


Figure 45: Structure of professional knowledge of pre-service chemistry teachers

Figure 45 shows the latent correlations between the dimensions: The largest correlations are found between the content knowledge dimensions and between the pedagogical content knowledge dimensions across the chemistry areas, i.e. CK_{CRep} with CK_{SPR} and PCK_{CRep} with PCK_{SPR} , respectively. The dimensions within each chemistry area also show similarly large correlations as well, i.e. CK_{CRep} with PCK_{CRep} and CK_{SPR} with PCK_{SPR} . The correlations across dimensions of different chemistry areas are smaller but still large, i.e. CK_{CRep} with PCK_{SPR} and PCK_{CRep} with CK_{SPR} , while the smallest correlations can be found between all chemistry-related dimensions and the pedagogical knowledge dimension.

The correlation patterns of the four chemistry dimensions with other constructs differ (Table 45); adding further strength to the fact that they are indeed separate dimensions.

Table 45: Correlations of the person parameters of the four content dimensions with other constructs

Instrument	CK _{CRep}	CK _{SPR}	PCK _{CRep}	PCK _{SPR}
Domain-specific INTEREST (<i>SIQ</i>), 6 items each				
• chemical content; $\alpha = .80$.10	.19**	.07	.18**
• teaching chemistry; $\alpha = .84$	-.13*	-.11	.00	.07
• pedagogy; $\alpha = .90$	-.18*	-.22**	-.06	-.12*
Domain-specific SELF-CONCEPT (<i>BEvaKomp</i>), 15 items each				
• chemical content; $\alpha = .87$.26**	.35**	.18**	.29**
• teaching chemistry; $\alpha = .91$.03	.00	.12*	.17**
• pedagogy; $\alpha = .89$	-.09	-.10	.03	-.05
Domain-specific BELIEFS				
Self-efficacy regarding experiments in the classroom (RABE); 9 items; $\alpha = .92$.08	.06	.18**	.13*
Self-efficacy regarding teaching activities (TSCHANNEN-MORAN & WOOLFOLK HOY); 4 items; $\alpha = .74$.13*	.22**	.19**	.19**
Nature of Science (<i>SUSS</i>); 23 items; $\alpha = .83$.30**	.26**	.23**	.35**

** $p < .01$, * $p < .05$

Regarding interest, both SPR-related dimensions correlate significantly with *interest in chemical content* while the CRep-related dimensions do not. Only the CK_{CRep} dimension shows a significant but negative correlation with *interest in teaching chemistry* while the CK_{SPR} and both PCK dimensions do not. Both SPR-related dimensions as well as the CK_{CRep} dimension show significant but negative correlations with *interest in pedagogy*.

Regarding the domain-specific self-concept, all four dimensions correlate significantly with *self-concept regarding chemical content*. Except for PCK_{SPR}, the correlations with the CK dimensions are higher. Both PCK dimensions correlate significantly with *self-concept regarding the teaching chemistry*. No knowledge dimension correlates significantly with *self-concept in pedagogy*.

Regarding the beliefs, only the PCK dimensions correlate significantly with *self-efficacy regarding experiments in the classroom*. All four dimensions correlate significantly with *self-efficacy regarding teaching activities* and *Nature of Science*.

6.5 Summary and discussion of findings

The results of study II confirm the finding from study I that both CK and PCK instruments on SPR are reliable, showing sufficient reliability values. Furthermore, the dimensions are found to be separable, emphasizing that CK and PCK are indeed two separate knowledge dimensions (cf. chapter 6.4.2). The latent correlation ($\psi = .72$) shows that both dimensions are closely related; the value being comparable to those reported for in-service math teachers in *COACTIV* ($r_{\text{latent}} = .79$, KRAUSS ET AL., 2013), pre-service math teachers in *MT21* ($r_{\text{latent}} = .79$, BLÖMEKE ET AL., 2010), pre-service physics teachers ($r_{\text{latent}} = .68$, RIESE & REINHOLD, 2012), or pre-service biology teachers ($r_{\text{latent}} = .68$, GROSSSCHEDL, HARMS, KLEICKMANN & GLOWINSKI, 2015) while being larger than those reported for the different sciences in *ProwiN* (KIRSCHNER ET AL., 2016). TASKIN also reports a large correlation between the subject-specific knowledge dimensions ($r = .50$, TASKIN, 2014). Therefore, the obtained result confirms the hypothesis for research question Q1. As can be seen from the Wright Map, for both SPR dimensions item difficulties match participants' abilities to a large degree; both instruments are slightly too difficult for the population but are suitable for assessing pre-service chemistry teachers' SPR knowledge, nonetheless. Overall it can be said that two paper-and-pencil tests regarding SPR have been successfully developed⁴⁷: I) a CK instrument containing 19 closed and open items and assessing pre-service chemistry teachers' knowledge of recognizing aspects regarding SPR as well as applying that knowledge in various contexts; and II) a PCK instrument consisting of 12 open items and measuring pre-service chemistry teachers' knowledge of potential problems their future students might have, of instructional strategies regarding the teaching of SPR phenomena and of selecting appropriate experiments and models.

In order to answer research question Q2, several variables – being differentiated into person variables, school variables and university variables – and their individual as well as joint effects on the test outcomes have been looked at (cf. chapter 6.4.3). If considered independently, many factors have an effect on the CK test outcome: the person variable *gender* (i.e. men achieved better results than women), the school

⁴⁷ The final test instruments can be found in Appendix III.

variables *GPA* (i.e. participants with better GPA achieved better results), *chemistry level in upper secondary* (i.e. pre-service teachers who had taken chemistry at upper secondary performed better than those without chemistry in upper secondary; and participants with at an advanced course outperformed those with only a basic course level), and *last chemistry grade* (i.e. students with good grades in their last secondary chemistry course achieved better results than those with poorer grades), as well as the university variables *year of studies* (i.e. students further along in their studies outperformed those closer to the beginning), *teacher track program* (i.e. participants enrolled in a program that enables them to later teach at the Gymnasium showed better performance than those aiming to teach at other school types), as well as the learning opportunities in *inorganic* and *organic chemistry* (i.e. for both factors, participants who had passed exams in that field outperformed those that had not yet taken any exams). Several of these factors have been found by other researchers as well. As an individual consideration of variables does not provide any insights into possibly confounding factors, a multivariate stepwise regression was calculated yielding only those factors that explain variance in test outcome when considered jointly. For the CK instrument, the five remaining factors are i) *GPA*, ii) *gender*, iii) *last school grade in chemistry*, iv) *course level at upper secondary*, and v) *successful exams in organic chemistry*. Of these five factors, many have been found by other researchers as factors influencing test outcome in their tests assessing content as well: GPA as a commonly considered indicator of student intelligence and predictor for study success has also been identified by FREYER (2013) in her test instrument assessing university chemistry knowledge, by RIESE AND REINHOLD (2012) in their CK dimension assessing pre-service physics teachers, and by GROSSCHEDL ET AL. (2015) in their test on university biology CK; stressing that the higher a student's GPA, the better his/her test result. In her test on beginning university students' chemistry concept knowledge, BUSKER (2010) showed that course level at upper secondary was a factor in students' success in her test. The same was found by RIESE AND REINHOLD (2012). Gender was also identified as a significant predictor by both TASKIN (2014) and RIESE AND REINHOLD (2012). As the items in this CK test on SPR knowledge include the depiction of compounds in two-dimensional drawings, one could safely say that the test requires the ability to mentally rotate and manipulate spatial information relating to the structure of a compound. Numerous studies have found

men to be better at tests involving spatial ability and provided various explanations for this difference (for a selection of literature see FISHER, MEREDITH & GRAY, 2017). Particularly in regard to organic chemistry, the group around BODNER presented significant correlations between spatial ability and performance in general and organic chemistry with items focusing on higher order cognitive skills such as problem-solving skills and mentally manipulating structures (cf. among others BODNER & McMILLAN, 1986; CARTER, LARUSSA & BODNER, 1987; PRIBYL & BODNER, 1987). FISHER ET AL. (2017) pointed out that it might not be the spatial ability itself that causes the differences but rather the way in which spatial information was presented to participants: when changing the stimulus from two-dimensional drawing to a photograph to a three-dimensional model that is only grasped visually to a three-dimensional touchable model, they found the gender difference to disappear. Based on these aspects, it is not surprising that the gender difference could be measured regarding the SPR test. According to FISHER ET AL. (2017) it would be expected that this factor disappears if all structures were to be presented as touchable models. As that is impossible for a paper-and-pencil test, one might ponder the idea of switching to a digital test administration including simulations of the structures in 3D.

Interestingly enough, students' academic track could not be found as a predictor for CK outcome in this study. While this finding replicates TASKIN's (2014) analysis, it is contrary to findings in many other research contexts (such as GROSSCHEDL ET AL., 2015; KLEICKMANN ET AL., 2013; KRAUSS ET AL., 2008; and RIESE & REINHOLD, 2012). One could speculate that the fundamentals of SPR thinking and much of the relevant learning opportunities might, therefore, be laid in the beginning of university education toward the middle, up to which point there is hardly any difference in content knowledge exposure between pre-service chemistry teachers studying for academic and non-academic tracks.

When looking at the amount of explained variance in test performance, the 34.6% for the CK instrument on SPR is comparable to numbers presented by RIESE AND REINHOLD (2012) who found that they could explain nearly 40% through gender, course level at upper secondary, number of university semester hours in the subject as well as subject-specific self-efficacy. Similarly, TASKIN (2014) reported that for the CK dimension regarding chemical representation, about 30% of the total variance of

students' achievement could be explained through the following factors: i) grade of school leaving certificate, ii) successfully passing the exam(s) in organic chemistry, iii) gender, iv) having studied chemistry in upper secondary at a basic and v) at an advanced level. Thus both the factors determined as well as the amount of variance explained is in line with other researchers' findings. The fact that TASKIN's (2014) finding are very similar to the findings of this study is not surprising as the CK instrument on CRep and SPR have been developed in close communication, in the research context of the *KiL*-project and administered to the identical cohort. As TASKIN (2014) reported, the majority of CK items in her instrument on chemical representations focused on organic chemistry and thus an identification of the successful completion of organic chemistry exams as a predictor for student success was not surprising – the same can be said for the CK instrument on SPR.

Regarding PCK, similar factors as in the CK analysis could be found to have an effect when considered independently: the person variable *gender*, the school variables *GPA*, *having had any chemistry course in upper secondary*, and *last chemistry grade*, as well as the university variables *year of studies*, *teacher track program*, as well as the *learning opportunities in inorganic, organic and the teaching of chemistry*. As with CK, a multivariate stepwise regression yielded a reduced number of factors that explain variance in PCK test outcome when considered jointly: i) *GPA*, ii) *gender*, iii) *progress in course of study*, iv) *successful exams in organic chemistry*, v) *successful exams in teaching chemistry*, and vi) *university teaching program*. Again, several of these factors have also been identified by other researchers. RIESE AND REINHOLD (2012) found GPA to be a predictor for PCK as well, so did TASKIN (2014) and GROSSCHEDL ET AL. (2015).

As with CK, students' academic track could not be found as a predictor for PCK outcome either even though it is described as a factor by GROSSCHEDL ET AL. (2015), KLEICKMANN ET AL. (2013), RIESE AND REINHOLD (2012), or WITNER AND TEPNER (2011). Again as with CK, a similar interpretation might hold: there is hardly any difference in pedagogical content knowledge exposure at university between pre-service chemistry teachers studying for academic and non-academic tracks up to the Master level.

With the above-mentioned six factors, the size of the explained variance for the PCK outcome regarding SPR is 23%. This is considerably larger than that reported by TASKIN (2014) or RIESE AND REINHOLD (2012). The former could identify the combination of i) grade of school leaving certificate, ii) having studied chemistry in upper secondary at an advanced level, iii) successfully passing the exam(s) in teaching chemistry to explain about 12% of variance (TASKIN, 2014). The latter could explain about 15 % of variance with the following factors: number of semester hours in the subject and in teaching the subject, gender and GPA (RIESE & REINHOLD, 2012).

Even though general pedagogical knowledge (PK) and SPR-specific CK might increase the explained variance in the PCK_{SPR} outcome, it is not possible at this point to add these data into the equation because that data have been obtained simultaneously. Analyses regarding the effect of these factors can only be evaluated in a longitudinal test design.

As the factors obtained have an effect on CK and PCK outcomes are similar to those found by other researchers, these findings are aspects for construct validity.

Regarding a possible development of pre-service chemistry teachers' SPR knowledge (research question Q3), the regression analysis reveals that progress in the course of study is a predictor only for PCK. The lack of this factor as a predictor for CK might be due to the fact that the majority of CK items in the instrument require only a sound understanding of knowledge usually acquired from upper secondary up to the middle of tertiary education. Therefore specialized knowledge possibly acquired towards the very end of a teaching program in chemistry is not needed. Furthermore, the teaching of content to pre-service teachers is concentrated more strongly in the Bachelor-part, thus from the beginning toward the middle of the university program, while the teaching of PCK rather begins in the middle of the university program and is continued towards the end. This might explain why it is a factor in the joint consideration of variables for PCK success.

When combining the instruments regarding SPR and chemical representations (research question Q4), a four-dimensional structure of the subject-specific professional knowledge is found. All dimensions show sufficient reliability values and the correlations between the dimensions per content area show a pattern that had been expected based on the design of the items: the CK dimensions correlate to a much higher extent than the PCK dimensions as they had been developed in close agreement (cf. chapter 4.1). A Wright Map of all four dimensions shows that the items regarding chemical representations tend to be easier while those on structure-property-relations appear to be more difficult. Taken together, however, the difficulties of the items match students' abilities and thus the combination of test instruments can be considered appropriate for the sample.

The correlations of the four dimensions with interest, self-concept and beliefs provide additional aspects for construct validity: The correlations with self-concept are in accordance with expectations: chemical content self-concept and teaching chemistry self-concept coincide with higher performance in the chemistry-related knowledge dimensions while no such relation could be found between pedagogy self-concept and performance in the chemistry-related knowledge dimensions. Students who reach higher outcomes in the tests also have higher chemistry-related self-concept while their test outcomes are independent of their pedagogical self-concept. The fact that both SPR-related dimensions correlate significantly positive with the interest in chemical content – while the CRep-related dimensions do not – might be explained by the fact that all SPR items refer to some macroscopic phenomenon while the CRep items remain at the representational level. As dealings with the real world tend to increase interest in most subjects, it seems reasonable that these dimensions correlate with higher test outcomes.

The comparison of mean test results for all four dimensions shows that there is an increase over the progress of the course of study most notably for the SPR dimensions, with the PCK_{SPR} dimension showing a steeper increase. The decline in CRep outcome for both CK and PCK for students in their final part of the teaching program might be explained as follows: Students rarely cover more in-depth content courses after year three during which they would have to use or apply their knowledge of chemical representations. This apparent lack of learning opportunities

is not as true for the other content knowledge construct as it focuses more on connections between the structures and properties and thus becomes relevant implicitly in many chemistry-related modules. The relatively high level of pre-service teachers' PCK_{CRep} at the onset of their university education might be explained by VAN DRIEL (2010), who argues that every future teacher already has a certain ability regarding PCK from his/her own experience as a student at school, thus strengthening the hypothesis that there is some experience gained as a by-product of secondary education regarding students' own or their classmates' difficulties, regarding representations and models used by the teacher, regarding experiments that could illustrate certain things and regarding a general structure of going from simple to more complex explanations – i.e. unstructured and non-integrated aspects of PCK (cf. Figure 1). Why this is noticed more dominantly for the PCK_{CRep} dimension than the PCK_{SPR} dimension is not as clear: An explanation could be that difficulties regarding chemical representations, which are likely to have been an explicit part of students' own experiences, might be recalled by pre-service teachers by simply seeing a chemical representation, while for recalling comparable aspects regarding structure-property-relations, the pre-service teacher will first have to assess two things: i) whether his/her experiences relate to this fundamental concept, and ii) that the task in front of him/her is actually asking SPR knowledge as well.

Considering the overall structure of pre-service chemistry teachers' professional knowledge (research question Q5), the correlation pattern between the five dimensions matches previous expectations: The largest correlations between the content knowledge dimensions and between the pedagogical content knowledge dimensions across the chemistry areas, i.e. CK_{CRep} with CK_{SPR} and PCK_{CRep} with PCK_{SPR} , respectively, might be explained due to the conceptualization of the original knowledge facets (cf. SHULMAN, 1987) and due to the teaching programs at university often being divided into content and teaching modules. An additional reason might be the communicative validation prior to CK and PCK item development: a common understanding of which aspects belonged to the knowledge sub-dimensions, which degree of specificity the items should contain, as well as an exchange on potential topics might have led to a more similar framework for item development. The

dimensions within each chemistry area also show similarly large correlations as well, i.e. CK_{CRep} with PCK_{CRep} and CK_{SPR} with PCK_{SPR} , which makes sense as they have been developed with the intention of containing each either CRep or SPR as common chemistry areas, a fact that had been demanded by researchers such as AUFSCHNAITER AND BLÖMEKE (2010), MULHALL ET AL. (2003), ROLLNICK ET AL. (2010), or WALPUSKI ET AL. (2012).

The correlations across dimensions of different chemistry areas are smaller but still large, i.e. CK_{CRep} with PCK_{SPR} and PCK_{CRep} with CK_{SPR} , suggesting that chemistry-related knowledge is connected even across fundamental concepts or competency areas. The smallest correlations can be found between all chemistry-related dimensions and the pedagogical knowledge dimension. With the described differences in correlation between the knowledge dimensions, the pattern is similar to that in other studies assessing the structure of professional knowledge (Figure 46):

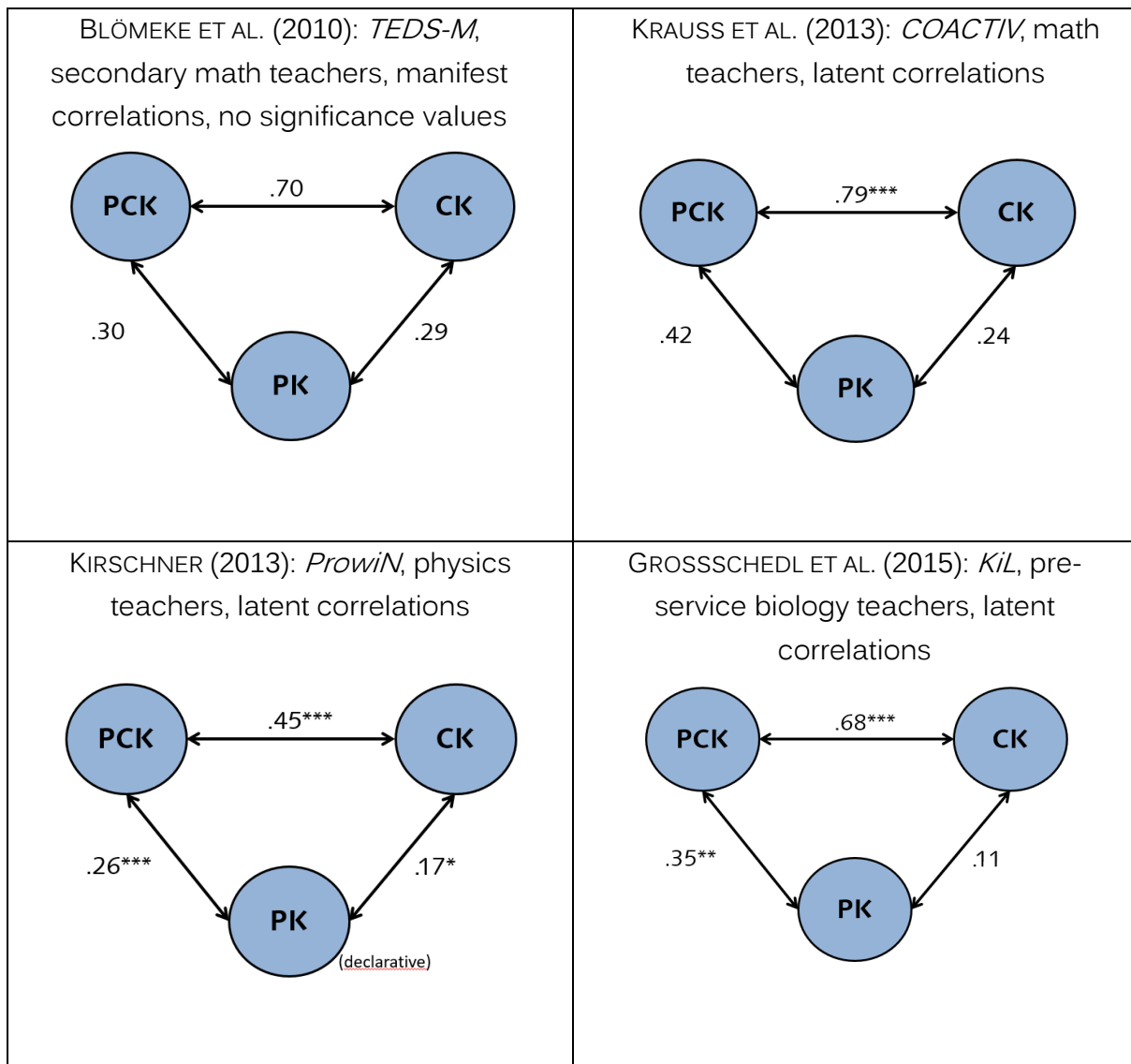


Figure 46: Structures of teachers' and pre-service teachers' professional knowledge

7 FINAL DISCUSSION AND IMPLICATIONS

Instead of repeating the results from both studies, a combined presentation of findings is undertaken at this point with the focus on the continuation of the *KiL*-project: the project *KeiLa* (acronym for the German project title: *Kompetenzentwicklung in mathematischen und naturwissenschaftlichen Lehramtsstudiengängen*, i.e. Development of competencies in mathematical and science-related teacher education programs). In that project, a true longitudinal testing is planned. With that intention, the project will allow for an analysis of knowledge development as a true process analysis that SCHMITZ (2006) still considers an exception in educational research. A few studies have attempted to assess professional knowledge at least twice in (pre-service) teachers' education in order to arrive at conclusions about a possible development: in *COACTIV-R*⁴⁸, (pre-service) math teachers are tested twice in their practical on-the-job phase (KUNTER, KLUSMANN, GENZ, BUSCHING & PETERS, 2009), while in *TEDS-FU*, (pre-service) math teachers provide knowledge data in their practical on-the-job phase and again four years later (BLÖMEKE ET AL., 2014). In *KeiLa*, pre-service math, chemistry, biology and physics teachers will fill in subject-specific knowledge tests at three to four points during their university education. Knowledge tests will be combined with questionnaires on learning opportunities at university to allow for a more detailed analysis of possible factors influencing professional knowledge development at the tertiary level. Through this combination of instruments, an in-depth look at the knowledge development is anticipated.

As a prerequisite for a developmental analysis of pre-service teachers' professional knowledge, subject-specific knowledge tests have to be developed. In the *KiL*-project, this could be achieved for the concept of structure-property-relations as well as for the combination of structure-property-relations and chemical representations. It has been statistically shown that CK and PCK instruments on SPR and CRep are separate dimensions which correlate to a large extent (study II). If one wishes to get a detailed account of pre-service chemistry teachers' professional

⁴⁸ The *R* in the project name stands for the German phrase *Referendariat* which is the practical teacher education phase after the more theoretical university education phase.

knowledge, the instruments used can, therefore, shed light into participants' knowledge regarding CRep and SPR and into the connection of both areas.

The added charm of the *KeiLa* project as opposed to assessing pre-service chemistry teachers' professional knowledge by itself will lie in a possible comparison between the science subjects regarding the development of the knowledge dimensions and influencing factors. While this has just been begun for the *KiL*-project as well (GROSSCHEDL, HARMS, HERZOG, MAHLER & STEFFENSKY, in prep.), comparing the actual development in the sciences will be very informative.

Even though several factors have been identified to influence the outcome on the subject-specific tests (cf. chapter 6.4.3) and even though content-related aspects in the form of learning opportunities at secondary and tertiary level have been identified, it could not be shown through this one-time assessment of participants' knowledge whether pre-service teachers' CK is truly a prerequisite for their PCK as is proposed by BALL, HILL AND BASS (2005), BAUMERT ET AL. (2010) or MA (1999), for example. As in the *KiL*-project the measures for CK and PCK regarding structure-property-relations had been newly developed instruments, using one to validate the other seemed rather presumptuous; in the following project with its cross-lagged-panel-design, however, this is possible. Similarly, it will be possible to find out if knowledge of chemical representations presents a prerequisite for structural-property-knowledge, as might be indicated by the Wright Maps. The design will also allow a detailed analysis of whether knowledge in the subject-specific dimensions develops in a continual process or whether results from study I are replicated. In which case, the differences in developments of the separate dimensions needs to be investigated further.

It is important to stress that the intention of the test has been to obtain purely diagnostic information. It is not meant to be used as a tool with which to make any predictions (as to how the knowledge will actually develop), selections or placement suggestions (as to which university teacher program or even courses to choose) for individual students. The instruments could, however, provide information when given to other cohorts, such as chemistry majors at university as they might have similar

problems in the beginning or even to students in advanced chemistry courses in upper secondary. In an additional study, this hypothesis should be tested. If the assumption holds true, secondary teachers or university lecturers using the instruments in their courses could draw inferences about the necessity of a possibly stronger focus on structure-property-thinking in their teaching.

As the PCK instruments have not been administered to students in their practical training or to teachers, the relevance of the items can only be inferred from a theoretical perspective as well as from communicational validation with selected teachers and university lecturers for the teaching of chemistry. That is why the instruments should be administered to practising teachers in an additional validation study to capture the complete knowledge development spanning from pre-service teachers at university to highly experienced teachers in the classroom.

The majority of the content knowledge items contained in the joint instrument focus on organic chemistry as this is a sub-discipline in which structure-property-thinking is immensely important. Even though this thinking is, of course, necessary for inorganic and physical chemistry as well, the decision to focus on organic chemistry has been made because there is a stepwise progression from simple to complex structures differing in their additive components of functional groups, which makes the development of conceptual understanding quite visible. Furthermore, it usually precedes physical chemistry but is normally not situated right at the beginning of chemistry programs at university as it requires some knowledge from inorganic chemistry. However, an inclusion of further items on physical chemistry would be necessary in order to allow conclusions for the whole realm of chemistry.

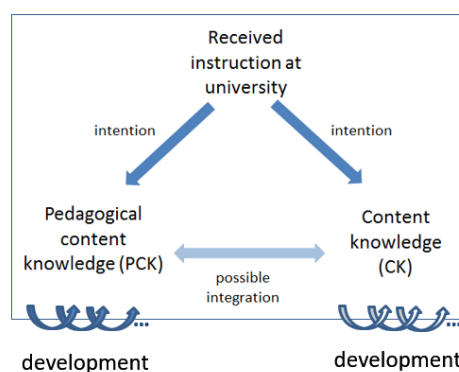


Figure 47: Knowledge acquisition at university

Even though many aspects could already be identified in this dissertation project, several questions remain that necessitate further research:

Data in this project shows that having prepared for organic chemistry exams is beneficial for PCK knowledge, while doing the same in chemistry teaching courses itself is not. This suggests that preparing for organic chemistry exams enhances pre-service teachers' understanding regarding possible learners' problems, difficulties of the topic, as well as approaches to explaining and visualizing it. While this aspect is plausible as it can indeed increase an understanding in PCK aspects, one might wonder why the same could not be found for preparing for chemistry teaching exams. Might it be that a preparation for chemistry teaching exams is less thorough or less demanding on the students? A longitudinal analysis incorporating pre-service teachers' use of learning opportunities might shed light into the question whether this aspect is an indication of differences in the way modules at university are perceived, in the modules' demands or in individual approaches to learning. It would also be interesting to investigate if this differentiation might be related to differences in participants' interest regarding the content of those modules, the teaching approaches used in those modules, the modules' main focus being macroscopic or submicroscopic, or perceived own benefit to a future teaching career.

Assumptions on the development of pre-service teachers' professional knowledge sometimes focus heavily on quantitative descriptions and statistical values. While these approaches are important and relevant in order to arrive at statements about how knowledge typically develops across a population, it is just as important to look at individual learning paths regarding knowledge acquisition (Figure 47). While this aspect is intended to be assessed in *KeiLa* through student questionnaires regarding their perceived learning opportunities, it seems necessary to accompany the quantitative approach with more qualitative approaches. One of those might focus on investigating the effects that an integration of particular learning opportunities into tertiary education such as those described in the following chapter might have on CK and PCK development regarding the SPR concept. In addition, analyses of university module descriptions and exemplary lectures might yield further insight into the intended as well as actual curriculum of teacher education and might provide a more thorough understanding of beneficial teaching approaches in university. One might wonder, for example, whether the development in students'

understanding could possibly be attributed to knowledge aspects being presented in a certain sequence or through more content- or competency-oriented courses.

Teaching and learning are complex activities and many of the relations in the process remain fuzzy. Even though HELMKE (2003) and others have proposed a model combining the concept of learning opportunities with uses of instruction that could be used to research teaching effectiveness (HELMKE, 2003), investigating all relations simultaneously in a single research project remains almost impossible. Assessing professional knowledge is a first approximation that has been undertaken in the *KiL* project, among others, which was complemented by measuring other aspects of professional competence such as beliefs, personality or motivation. Accompanying this with information on perceived as well as actual learning opportunities and (qualitative) investigations on how pre-service teachers make use of certain learning opportunities might lead to a deeper understanding of potential influencing parameters in generating good teachers.

8 ADDENDUM – A PRACTITIONER’S PERSPECTIVE

A possible implication of the findings of particularly study I, in which many of the problems and difficulties reported in connection with the SPR concept could be replicated with pre-service teachers' answers to exemplary items, seems to be the continued call for making the reasoning process for gaining conceptual understanding more transparent to students – regardless of educational level. This process includes two main aspects: i) repeatedly and explicitly integrating all perspectives from the chemical tetrahedron in a way that makes sense to the students, applying various methods and scaffolding measures, and ii) explicitly reasoning through problems so that the students can see how factual knowledge is combined and reintegrated to form a knowledge base less reliant on the retrieval of individual knowledge pieces but more on the creation of connections between knowledge pieces that enables the learner to form information networks in their minds. Both of these aspects have been demanded by other researchers as well, among those MAHAFFY (2006) as well as GILBERT AND TREAGUST (2009) for the former aspect, and TAAGEPERA AND NOORI (2000), MAEYER AND TALANQUER (2010), as well as BHATTACHARYYA AND BODNER (2005) for the latter aspect. As with most other abilities, this knowledge network needs own experience to form and re-form, so that learners need to be given the opportunity as well as need to be required to practice verbalizing their logical reasoning when answering tasks or problems. Through this, learners might be able to go beyond pseudo-conceptual thinking (cf. VINNER, 1997) and come to a fundamental understanding of science concepts.

As both secondary students and pre-service teachers show difficulty with the fundamental concept of SPR, both groups need to be presented with further appropriate learning opportunities that enable them to develop a systematic understanding of structure-property-relations. Apart from attempts by COOPER ET AL. (2012b) and NGAI ET AL. (2014), for example, proposing learning progressions for this or adjacent fundamental concepts, a different approach has been attempted in the EU-project *ESTABLISH* (acronym for *European Science and Technology in Action: Building Links with Industry, Schools and Home*)⁴⁹. Project members in eleven

⁴⁹ <http://www.establish-fp7.eu/>

European countries developed, trialed, and disseminated teaching materials with an inquiry-based approach. These units were designed to either address both single science subjects and integrated science subjects, included explicit connections to industry and research – thus asking questions about the development and optimization of products – and were designed for lower as well as upper secondary education. Selected contexts of these overarching science units included “Medical imaging”, “Forensic science”, and “Renewable energy”. The contributions by (or generated in cooperation with) the IPN group were all chemistry-related units:

- “Chitosan – fat burner?” (focusing on the production of and, in part, controversial use of chitosan made from crab shells as well as its ecological aspects),
- “Chemical Care” (focusing on forces and interactions between substances such as different household materials, cleaners and different types of dirt as well as functional materials), and
- “Exploring Holes” (focusing on gaps or spaces that are created due to a certain structural makeup; resulting in pores that are relevant in the sciences and everyday lives, from coffee filters to polymer films in dialysis or antimicrobial coatings).

All three units focus on structure-property-relations. They can be used to support SPR-learning for two separate target groups: secondary students who should be supported in interpreting structure-property-relations on various levels of abstraction, as well as pre-service teachers who can use (parts of) these units in their own first teaching experiences. As the units also focus on both scientific and product-oriented aspects, they could ultimately be used to foster students’ motivation to ask and answer research questions both from a scientific and an industrial perspective. This is realized through an inquiry-based learning approach with different aims and comparable paths (Table 46).

Table 46: Comparison of an inquiry-based approach with different aims and quite similar paths (HERZOG ET AL., 2015)

Perspective	Research/Science	Development/Industry
Aim	Gain information on how knowledge is developed in the sciences	Gain information on how to develop and/or improve a product
Path	From phenomena to - Research questions - Model-based hypotheses - Investigations - Evaluation and application	From phenomena to - Product questions - Model-based hypotheses - Prototypes and tests - Evaluation and further ideas

In the following paragraphs, it shall be described how these ESTABLISH units can foster a more sophisticated understanding for the secondary school students is described before moving toward a discussion of potential benefits for the second target group, the pre-service teachers.

Especially the idea that industrial and every-day products are designed to show very specific properties lends itself to an investigation of phenomena through the structure-property-lens. As reported by HERZOG, FINLAYSON, HICKMANN AND PARCHMANN (2015) as part of a teaching module developed in the ESTABLISH-project, students are to investigate that a product's properties result from certain features regarding i) the material itself, ii) the chemical compounds, and iii) the submicroscopic level. In order to explain these things, students use models, from simple ones explaining things on a meso level to more complex ones clarifying the submicroscopic situation.

Considering the topic of *plastic*, HERZOG ET AL. (2015) suggest that abstraction levels in a learning development be described as follows: the monomers of a polymer molecule are responsible for the molecule's degree of interlinkage, which in turn determines properties such as the ability to be molded and melted, thus resulting in the different uses of plastic products, e.g. plastic bottles used for transporting water.

For other products, the joint consideration of macroscopic as well as submicroscopic functional units (that could be assigned to the meso-level between material and molecules) is relevant. For clothing items these units could be fibers while for superabsorbers, for example, a proposed SPR developmental model referring to these abstraction levels could look like Figure 48.

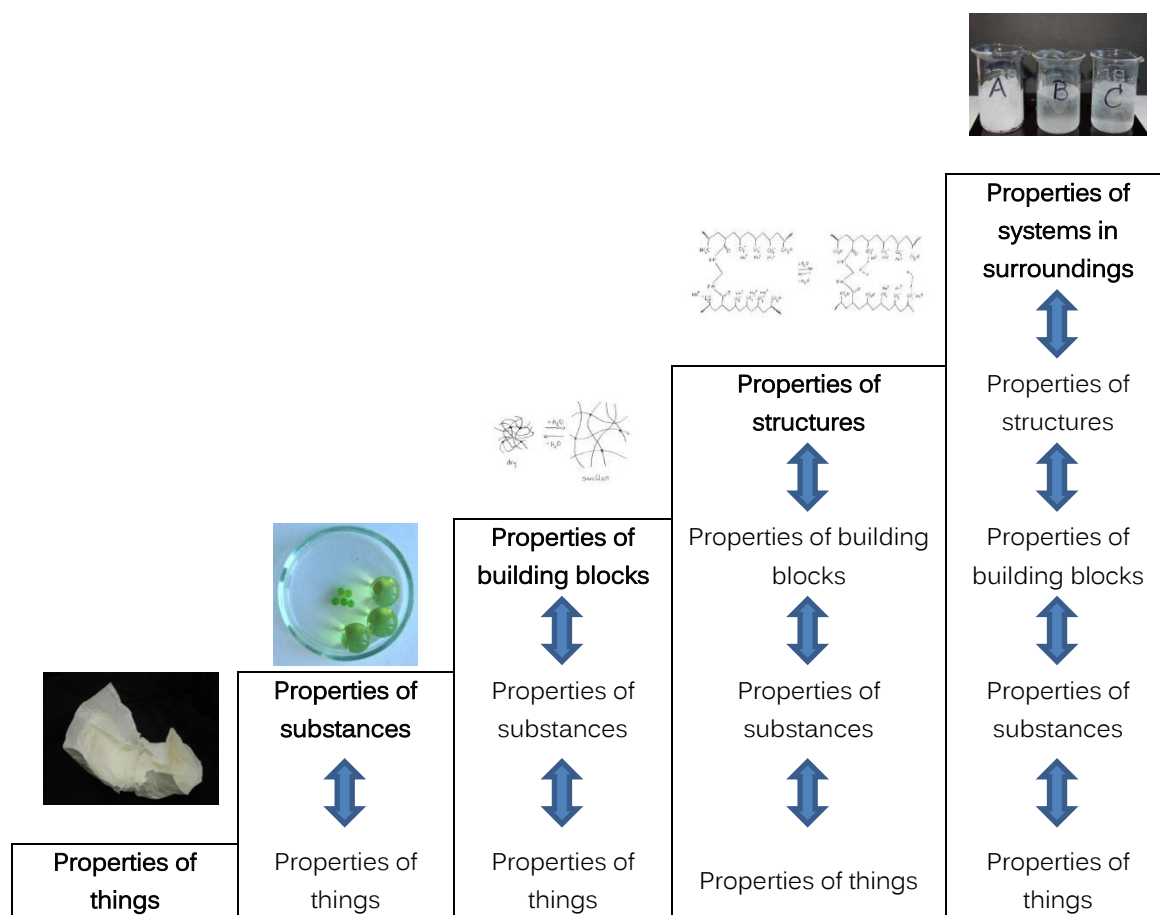


Figure 48: Developmental model for the SPR concept (HERZOG ET AL., 2015)

Working with products that students are introduced to in the ESTABLISH unit “Exploring Holes”, for example, is intended to improve the learners’ understanding of permeability as well as to guide their attention toward the importance of spaces within macromolecules. Through this, the unit aims to reduce some well-known student misconceptions: students i) usually assign macroscopic properties to individual atoms (cf. STEFFENSKY ET AL., 2005) or ii) have a problem with conceptualizing that spaces between particles are actually empty (cf. BARKE ET AL., 2009; or NOVICK & NUSSBAUM, 1978). HERZOG ET AL. (2015) point out that it is exactly this empty space that is relevant for allowing many macroscopic properties. These spaces, also called pores, can then be used for two main purposes: either to let selective substances pass through them or to retain certain substances. In both purposes, the functionality is approached via everyday products: holes in clothing

items for the former and baby diapers for the latter purpose. Both purposes are described in HERZOG ET AL. (2015) as follows:

Holes in clothes – functional spaces to let substances pass through

When asked if they know of clothing items with holes in them, students might think of plastic shoes for gardening that have visible holes (such as Crocs™, for example). At the abstraction level “properties of things” it is easy for students to hypothesize about the function of those holes: air and humidity can easily move from the inside to the outside, while you can just as easily get dirty feet in the sand and wet feet when it is raining. As the latter is a property that might not be so desirable, students almost intuitively raise the question how certain properties can be maintained while others are eliminated, thus arriving at the wish to assign particular properties to clothes: the size of the holes must be decreased to a level that prevents dirt and rain drops from passing through. In order to find out how small the holes must be, students then analyze different clothes items based on two things: their behavior towards permeability of air, steam, and water drops, as well as their pore sizes. The latter can be done via different optical devices such as cameras or light microscopes, showing students that pores are always present – it only depends on the device whether humans can make them visible. Investigating the former properties of different textiles, and thus moving toward the abstraction level “properties of substances” when differentiating between water in its different aggregation states, several experiments can be developed by students themselves and can in a subsequent step even be optimized.

In order to investigate a textile’s permeability of steam, for example, students may suggest or be asked to cover steaming water with a textile to find out if the steam condenses on the other side of the textile. A possible experimental setup might look like Figure 49 (boiling water is filled into a glass, the textile is stretched across the opening and held into place by a perforated lid, and a second glass is held upside-down over the perforated lid). Depending on the textile, students will either see water condensing inside the upper glass or not. With rain jackets, for example, it is especially motivating for students to see that sometimes it depends



Figure 49: Setup for testing steam permeability of a textile

on the “orientation” of the textile in the setup whether water vapor passes through (as most rain jackets have an inner and outer side that are intentionally designed to behave differently towards vapor permeability). In order to explore the permeability regarding water drops, students may suggest or be asked to stretch the textile over an empty glass, using a rubber band to hold it in place, adding a single water drop (or a small jet of water) to the surface of the textile and observing the water droplet spread. For more sophisticated analyses, all experiments can also be quantified in order to compare the properties of different textiles. The time of steam condensation can be measured and the contact angle between drop and textile can be determined⁵⁰.

For more advanced students, the mere phenomenological level can be accompanied by an analysis of hydrophilic and hydrophobic properties which are based on polar and non-polar parts of compounds and intermolecular forces, thus progressing to the abstraction level “properties of substances”. Differences between permeability regarding water drops and permeability regarding water vapor can be explained through the surface tension of water drops. By moving between the different abstraction levels and using them dependent on the focus in question, students develop a more thorough understanding of structure-property-relations. The same can be said for the other function of pores: retaining substances. This shall be described in the following section.

⁵⁰ <http://www.spektrum-neo.de/news/dem-lotuseffekt-auf-der-spur/1195578> (last accessed June 25, 2014).

Secrets of a diaper – functional spaces to retain substances

As with the visible holes in clothes, a product often known to students with smaller siblings is presented as a starting point: a diaper. When being shown that 20 mL of water are poured into a diaper and remain “caught” in it, the result is hardly surprising for many students because that is exactly what this product is made for (the difference between water and baby’s urine is usually mentioned by students and can be investigated at a subsequent step). When pouring a whole liter of water into the diaper without any water dripping out, students will, however, be surprised. They can then investigate the phenomenon according to the different abstraction levels:

- i) at the transition from “properties of things” to “properties of substances”, students can find out which substances in the diaper are responsible for water uptake and retention,
- ii) at the transition from “properties of substances” to “properties of building blocks”, students can look into simple models explaining the process of water retention,
- iii) at the transition from “properties of simple building blocks” to “properties of atomic structures”, students can use more sophisticated structural models to explain why water diffuses into the diaper and how water molecules attach to the polymer structure,
- iv) at the transition from “properties of atomic structures” to “properties of systems in surroundings”, students can investigate how the degree of water retention is influenced if different salt solutions are used (thus making the substances under investigation closer to its real-world counterpart, i.e. baby urine)

Depending on the level of the learners, this unit on diapers can therefore be used to stress either scientific methods (for younger learners), to repeat or integrate different concepts such as ion formation and ionic bonds, hydrate formation, osmosis, or functional groups in organic chemistry (for medium learners), or to introduce advanced learners to different aspects of polymer chemistry.

For the second target group, i.e. university students, the approach can also facilitate own structure-property-thinking both from a learner and a future teacher perspective. In tutorials or laboratory courses, students themselves explicitly reason through the increasingly complex levels of explaining a structure’s properties or investigate properties of substances and deduce their submicroscopic structure. This approach can be selected for courses regardless of the students’ university program. For pre-service teachers, integration in their laboratory courses makes sense as these courses should prepare the pre-service teachers for performing and discussing similar experiments in their future career. Thus they can be guided toward gauging the suggested experiments regarding their potential benefit or pitfalls for students of different abilities and hopefully transfer this knowledge to other experiments as well. Similarly, additional coverage of SPR in teaching chemistry courses seems necessary, thus introducing them to possible student misconceptions as well as problematic topics and to diverse options on how to deal with these, i.e. weighing specific instructional strategies, experiments and models. Therefore, the ESTABLISH material can be used here as well, but with a slightly different focus: pre-service teachers are to hypothesize why certain approaches (i.e. teaching aldehydes before alcohols, for example) or traditional practices (such as formulating net ion equations for simplification purposes) might contribute to student confusion. They should then argue how to approach the topic differently. The above-mentioned teaching units lend themselves to this idea because for a hypothetical teaching situation, pre-service teachers must think of several content prerequisites that should to be covered before the dealing with the unit content, as well as of possible student answers with varying levels of subject-specific language.

9 REFERENCES

- Akaike, H. (1981). Likelihood of a model and information criteria. *Journal of Econometrics*, *16*, 3-14.
- Andersson, B. (1986). The experiential gestalt of causation: a common core to pupils' preconceptions in science. *International Journal of Science Education*, *8*(2), 155-171.
- Arzi, H.J. & White, R.T. (2008). Change in Teachers' Knowledge of Subject Matter – A 17-Year Longitudinal Study. *Science Education*, *92*(2), 221-251.
- Aufschnaiter, C. V. & Blömeke, S. (2010). Professionelle Kompetenz von (angehenden) Lehrkräften erfassen – Desiderata [Assessing professional competencies of (pre-service) teachers – necessary research objectives]. *Zeitschrift für Didaktik der Naturwissenschaften*, *16*, 361-367.
- Ball, D. L., Hill, H. C. & Bass, H. (2005). Knowing mathematics for teaching: Who knows mathematics well enough to teach third grade, and how can we decide? *American Educator*, *29*(1), 14-17, 20-22, 43-46.
- Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching what makes it special? *Journal of teacher education*, *59*(5), 389-407.
- Banks, F., Leach, J. & Moon, B. (2005). Extract from new understandings of teachers' pedagogic knowledge. *Curriculum Journal*, *16*(3), 331-340.
- Barke, H.-D., Hazari, A. & Yitbarek, S. (2009). *Misconceptions in Chemistry: Addressing Perceptions in Chemical Education*. Heidelberg: Springer.
- Baumert, J. & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften [Keyword: Professional competence of teachers]. *Zeitschrift für Erziehungswissenschaft*, *9*(4), 469-520.
- Baumert, J. & Kunter, M. (2013). The COACTIV model of teachers' professional competence. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Eds). *Cognitive activation in the mathematics classroom and professional competence of teachers* (pp. 25-48). New York: Springer.
- Baumert, J., Kunter, M., Blum, W., Brunner, M., Voss, T., Jordan, A., Klusmann, U., Krauss, S., Neubrand, M. & Tsai, Y.-M. (2010). Teachers' mathematical knowledge, cognitive activation in the classroom, and student progress. *American Educational Research Journal*, *47*(1), 133-180.
- Bennett, J. & Lubben, F. (2006). Context-based chemistry – the Salters approach. *International Journal of Science Education*, *28*(9), 999-1016.
- Ben-Zvi, R., Eylon, B. & Silberstein, J. (1986). Is an Atom of Copper malleable? *Journal of Chemical Education*, *63*(1), 64-66.
- Bernholt, S. (2010). *Kompetenzmodellierung in der Chemie - Theoretische und empirische Reflexion am Beispiel des Modells hierarchischer Komplexität* [Modelling of competence in chemistry – theoretical

- and empirical reflection based on the example of the Model of Hierarchical Complexity]. Dissertation. Berlin: LOGOS.
- Bernholt, S. & Parchmann, I. (2011). Assessing the complexity of students' knowledge in chemistry. *Chemistry Education Research and Practice*, 12, 167-173.
- Bernholt, S., Walpuski, M., Sumfleth, E. & Parchmann, I. (2009). Kompetenzentwicklung im Chemieunterricht – Mit welchen Modellen lassen sich Kompetenzen und Aufgaben differenzieren? [Development of competence in chemistry education – Which models can be used to differentiate competences and items?]. *Naturwissenschaften im Unterricht Chemie*, 20(111/112), 78-85.
- Bhattacharyya, G. & Bodner, G. M. (2005). "It gets me to the product": how students propose organic mechanisms. *Journal of Chemical Education*, 82(9), 1402-1407.
- Blömeke, S., Kaiser, G., Döhrmann, M. & Lehmann, R. (2010). Mathematisches und mathematikdidaktisches Wissen angehender Sekundarstufen-I-Lehrkräfte im internationalen Vergleich [Mathematical content knowledge and pedagogical content knowledge of prospective math teachers for lower secondary in an international comparison]. In S. Blömeke, G. Kaiser & R. Lehmann (Eds.). *TEDS-M 2008 – Professionelle Kompetenz und Lerngelegenheiten angehender Mathematiklehrkräfte für die Sekundarstufe I im internationalen Vergleich* (pp. 197-238) [TEDS-M 2008 – Professional competence and learning opportunities of prospective math teachers for lower secondary in an international comparison]. Münster: Waxmann.
- Blömeke, S., Kaiser, G. & Lehmann, R. (Eds.) (2008). *Professionelle Kompetenz angehender Lehrerinnen und Lehrer. Wissen, Überzeugungen und Lerngelegenheiten deutscher Mathematikstudierender und -referendare – Erste Ergebnisse zur Wirksamkeit der Lehrerausbildung* [Professional competence of future teachers. Knowledge, beliefs and learning opportunities of German math students and students in their practical training phase after the university education – first results regarding teacher education effectiveness]. Münster: Waxmann.
- Blömeke, S. & König, J. (2010). Pädagogisches Wissen angehender Mathematiklehrkräfte für die Sekundarstufe I im internationalen Vergleich [Pedagogical knowledge of prospective math teachers for lower secondary in an international comparison]. In S. Blömeke, G. Kaiser & R. Lehmann (Eds.). *TEDS-M 2008 – Professionelle Kompetenz und Lerngelegenheiten angehender Mathematiklehrkräfte für die Sekundarstufe I im internationalen Vergleich* (pp. 265-278) [TEDS-M 2008 – Professional competence and learning opportunities of prospective math teachers for lower secondary in an international comparison]. Münster: Waxmann.
- Blömeke, S., König, J., Busse, A., Suhl, U., Benthien, J., Döhrmann, M. & Kaiser, G. (2014). Von der Lehrerausbildung in den Beruf – Fachbezogenes Wissen als Voraussetzung für Wahrnehmung, Interpretation und Handeln im Unterricht [From teacher education into the profession – subject-specific knowledge as prerequisite for perception, interpretation and acting in the classroom]. *Zeitschrift für Erziehungswissenschaft*, 17(3), 509-542.

-
- Blömeke, S., Suhl, U. & Kaiser, G. (2011). Teacher Education Effectiveness: Quality and Equity of Future Primary Teachers' Mathematics and Mathematics Pedagogical Content Knowledge. *Journal of Teacher Education*, 62(2), 154-171.
- Bodner, G. M. & McMillen, T. L. (1986). Cognitive restructuring as an early stage in problem solving. *Journal of Research in Science Teaching*, 23(8), 727-737.
- Borko, H. & Putnam, R. T. (1996). Learning to teach. In D. C. Berliner & R. C. Calfee (Eds.). *Handbook of Educational Psychology* (pp. 673-708). New York: Macmillan.
- Braun, E., Gusy, B., Leidner, B. & Hannover, B. (2008). Das Berliner Evaluationsinstrument für selbsteingeschätzte, studentische Kompetenzen (BEvaKomp) [The Berlin Evaluation instrument for self-evaluated student competences (BEvaKomp)]. *Diagnostica*, 54(1), 30-42.
- Bromme, R. (1997). Kompetenzen, Funktionen und unterrichtliches Handeln des Lehrers [Competences, functions and classroom activities of teachers]. In F.E. Weinert (Ed.). *Enzyklopädie der Psychologie: Themenbereich D, Praxisgebiete: Serie 1, Pädagogische Psychologie, Bd. 3* (pp. 177-212) [Encyclopedia of psychology – pedagogical psychology]. Göttingen: Hogrefe.
- Bromme, R. & Haag, L. (2008). Forschung zur Lehrerpersönlichkeit [Research on teacher personality]. In W. Helsper & J. Böhme (Eds.). *Handbuch der Schulforschung*, 2. Ed. (pp. 803-819) [Handbook of school research]. Wiesbaden: Verlag für Sozialwissenschaften.
- Brouwer, C. N. (2010). Determining long term effects of teacher education. In P. Peterson, E. Baker, & B. McGaw (Eds.). *International encyclopedia of education*. 7. Ed. (pp.503-510). Amsterdam: Elsevier.
- Bruner, J. S. (1977). *The process of education*. Cambridge, MA: Harvard University Press.
- Busker, M. (2010). *Entwicklung einer adressatenorientierten Übungskonzeption im Übergang Schule-Universität auf Basis empirischer Analysen von Studieneingangsvoraussetzungen im Fach Chemie* [Developing a tutorial concept for beginning chemistry university students based on empirical analyses of those students' entry conditions]. Dissertation. Tönning: Der Andere Verlag.
- Butts, B. & Smith, R. (1997). HSC Chemistry students' understanding of the structure and properties of molecular and ionic compounds. *Research in Science Education*, 17, 192-201.
- Carter, C. S., Larussa, M. A. & Bodner, G. M. (1987). A study of two measures of spatial ability as predictors of success in different levels of general chemistry. *Journal of research in science teaching*, 24(7), 645-657.
- Centrum für Hochschulentwicklung (Ed.) (2016). *Monitor Lehrerbildung* [Monitor teacher education]. Last accessed 2017-08-25 at <http://www.monitor-lehrerbildung.de/web/universitaet/index.html>
- Childs, P. E. & Sheehan, M. (2009). What's difficult about chemistry? An Irish perspective. *Chemistry Education Research and Practice*, 10(3), 204-218.
- Cochran, K. F., DeRuiter, J. A. & King, R. A. (1993). Pedagogical content knowing: An integrative model for teacher preparation. *Journal of teacher Education*, 44(4), 263-272.

-
- Cochran, K. F. & Jones, L. L. (1998). The subject matter knowledge of preservice science teachers. In B. Fraser & K. Tobin (Eds.). *International handbook of science education*, Vol. 2 (pp. 707-718). Dordrecht, Boston, London: Kluwer.
- Cohen, J. (1992). A power primer. *Psychological bulletin*, *112*(1), 155-159.
- Cooper, M. M., Grove, N., Underwood, S. M. & Klymkowsky, M. W. (2010). Lost in Lewis structures: An investigation of student difficulties in developing representational competence. *Journal of Chemical Education*, *87*(8), 869-874.
- Cooper, M. M., Underwood, S. M., & Hilley, C. Z. (2012a). Development and validation of the implicit information from Lewis structures instrument (IILSI): do students connect structures with properties? *Chemistry Education Research and Practice* *13*(3), 195-200.
- Cooper, M. M., Underwood, S. M., & Hilley, C. Z. (2012b). Development and Assessment of a Molecular Structure and Properties Learning Progression. *Journal of Chemical Education*, *89*(8), 1351-1357.
- Davidowitz, B. & Chittleborough, G. (2009). Linking the Macroscopic and Sub-microscopic Levels – Diagrams. In J. Gilbert & D. Treagust (Eds.). *Multiple Representations in Chemical Education* (pp. 169-191). New York: Springer.
- Dicke, T., Parker, P. D., Holzberger, D., Kunina-Habenicht, O., Kunter, M. & Leutner, D. (2015). Beginning teachers' efficacy and emotional exhaustion: Latent changes, reciprocity, and the influence of professional knowledge. *Contemporary Educational Psychology*, *41*, 62-72.
- Ding, L., Chabay, R., Sherwood, B. & Beichner, R. (2006). Evaluating an electricity and magnetism assessment tool: Brief electricity and magnetism assessment. *Physical review special Topics - Physics education research*, *2*(1), 010105.
- Döhrmann, M., Kaiser, G. & Blömeke, S. (2010). Messung des mathematischen und mathematikdidaktischen Wissens: Theoretischer Rahmen und Teststruktur [Measuring mathematical content knowledge and pedagogical content knowledge – Theoretical framework and test structure]. In S. Blömeke, G. Kaiser & R. Lehmann (Eds.). *TEDS-M 2008 - Professionelle Kompetenz und Lerngelegenheiten angehender Mathematiklehrkräfte für die Sekundarstufe I im internationalen Vergleich* (pp. 169-196) [TEDS-M 2008 – Professional competence and learning opportunities of prospective math teachers for lower secondary in an international comparison]. Münster: Waxmann.
- Dollny, S. (2011). *Entwicklung und Evaluation eines Testinstruments zur Erfassung des fachspezifischen Professionswissens von Chemielehrkräften* [Development and evaluation of a test instrument for measuring subject-specific professional knowledge of chemistry teachers]. Dissertation. Berlin: Logos.
- Draney, K. & Wilson, M. (2008). A LLTM approach to the examination of teachers' ratings of classroom assessment tasks. *Psychology Science Quarterly*, *50*(3), 417-432.

-
- Duis, J. (2011). Organic Chemistry Educators' Perspectives on Fundamental Concepts and Misconceptions: An Exploratory Study. *Journal of Chemical Education*, *88*(3), 346-350.
- Eurydice (Ed.) (2002). *Key topics in education in Europe – Vol. 3: The teaching profession in Europe – profile, trends and concerns. Report I: Initial training and transition to working life*. Brussels.
- Fernández-Balboa, J. M. & Stiehl, J. (1995). The generic nature of pedagogical content knowledge among college professors. *Teaching and Teacher Education*, *11*(3), 293-306.
- Field, A. (2009). *Discovering statistics using IBM SPSS*. 3rd Ed. Los Angeles: Sage.
- Fisher, M. L., Meredith, T. & Gray, M. (2018). Sex Differences in Mental Rotation Ability Are a Consequence of Procedure and Artificiality of Stimuli. *Evolutionary Psychological Science*, *4*(2), 124-133.
- Freyer, K. (2013). Zum Einfluss von Studieneingangsvoraussetzungen auf den Studienerfolg Erstsemesterstudierender im Fach Chemie [The influence of study pre-requisites on study success of first semester chemistry students] (Vol. 156). Berlin: Logos.
- Fulmer, G. W., Chu, H. E., Treagust, D. F. & Neumann, K. (2015). Is it harder to know or to reason? Analyzing two-tier science assessment items using the Rasch measurement model. *Asia-Pacific Science Education*, *1*(1), 1-16, DOI 10.1186/s41029-015-0005-x.
- Gabel, D. (1999). Improving Teaching and Learning through Chemistry Education Research: A Look to the Future. *Journal of Chemical Education*, *76*(4), 548-554.
- Garnett, P.J., Garnett, P.J. & Hackling, M.W. (1995). Students' Alternative Conceptions in Chemistry – A Review of Research and Implications for Teaching and Learning. *Studies in Science Education*, *25*, 69-95.
- Geddis, A. N. (1993). Transforming subject-matter knowledge: the role of pedagogical content knowledge in learning to reflect on teaching. *International Journal of Science Education*, *15*(6), 673-683.
- Gess-Newsome, J. (1999). Pedagogical content knowledge: An introduction and orientation. In J. Gess-Newsome & N.G. Lederman (Eds.). *Examining pedagogical content knowledge* (pp. 3-17). Dordrecht: Kluwer.
- Gess-Newsome, J. & Carlson, J. (2013). The PCK summit consensus model and definition of pedagogical content knowledge. *The Symposium "Reports from the Pedagogical Content Knowledge (PCK) Summit"*. ESERA conference, Nicosia, September 2-7, 2013.
- GDCh – Gesellschaft Deutscher Chemiker (Ed.) (2008). *Das Bachelor-/Master-Studium für das Lehramt Chemie – Vorschläge zur Restrukturierung und curricularen Entwicklung* [Bachelor-/Master-studies for chemistry teachers – suggestions for a restructuring and curricular development]. Frankfurt am Main.

-
- GDCh – Gesellschaft Deutscher Chemiker (Ed.) (2015). *Empfehlungen der GDCh-Studienkommission zum Bachelor-Studium Chemie an Universitäten* [Recommendations of the GDCh study commission for the bachelor degree in chemistry at universities]. Frankfurt am Main.
- Gilbert, J. K. & Treagust, D. F. (Eds.) (2009). *Multiple Representations in Chemical Education*. New York: Springer.
- Gold, B. & Holodyski, M. (2015). Development and Construct Validation of a Situational Judgment Test of Strategic Knowledge of Classroom Management in Elementary Schools. *Educational Assessment, 20*(3), 226-248.
- Graulich, N., Hopf, H. & Schreiner, P. (2010). Heuristic thinking makes a chemist smart. *Chemical Society Reviews, 39*, 1503-1512.
- Grossman, P. L. (1990). *The making of a teacher: Teacher knowledge and teacher education*. Teachers College Press, Teachers College, Columbia University.
- Grossschedl, J., Harms, U., Herzog, S., Mahler, D. & Steffensky, M. (in prep.). Fächerkombination als Determinante professioneller Kompetenz angehender Lehrkräfte in den naturwissenschaftlichen Fächern [Subject combination as a determining factor of professional competence of pre-service teachers in the sciences].
- Grossschedl, J., Harms, U., Kleickmann, T. & Glowinski, I. (2015). Preservice Biology Teachers' Professional Knowledge: Structure and Learning Opportunities. *Journal of Science Teacher Education, 26*(3), 291-318.
- Hashweh, M. Z. (2005). Teacher pedagogical constructions: A reconfiguration of pedagogical content knowledge. *Teachers and Teaching: Theory and Practice, 11*(3), 273-292.
- Hassan, A. K., Hill, R. A. & Reid, N. (2004). Ideas underpinning success in an introductory course of organic chemistry. *University Chemistry Education, 8*, 40-51.
- Hattie, J. A. C. & Hansford, B. C. (1982). Self measures and achievement: Comparing a traditional review of literature with meta-analysis. *Australian Journal of Education, 26*, 71-75.
- Heller, K. A. & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12. Klassen, Revision: KFT 4-12+ R* [Cognitive Ability Test for 4th grade to 12th grade, Revision]. Beltz-Test.
- Helmke, A. (2003). *Unterrichtsqualität - erfassen, bewerten, verbessern* [Teaching quality – assessing, evaluating and improving it]. Seelze: Kallmeyer.
- Henderleiter, J., Smart, R., Anderson, J. & Elian, O. (2001). How do Organic Chemistry Students Understand and Apply Hydrogen Bonding? *Journal of Chemical Education, 78*(8), 1126-1130.
- Herzog, S., Finlayson, O., Hickmann, K. & Parchmann, I. (2015). Poren nach Maß – Struktur-Eigenschafts-Beziehungen unter der chemiedidaktischen Lupe [Pores to fit – structure-property-relations under the chemical magnifying glass]. *CHEMKON, 22*(1), 29-36.

-
- Hiebert, J., Gallimore, R. & Stigler, J. W. (2002). A knowledge base for the teaching profession: What would it look like and how can we get one? *Educational Researcher*, 31(5), 3–15.
- Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, 42(2), 371-406.
- Hill, H.C., Schilling, S.G. & Ball, D.L. (2004). Developing Measures of Teachers' Mathematics Knowledge for Teaching. *Elementary School Journal*, 105, 11-30.
- Johnstone, A. H. (1991). Why is science difficult to learn? Things are seldom what they seem. *Journal of computer assisted learning*, 7(2), 75-83.
- Johnstone, A. H. (2000). Teaching of Chemistry – Logical or Psychological? *Chemistry Education: Research and Practice in Europe*, 1(1), 9–15.
- Jonassen, D. H. & Ionas, I. G. (2008). Designing effective supports for causal reasoning. *Educational Technology Research and Development*, 56(3), 287–308.
- Kauertz, A. (2014). Entwicklung eines Rasch-skalierten Leistungstests [Development of a Rasch scaled achievement test]. In D. Krüger, I. Parchmann & H. Schecker (Eds.). *Methoden in der naturwissenschaftsdidaktischen Forschung* (pp. 341-353) [Methods in science education research]. Berlin, Heidelberg: Springer Spektrum.
- Kaya, O. N. (2009). The Nature of Relationships among the Components of Pedagogical Content Knowledge of Preservice Science Teachers: 'Ozone layer depletion' as an example. *International Journal of Science Education*, 31(7), 961-988.
- Keig, P.F. & Rubba, P.A. (1993). Translation of representations of the structure of matter and its relationship to reasoning, gender, spatial reasoning, and specific prior knowledge. *Journal of Research in Science Teaching*, 30(8), 883–903.
- Kelava, A. & Moosbrugger, H. (2008). Deskriptivstatistische Evaluation von Items (Itemanalyse) und Testwertverteilungen [Descriptive statistic evaluation of items and test value distribution]. In H. Moosbrugger & A. Kelava (Eds.). *Testtheorie und Fragebogenkonstruktion* (pp. 73-98) [Test theory and questionnaire construction]. Heidelberg: Springer.
- KFC – Konferenz der Fachbereiche Chemie an den wissenschaftlichen Hochschulen in Deutschland (Ed.) (2014). *Chemielehrerausbildung an den Hochschulen im deutschsprachigen Raum* [Chemistry teacher education at universities in German-speaking countries]. Last accessed: 2016-11-29 at: http://www.kfc.cup.uni-muenchen.de/la_website/index.html
- Kind, V. (2004). *Beyond Appearances: Students' Misconceptions about Basic Chemical Ideas*; 2nd Ed. London: Royal Society of Chemistry. Last accessed 2012-09-05 at: http://www.rsc.org/images/Misconceptions_update_tcm18-188603.pdf.
- Kind, V. (2009). A Conflict in Your Head: An exploration of trainee science teachers' subject matter knowledge development and its impact on teacher self-confidence. *International Journal of Science Education*, 31(11), 1529–1562.

- Kirschner, S. (2013). *Modellierung und Analyse des Professionswissens von Physiklehrkräften* [Modelling and analysis of physics teachers' professional knowledge]. Dissertation. Studien zum Physik- und Chemielernen: Vol. 161. Berlin: Logos.
- Kirschner, S., Sczudlek, M., Tepner, O., Borowski, A., Fischer, H. E., Lenkse, G., Leutner, D., Neuhaus, B. J., Sumfleth, E., Thillmann, H. & Wirth, J. (2016). Professionswissen in den Naturwissenschaften (ProwiN) [Professional knowledge in the Sciences (ProwiN)]. In C. Gräsel & K. Trempler (Eds.). *Entwicklung von Professionalität pädagogischen Personals. Interdisziplinäre Betrachtungen, Befunde und Perspektiven* (pp. 113-130) [Development of pedagogical personnel's professionalism. Interdisciplinary considerations, results and perspectives]. Wiesbaden: Springer.
- Kleickmann, T., Großschedl, J., Harms, U., Heinze, A., Herzog, S., Hohenstein, F., Köller, O., Kröger, J., Lindmeier, A., Loch, C., Mahler, D., Möller, J., Neumann, K., Parchmann, I., Steffensky, M., Taskin, V. & Zimmermann, F. (2014). Professionswissen von Lehramtsstudierenden der mathematisch-naturwissenschaftlichen Fächer – Testentwicklung im Rahmen des Projekts KiL [Professional knowledge of pre-service math and science teachers – test development within the KiL-project]. *Unterrichtswissenschaft*, 42, 280-288.
- Klieme, E. (2004). Was sind Kompetenzen und wie lassen sie sich messen? [What are competences and how can they be measured?]. *Pädagogik*, 6, 10-13.
- Klieme, E., Hartig, J. & Rauch, D. (2008). The concept of competence in educational contexts. In J. Hartig, E. Klieme & D. Leutner (Eds.). *Assessment of competencies in educational contexts* (pp. 3-22). Göttingen: Hogrefe.
- Klostermann, M. (2014). *Lehr-/ Lern-Überzeugungen von Studierenden und Lehrenden im Fach Chemie im ersten Studienjahr* [Beliefs of teaching and learning of students and professors of chemistry in the first year at university]. Dissertation. Kiel.
- KMK – Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (Ed.) (2004). *Beschlüsse der Kultusministerkonferenz – Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss* [Resolutions of the Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic – Educational standards in chemistry for the Realschule leaving certificate]. Beschluss vom 16.12.2004.
- KMK – Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (Ed.) (2005). *Bildungsstandards im Fach Chemie für den Mittleren Schulabschluss* [Educational standards for the subject of chemistry relating to the medium leaving certificate]. Luchterhand: München, Germany.
- KMK – Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (Ed.) (2008). *Ländergemeinsame inhaltliche Anforderungen für die Fachwissenschaften und Fachdidaktiken in der Lehrerbildung* [Content requirements for subject-related studies and subject-related didactics in teacher training which apply to all Länder]. Last accessed 2011-02-15 at:

http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung.pdf

- KMK – Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (Ed.) (2015). *The Education System in the Federal Republic of Germany 2013/2014*. Last accessed 2016-07-22 at: <https://www.kmk.org/dokumentation-und-statistik/informationen-zum-deutschen-bildungssystem/dossier-englisch.html>
- Koballa, T. R., Gräber, W., Coleman, D. & Kemp, A. C. (1999). Prospective teachers' conceptions of the knowledge base for teaching chemistry at the German gymnasium. *Journal of Science Teacher Education*, 10(4), 269-286.
- Krauss, S., Blum, W., Brunner, M., Neubrand, M., Baumert, J., Kunter, M., Besser, M. & Elsner, J. (2013). Mathematics Teachers' Domain-Specific Professional Knowledge: Conceptualization and Test Construction in COACTIV. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.). *Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers – Results from the Coactiv project* (pp. 147-174). New York: Springer.
- Krauss, S. & Brunner, M. (2011). Schnelles Beurteilen von Schülerantworten: Ein Reaktionszeittest für Mathematiklehrer/innen [Fact assessment of student answers: a test measuring reaction times of math teachers]. *Journal für Mathematikdidaktik*, 32, 233–251.
- Krauss, S., Brunner, M., Kunter, M., Baumert, J., Blum, W., Neubrand, M. & Jordan, A. (2008). Pedagogical content knowledge and content knowledge of secondary mathematics teachers. *Journal of Educational Psychology*, 100(3), 716–725.
- Kruse, R. A. & Roehrig, G. H. (2005). A comparison study: Assessing teachers' conceptions with the chemistry concepts inventory. *Journal of Chemical Education*, 82(8), 1246-1250.
- Kunter, M., Klusmann, U., Genz, K., Busching, R. & Peters, R. (2009). *COACTIV-R: Kompetenzerwerb von Lehramtskandidat(inn)en im Vorbereitungsdienst* [COACTIV-R Competence acquisition of pre-service teachers in their practical teaching phase]. Berlin: Max-Planck-Institut für Bildungsforschung.
- Lampert, M. & Ball, D. L. (1999). Aligning teacher education with contemporary K-12 reform visions. In L. Darling-Hammond & G. Sykes (Eds.). *Teaching as the learning profession: Handbook of policy and practice* (pp. 33–53). San Francisco, CA: Jossey-Bass.
- Lawrenz, F. (1986). Misconceptions of physical science concepts among elementary school teachers. *School Science and Mathematics*, 86(8), 654-660.
- Lee, E. & Luft, J. A. (2008). Experienced secondary science teachers' representation of pedagogical content knowledge. *International Journal of Science Education*, 30(10), 1343-1363.
- Lee, O., Eichinger, D.C., Anderson, C.W., Berkheimer, G.D. & Blakeslee, T.D. (1993). Changing Middle School Students' Conceptions of Matter and Molecules. *Journal of Research in Science Teaching*, 30(3), 249–270.

-
- Leinhardt, G. & Smith, D. A. (1985). Expertise in mathematics instruction: Subject matter knowledge. *Journal of Educational Psychology*, 77(3), 247–271.
- Levy Nahum, T., Hofstein, A., Mamlok-Naaman, R. & Bar-Dov, Z. (2004). Can Final Examinations Amplify Students' Misconceptions in Chemistry? *Chemistry Education Research and Practice*, 5(3), 301-325.
- Levy Nahum, T., Mamlok-Naaman, R., Hofstein, A. & Kronik, L. (2008). A New „Bottom-Up“ Framework for Teaching Chemical Bonding *Journal of Chemical Education*, 85(12), 1680-1685.
- Liang, L. L., Chen, S., Chen, X., Kaya, O. N., Adams, A. D., Macklin, M. & Ebenezer, J. (2006). Student Understanding of Science and Scientific Inquiry (SUSSI): Revision and Further Validation of an Assessment Instrument. Paper prepared for NARST conference in San Francisco, CA, April 3-6, 2006, 1-37.
- Lipowsky, F. (2006). Auf den Lehrer kommt es an. Empirische Evidenzen für Zusammenhänge zwischen Lehrerkompetenzen, Lehrerhandeln und dem Lernen der Schüler [It depends on the teacher: empirical evidence for relationships between teacher competence, teacher actions and student learning]. 51. Beiheft der Zeitschrift für Pädagogik: Kompetenzen und Kompetenzentwicklung von Lehrerinnen und Lehrern: Ausbildung und Beruf, 47-70.
- Llinares, S. & Krainer, K. (2006). Mathematics (student) teachers and teacher educators as learners. In A. Gutiérrez & P. Boero (Eds.). *Handbook of Research on the psychology of mathematics education: Past, present and future* (pp. 429-459). Rotterdam: Sense Publisher.
- Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Hillsdale, NJ: Lawrence Erlbaum.
- Maeyer, J. & Talanquer, V. (2010). The Role of Intuitive Heuristics in Students' Thinking – Ranking of Chemical Substances. *Science Education*, 94, 963-984.
- Magnusson, S., Krajcik, J. & Borko, H. (1999). Nature, sources, and development of pedagogical content knowledge for science teaching. In J. Gess-Newsome & N.G. Lederman (Eds.). *Examining pedagogical content knowledge* (pp. 95-132). Dordrecht: Kluwer.
- Mahaffy, P. (2006). Moving Chemistry Education into 3D – A Tetrahedral Metaphor for Understanding Chemistry. *Journal of Chemical Education*, 83(1), 49–55.
- Marks, R. (1990). Pedagogical Content Knowledge: From a Mathematical Case to a Modified Conception. *Journal of Teacher Education*, 41(3), 3-11.
- Martinez, M.E. (1999). Cognition and the Quest of Test Item Format. *Educational Psychologist*, 34(4), 207-218.
- Mattern, J. & Bauer, J. (2014). Does teachers' cognitive self-regulation increase their occupational well-being? The structure and role of self-regulation in the teaching context. *Teaching and Teacher Education*, 43, 58-68.

-
- Mayer, J. & Wellnitz, N. (2014). Die Entwicklung von Kompetenzstrukturmodellen [Development of models showing structure of competence]. In D. Krüger, I. Parchmann & H. Schecker (Eds.). *Methoden in der naturwissenschaftsdidaktischen Forschung* (pp. 19-29) [Methods in science education research]. Berlin, Heidelberg: Springer Spektrum.
- Mayer, R. E. (1975). Information processing variables in learning to solve problems. *Review of Educational Research*, 45(4), 525-541.
- McCormack, L., Finlayson, O. E. & McCloughlin, T. (2010). The cognitive developmental levels of a sample of first year university science students. Paper presented at the International Conference on Engaging Pedagogy, Maynooth, Ireland. Last accessed 2014-05-07 at: <http://icep.ie/paper-template/?pid=52>
- Meijer, M.R. (2011). *Macro-meso-micro thinking with structure-property relations for chemistry education - An explorative design-based study*. Dissertation. Utrecht University.
- Meinhardt, C., Rabe, T. & Krey, O. (2016). Selbstwirksamkeitserwartungen in physikdidaktischen Handlungsfeldern. Skalendokumentation (Version 1.0) [Self-efficacy in activities in the physics classroom. Scale documentation]. URN: urn:nbn:de:0111-pedocs-118180.
- Messner, H. & Reusser, K. (2000). Die berufliche Entwicklung von Lehrpersonen als lebenslanger Prozess [Professional development of teachers as a life-long process]. *Beiträge zur Lehrerbildung*, 18(2), 157-171.
- Miles, J., & Shevlin, M. (2001). *Applying regression & correlation – a guide for students and researchers*. London: SAGE Publications.
- Ministerium für Schule und Berufsbildung des Landes Schleswig-Holstein (Ed.) (2016). *Fachanforderungen Chemie – Allgemeinbildende Schulen, Sekundarstufe I - Gymnasium, Sekundarstufe II* [Chemistry Curriculum, general education schools, lower secondary, Gymnasium upper secondary]. Kiel. Schmidt & Klaunig. Last accessed 2017-01-10 at: <http://lehrplan.lernnetz.de>
- MNU – Mathematisch Naturwissenschaftlicher Unterricht (Ed.) (2004). *Empfehlungen zur Ausbildung von Chemielehrern in CHEMIEDIDAKTIK an Hochschule und Seminar – Ausbildungsstandards und Projektideen* [Suggestions for educating chemistry teachers in PCK at university and in the seminar – Standards for education and project ideas]. Results from a joint conference in Bad Honnef, 14th - 17th March, 2004.
- Möller, J. & Köller, O. (2004). Die Genese akademischer Selbstkonzepte [On the development of academic self-concepts: The impact of social and dimensional comparisons]. *Psychologische Rundschau*, 55(1), 19-27.
- Müller-Fohrbrodt, G., Cloetta, B. & Dann, H. D. (1978). *Der Praxischock bei jungen Lehrern* [Praxis shock in young teachers]. Stuttgart: Klett.
- Mulford, D. R. & Robinson, W. R. (2002). An Inventory of Alternate Conceptions among First-Semester General Chemistry Students. *Journal of Chemical Education*, 79(6), 736–744.

-
- Mulhall, P., Berry, A. & Loughran, J. (2003). Frameworks for representing science teachers' pedagogical content knowledge. *Asia-Pacific Forum on Science Learning and Teaching*, 4(2), 1-25.
- Multon, K. D., Brown, S. D. & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of counseling psychology*, 38(1), 30-38.
- Munby, H., Russell, T. & Martin, A. K. (2001). Teachers' knowledge and how it develops. In V. Richardson (Ed.). *Handbook of Research on Teaching* (pp. 877-904). Washington: American Educational Research Association.
- Nakhleh, M. B., Samarapungavan, A. & Saglam, S. (2005). Middle School Students' Beliefs about Matter. *Journal of Research in Science Teaching*, 42(5), 581-612.
- National Research Council (Ed.) (2012). *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*. Committee on a Conceptual Framework for New K-12 Science Education Standards. Board on Science Education, Division of Behavioral and Social Sciences and Education. Washington, DC: The National Academies Press.
- Nentwig, P., Demuth, R., Parchmann, I., Gräsel, C. & Ralle, B. (2007). Chemie im Kontext: Situated Learning in Relevant Contexts while Systematically Developing Basic Chemical Concepts. *Journal of Chemical Education*, 84, 1439-1444.
- Neumann, K., Kauertz, A., Lau, A., Notarp, H. & Fischer, H. E. (2007). Die Modellierung physikalischer Kompetenz und ihrer Entwicklung [Modelling structure and development of students' physics competence]. *Zeitschrift für die Didaktik der Naturwissenschaft*, 13, 101-121.
- Neumann, K., Fischer, H. E. & Kauertz, A. (2010). From PISA to educational standards: The impact of large-scale assessments on science education in Germany. *International Journal of Science and Mathematics Education*, 8(3), 545-563.
- Ngai, C., Sevian, H. & Talanquer, V. (2014). What is this substance? What makes it different? Mapping progression in students' assumptions about chemical identity. *International Journal of Science Education*, 36(14), 2438-2461.
- NGSS – Next Generation Science Standards Lead States (Ed.) (2013). *Next Generation Science Standards: For States, By States*. Washington, DC: The National Academies Press.
- Niedersächsisches Kultusministerium (Ed.) (2009). *Kerncurriculum Chemie für das Gymnasium – gymnasiale Oberstufe, die Gesamtschule – gymnasiale Oberstufe, das Fachgymnasium, das Abendgymnasium, das Kolleg* [Chemistry curriculum for Gymnasium, upper secondary and comparable school types]. Hannover. Unidruck. Last accessed 2012-03-15 at: <http://db2.nibis.de/1db/cuvo/ausgabe/>
- Nilsson, P. (2008). Teaching for Understanding: The complex nature of pedagogical content knowledge in pre-service education. *International Journal of Science Education*, 30(10), 1281-1299.
- Novick, S. & Nussbaum, J. (1978). Pupils' understanding of the Particulate Nature of Matter: An interview study. *Science Education*, 62(3), 187-195.

-
- Özmen, H. (2004). Some Students' Misconceptions in Chemistry: A Literature Review of Chemical Bonding. *Journal of Science Education and Technology*, 13(2), 147–159.
- O'Dwyer, A. & Childs, P.E. (2011). *Second Level Irish pupils' and teachers' view of Difficulties in Organic Chemistry*. Presented at ESERA Conference 5-9 Sept 2011 in Lyon, France.
- Parchmann, I., Broman, K., Busker, M. & Rudnik, J. (2015). Context-Based Teaching and Learning on School and University Level. In J. García-Martínez & E. Serrano-Torregrosa (Eds). *Chemistry Education – Best Practices, Opportunities and Trends* (pp. 259-278). Weinheim: Wiley-VCH.
- Parchmann, I., Bündler, W., Demuth, R., Freienberg, J., Klüter, R. & Ralle, B. (2006). Lernlinien zur Verknüpfung von Kontextlernen und Kompetenzentwicklung [Learning developments to link context-based learning and competence development]. *CHEMKON*, 13(3), 124-131.
- Parchmann, I., Graesel, C., Baer, A., Nentwig, P., Demuth, R., Ralle, B. & the ChiK Project Group (2006). „Chemie im Kontext“: a symbiotic implementation of a context-based teaching and learning approach. *International Journal of Science Education*, 28(9), 1087–1112.
- Paris, S. G., Lipson, M. Y. & Wixson, K. K. (1983). Becoming a Strategic Reader. *Contemporary Educational Psychology*, 8, 293-316.
- Park, S., & Oliver, J. S. (2008). Revisiting the conceptualisation of pedagogical content knowledge (PCK): PCK as a conceptual tool to understand teachers as professionals. *Research in Science Education*, 38(3), 261-284.
- Peterson, R.F., Treagust, D. & Garnett, P. (1989). Development and use of diagnostic tests to evaluate students' misconceptions in science. *International Journal of Science Education*, 26, 301–314.
- Pintrich, P.R. (2002). The Role of Metacognitive Knowledge in Learning, Teaching, and Assessing. *Theory Into Practice*, 41(4), 219-225, DOI: 10.1207/s15430421tip4104_3.
- Potgieter, M., Rogan, J. M. & Howie, S. (2005). Chemical concepts inventory of grade 12 learners and UP foundation year students. *African Journal of Research in Mathematics, Science and Technology Education*, 9(2), 121-134.
- Pribyl, J. R. & Bodner, G. M. (1987). Spatial ability and its role in organic chemistry: A study of four organic courses. *Journal of research in science teaching*, 24(3), 229-240.
- Rehm, M. & Bölsterli, K. (2014). Entwicklung von Unterrichtsvignetten [Development of teaching vignettes]. In D. Krüger, I. Parchmann & H. Schecker (Eds.). *Methoden in der naturwissenschaftsdidaktischen Forschung* (pp. 213-226) [Methods in science education research]. Berlin, Heidelberg: Springer Spektrum.
- Riese, J. (2009). *Professionelles Wissen und professionelle Handlungskompetenz von (angehenden) Physiklehrkräften* [Professional knowledge and professional action competence of (prospective) physics teachers]. Dissertation. Berlin: Logos.

-
- Riese, J. & Reinhold, P. (2010). Empirische Erkenntnisse zur Struktur professioneller Handlungskompetenz von angehenden Physiklehrkräften [Empirical findings regarding the structure of future physics teachers' competence regarding professional action]. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 167-187.
- Riese, J. & Reinhold, P. (2012). Die professionelle Kompetenz angehender Physiklehrkräfte in verschiedenen Ausbildungsformen: Empirische Hinweise für eine Verbesserung des Lehramtsstudiums [Professional competence of pre-service physics teachers in different teaching tracks: empirical hints for an improvement of teacher education studies]. *Zeitschrift für Erziehungswissenschaft*, 15(1), 111-143.
- Risch, B. (Ed.). (2010a). *Teaching Chemistry around the World*. Münster: Waxmann.
- Risch, B. (2010b). Germany. In B. Risch (Ed.). *Teaching Chemistry around the World* (pp. 267-279). Münster: Waxmann.
- Rohaan, E. J. (2009). *Testing teacher knowledge for technology teaching in primary schools*. Dissertation. Technische Universiteit Eindhoven, Printservice TU/e.
- Rollnick, M., Bennett, J., Rhemtula, M., Dharsey, N. & Ndlovu, T. (2010). The Place of Subject Matter Knowledge in Pedagogical Content Knowledge: A case study of South African teachers teaching the amount of substance and chemical equilibrium. *International Journal of Science Education*, 30(10), 1365-1387.
- Ropohl, M. (2010). *Modellierung von Schülerkompetenzen im Basiskonzept Chemische Reaktion: Entwicklung und Analyse von Testaufgaben*. [Modelling of student competences in the fundamental concept of chemical reaction: development and analysis of test items] Dissertation. Berlin: Logos.
- Rost, J. (2004). *Lehrbuch Testtheorie – Testkonstruktion*, 2nd Ed. [Test theory, test construction]. Bern: Verlag Hans Huber.
- Rushton, G. T., Hardy, R. C., Gwaltney, K. P. & Lewis, S. E. (2008). Alternative conceptions of organic chemistry topics among fourth year chemistry students. *Chemistry Education Research and Practice*, 9(2), 122–130.
- Sanmarti, N., Izquierdo, M. & Watson, R. (1995). The Substantialisation of Properties in Pupils' Thinking and in the History of Science. *Science & Education*, 4(4), 349–369.
- Schecker, H. & Parchmann, I. (2006). Modellierung naturwissenschaftlicher Kompetenz [Modelling scientific competences]. *Zeitschrift für die Didaktik der Naturwissenschaften*, 12, 45-66.
- Schecker, H. & Parchmann, I. (2007). Standards and competence models: The German situation. In D. Waddington, P. Nentwig & S. Schanze (Eds.). *Making it comparable. Standards in science education* (pp. 147-164). Münster: Waxmann.
- Scheffel, L., Brockmeier, W. & Parchmann, I. (2009). Historical Material in Macro-Micro Thinking – Conceptual Change in Chemistry Education and the History of Chemistry. In J. Gilbert & D. Treagust (Eds.). *Multiple Representations in Chemical Education* (pp. 215-250). New York: Springer.

- Scheffel, L., Herzog, S. & Parchmann, I. (2011). *Educational Reconstruction of Structure-Property-Relations*. ESERA: Science Learning and Citizenship, Sept 5-9, 2011.
- Schiefele, U., Krapp, A., Wild, K. P. & Winteler, A. (1993). Der Fragebogen zum Studieninteresse (FSI) [Study interest questionnaire (SIQ)]. *Diagnostika*, 39(4), 335-351.
- Schiepe-Tiska, A., Schöps, K., Rönnebeck, S., Köller, O. & Prenzel, M. (2013). Naturwissenschaftliche Kompetenz in PISA 2012: Ergebnisse und Herausforderungen [Scientific competence in PISA 2012: Results and challenges]. In M. Prenzel, C. Sälzer, E. Klieme & O. Köller (Eds.). *PISA 2012. Fortschritte und Herausforderungen in Deutschland* (pp.189-215) [PISA 2012. Progress and challenges in Germany]. Münster: Waxmann.
- Schilling, S. G. (2007). The role of psychometric modeling in test validation: An application of multidimensional item response theory. *Measurement*, 5(2-3), 93-106.
- Schermelleh-Engel, K., Moosbrugger, H. & Müller, H. (2003). Evaluating the fit of structural equation models: Test of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research – Online*, 8, 23-74.
- Schmelzing, S. (2010). *Das fachdidaktische Wissen von Biologielehrkräften: Konzeptionalisierung, Diagnostik, Struktur und Entwicklung im Rahmen der Biologielehrerbildung* [Pedagogical content knowledge of biology teachers – conceptualization, diagnostics, structure and development within biology teacher education]. Dissertation. Berlin: Logos.
- Schmidkunz, H. & Lindemann, H. (1992). *Das forschend-entwickelnde Unterrichtsverfahren* [Teaching concept of research and inquiry]. Essen: Westarp.
- Schmidt, J.-H. (1996). Students' understanding of molecular structure and properties of organic compounds. *Paper presented at the Annual Meeting of the National Association of Research in Science Teaching (NARST)*, St. Louis, MO.
- Schmidt, W. H., Blömeke, S., Tatto, M. T., Hsieh, F. J., Cogan, L., Houang, R. T. & Schille, J. (2011). *Teacher education matters: A study of middle school mathematics teacher preparation in six countries*. New York: Teachers College Press, Columbia University.
- Schmitz, B. (2006). Advantages of studying processes in educational research. *Learning and Instruction*, 16(5), 433-449.
- Schoenfeld, A. H. & Kilpatrick, J. (2008). Toward a Theory of Proficiency in Teaching Mathematics. In D. Tirosh & T. Wood (Eds.). *International handbook of mathematics teacher education. Vol. 2: Tools and processes in mathematics teacher education* (pp. 321-354). Rotterdam: Sense Publishers.
- Schumacher, K. & Lind, G. (2000). *Praxisbezug im Lehramtsstudium — Bericht einer Befragung von Konstanzer LehrerInnen und Lehramtsstudierenden* [Practice in teacher education — report on an examination of teacher students]. Forschungsbericht. Konstanz: Universität Konstanz.
- Schwarzer, S., Rudnik, J. & Parchmann, I. (2013). Chemische Schalter als potenzielle Lernschalter - Fachdidaktische Begleitung eines Sonderforschungsbereichs [Chemical switches as potential

-
- learning switches – chemical education accompaniment of a special research field]. *CHEMKON*, 20(4), 175-181.
- Seidel, T. & Prenzel, M. (2008). Wie Lehrpersonen Unterricht wahrnehmen und einschätzen—Erfassung pädagogisch-psychologischer Kompetenzen mit Videosequenzen [How teachers notice and assess lessons – measuring pedagogical-psychological competences with video clips]. In M. Prenzel, I. Gogolin & H.H. Krüger (Eds.). *Kompetenzdiagnostik* (pp. 201-216) [Diagnosis of competence]. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Sevian, H. & Talanquer, V. (2014). Rethinking chemistry: a learning progression on chemical thinking. *Chemistry Education Research and Practice*, 15, 10-23.
- Shavelson, R. (2010). On the measurement of competency. *Empirical Research in Vocational Education and Training*, 2(1), 41-63.
- Shavelson, R. J., Baxter, G. P & Pine, J. (1991). Performance assessment in science. *Applied Measurement in Education*, 4(4), 347-262.
- Sheehan M., Childs, P.E. & Hayes, S. (2011). *Identification of Irish Pre-service science teachers' misconceptions in chemistry*. Presented at ESERA Conference 5-9 Sept 2011 in Lyon, France.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4-14.
- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1-23.
- Smith, D. C. & Neale, D. C. (1989). The construction of subject matter knowledge in primary science teaching. *Teaching and teacher Education*, 5(1), 1-20.
- Steffensky, M., Parchmann, I. & Schmidt, S. (2005). Die Teilchen saugen das Aroma aus dem Tee – Alltagsvorstellungen und chemische Erklärungskonzepte [Particles are sucking the aroma out of the tea – misconceptions from every-day life and chemical explanatory concepts]. *Chemie in unserer Zeit*, 39, 274–278.
- Steinkamp, M. W. & Maehr, M. L. (1983). Affect, ability, and science achievement: A quantitative synthesis of correlational research. *Review of Educational Research*, 53(3), 369-396.
- Taagepera, M. & Noori, S. (2000). Mapping Students' Thinking Patterns in Learning Organic Chemistry by the Use of Knowledge Space Theory. *Journal of Chemical Education*, 77(9), 1224-1229.
- Taber, K. S. (1997). Student Understanding of Ionic Bonding: Molecular versus Electrostatic Framework? *School Science Review*, 78(285), 85–95.
- Taber, K. S. (2001). Building the Structural Concepts of Chemistry – Some Considerations from Educational Research. *Chemistry Education Research and Practice in Europe*, 2(2), 123-158.
- Taber, K. S. (2002). *Chemical misconceptions - prevention, diagnosis, and cure*. Royal Society of Chemistry: London.

- Tamir, P. (1988). Subject matter and related pedagogical knowledge in teacher education. *Teaching and Teacher Education*, 4(2), 99-110.
- Taskin, V. (2014). *Untersuchung des Fachwissens und fachdidaktischen Wissens von Chemie-Lehramtsstudierenden bezüglich chemischer Repräsentationen* [Assessing pre-service chemistry teachers' content knowledge and pedagogical content knowledge regarding chemical representations]. Dissertation, Kiel.
- Teodorescu, T. (2006). Competence versus competency: What is the difference? *Performance Improvement*, 45(10), 27-30.
- Terhart, E. (2002). *Standards für die Lehrerbildung. Eine Expertise für die Kultusministerkonferenz* [Standards for teacher education – an expertise for the Standing Conference of the Ministers of Education and Cultural Affairs of the States in the Federal Republic]. Westfälische Wilhelms-Universität Münster.
- Tepner, O., Borowski, A., Dollny, S., Fischer, H. E., Jüttner, M., Kirschner, S., Leutner, D., Neuhaus, B., Sandmann, A., Sumfleth, E., Thillmann, H. & Wirth, J. (2012). Modell zur Entwicklung von Testitems zur Erfassung des Professionswissens von Lehrkräften in den Naturwissenschaften [Modell for developing test items for measuring science teachers' professional knowledge]. *Zeitschrift für Didaktik der Naturwissenschaften*, 18, 7-28.
- Trautwein, P. D. U., Maaz, K., Lüdtke, O., Nagy, G., Husemann, N., Watermann, R. & Köller, O. (2006). Studieren an der Berufsakademie oder an der Universität, Fachhochschule oder Pädagogischen Hochschule? [Studying at the Berufsakademie or university, university of applied science or university for teaching]. *Zeitschrift für Erziehungswissenschaft*, 9(3), 393-412.
- Treagust, D. F. (1988). The development and use of diagnostic instruments to evaluate students' misconceptions in science. *International Journal of Science Education*, 10, 159-169.
- Tschannen-Moran, M. & Woolfolk Hoy, A. (2001). Teacher efficacy: Capturing an elusive construct. *Teaching and teacher education*, 17(7), 783-805.
- Tynjälä, P. & Heikkinen, H. L. T. (2011). Beginning teachers' transition from pre-service education to working life. *Zeitschrift für Erziehungswissenschaft*, 14, 11-33.
- Van Berkel, B., Pilot, A. & Bulte, A.M.W. (2009). Micro-Macro Thinking in Chemical Education – Why and how to Escape. In J. Gilbert & D. Treagust (Eds.). *Multiple Representations in Chemical Education* (pp. 31-54). New York: Springer.
- Van Dijk, E. M. & Kattmann, U. (2010). Evolution im Unterricht: Eine Studie über fachdidaktisches Wissen von Lehrerinnen und Lehrern [Evolution in the classroom – a study about teachers' pedagogical content knowledge]. *Zeitschrift für Didaktik der Naturwissenschaften*, 16, 7-21.
- Van Driel, J. H., Verloop, N. & De Vos, W. (1998). Developing science teachers' pedagogical content knowledge. *Journal of Research in Science Teaching*, 35(6), 673-695.

- Veal, W. R. & MaKinster, J. G. (1999). Pedagogical content knowledge taxonomies. *Electronic Journal of Science Education*, 3(4).
- Vereinigung der Bayerischen Wirtschaft e.V. (Ed.) (2010). *Bildungsautonomie: Zwischen Regulierung und Eigenverantwortung–die Bundesländer im Vergleich. Expertenrating der Schul- und Hochschulgesetze der Länder zum Jahresgutachten 2010* [Educational autonomy: From regulations to self-governance – federal states compared: expert rating of states' school and university laws for the 2010 expertise]. Wiesbaden: VS Verlag für Sozialwissenschaften.
- Vinner, S. (1997). The pseudo-conceptual and the pseudo-analytical thought processes in mathematics learning. *Educational Studies in Mathematics*, 34(2), 97-129.
- Voss, T. & Kunter, M. (2013). Teachers' General Pedagogical/ Psychological Knowledge. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss & M. Neubrand (Eds.). *Cognitive activation in the mathematics classroom and professional competence of teachers* (pp. 207-228). New York: Springer.
- Waddington, D., Nentwig, P. & Schanze, S. (Eds.) (2007). *Making it comparable. Standards in science education*. Münster: Waxmann.
- Walpuski, M., Kauertz, A., Kampa, N., Fischer, H.E., Mayer, J., Sumfleth, E. & Wellnitz, N. (2010). ESNaS - Evaluation der Standards für die Naturwissenschaften in der Sekundarstufe I [ESNaS – evaluation of standards of the sciences in lower secondary]. In A. Gehrman, U. Hericks & M. Lüders (Eds.). *Bildungsstandards und Kompetenzmodelle – Beiträge zu einer aktuellen Diskussion über Schule, Lehrerbildung und Unterricht* (pp. 171-184) [Educational standards and competence models – contributions to a current discussion about school, teacher education and lessons]. Bad Heilbrunn: Klinkhardt Verlag.
- Walpuski, M., Tepner, O., Sumfleth, E., Dollny, S., Hostenbach, J. & Pollender, T. (2012). Multiple perspectives on students' scientific communication & reasoning in chemistry education: VISIONS 2011: Teaching. *Acta Didactica Norge*, 6(1). Last accessed 2012-07-04 at: <http://adno.no/index.php/adno/article/view/206>
- Wandersee, J. H., Mintzes, J. J. & Novak, J. D. (1994). Research on alternative conceptions in science. In D. L. Gabel (Eds.). *Handbook of research on science teaching and learning* (pp. 177-210). New York: Simon & Schuster and Prentice Hall International.
- Weinert, F.E. (2001a). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit [Comparative performance measurement in schools – a controversial matter of course]. In F.E. Weinert (Ed.). *Leistungsmessung an Schulen* (pp. 17-31) [Performance measurements in schools]. Weinheim & Basel: Beltz Verlag.
- Weinert, F. E. (2001b). Concept of competence: A conceptual clarification. In D. S. Rychen & L. H. Salganik (Eds.). *Defining and selecting key competencies* (pp. 45-65). Kirkland, WA: Hogrefe and Huber.

- Weißeno, G. & Eck, V. (2013). *Wissen, Selbstkonzept und Fachinteresse: Ergebnisse einer Interventionsstudie zur Politikkompetenz* [Knowledge, self-concept and interest in the subject: results of an interventional study regarding competence in politics]. Münster: Waxmann.
- Wilson, M., De Boeck, P. & Carstensen, C. (2008). Explanatory item response models: A brief introduction. In J. Hartig, E. Klieme & D. Leutner (Eds.). *Assessment of competencies in educational contexts: State of the art and future prospects* (pp. 91-120). Göttingen: Hogefe & Huber.
- Winther, E. (2010). *Kompetenzmessung in der beruflichen Bildung* [Measuring competence in vocational education]. Habilitation. Bielefeld: W. Bertelsmann Verlag.
- Witner, S. & Tepner, O. (2011). Entwicklung geschlossener Testaufgaben zur Erhebung des fachdidaktischen Wissens von Chemielehrkräften [Development of closed test items for assessing chemistry teachers' pedagogical content knowledge]. *Chimica et ceterae artes rerum naturae didacticae*, 37(104), 113-137.
- Wu, M. L., Adams, R. J., Wilson, M. R. & Haldane, S. A. (2007). *ACER ConQuest version 2.0: generalised item response modelling software*. Camberwell: ACER Press.
- Yarroch, W.L. (1985). Student understanding of chemical equation balancing. *Journal of Research in Science Teaching*, 22(5), 449-459.
- Zecherle, H. (2011). *2020 – Die Zukunft der Lehrerausbildung [2020 – The Future of Teacher Education]*. Berlin: Epubli.

10 APPENDICES

10.1 Overview of items

As the items are still being used in subsequent research projects, they are not printed here. If the items are of interest, please contact the author.

10.2 SPR items in study II

Item label study II	Corresponding label from study I
FW7	E 4
FW8	E 22
FW9	E 8
FW10	E 7, <i>slightly adapted</i>
FW11	R 11_1, <i>slightly adapted</i>
FW12	R 11_2
FW13	E 20
FW14	E 23
FW28	R 14
FW29	R 18
FW30	R 22
FW31	R 28
FW32	E 1
FW33	E 5
FW34	E 16, <i>slightly adapted</i>
FW35	E 17
FW36	R12
Z1	E 30
Z2	E 24, <i>slightly adapted</i>
Z3	R 23
FDW13	I 15
FDW14	EM 3
FDW15	EM 4
FDW16	EM 7
FDW17	EM 6

FDW18	EM 9
FDW19	EM 10
FDW20	I 2
FDW21	I 6
FDW22	I 8
FDW23	S 4
FDW24	S 7
FDW25	S 10
FDW26	S 12

10.3 Overview of items in Wright Map (SPR only)

Item number in Wright Map	Item label	Dimension
1	FW_7_Bew	CK _{SPR}
2	FW_8_Bew_B	
3	FW_9_Bew	
4	FW_10_Bew	
5	FW_11_ges_A_Bew_weich	
6	FW_12_Bew	
7	FW_SEB_13_end	
8	FW_14_Bew	
9	FW_28_Bew	
10	FW_29_Bew	
11	FW_30_Bew_B	
12	FW_31_Bew	
13	FW_32_ges_B_Bew_hart	
14	FW_33_Bew	
15	FW_34_ges_Bew	
16	FW_35_ges_Bew	
17*	FW_36_Bew	
18	FW_Z1_Bew	
19	FW_Z2_ges_Bew	
20	FW_Z3_Bew	
21	FDW13_bew	PCK _{SPR}
22*	FDW14_bew	
23	FDW15_recod	
24	FDW16_bew2	
25	FDW17_ges_bew	
26	FDW18_bew	
27	FDW19_bew	
28	FDW20_bew	
29	FDW21_bew	
30	FDW22_ges_bew	
31	FDW23_bew	
32*	FDW24_bew	
33	FDW25_hart_bew	
34	FDW26_bew	

*Items were excluded due to statistical considerations

10.4 Overview of items in Wright Map (CRep and SPR)

Item number in Wright Map	Item label	Dimensions
1	FW1_end_0	CK _{CRep}
2	FW2_end_0	
3	FW3_end_streng_0	
4	FW4_end_0	
5	FW5_end_0	
6	FW6_end_0	
7	FW15_end_0	
8	FW16_end_0_sum	
9	FW17_end_0_sum	
10	FW18_end_streng_0	
11	FW19_end_0	
12	FW20_end_manuell4_0	
13	FW21_end_manuell_0	
14	FW22_end_0	
15	FW23_end_0	
16	FW24_end_0	
17	FW25_end_0	
18*	FW26_end_0	
19	FW27_end_0	
20*	FW_SEB_7_Bew	CK _{SPR}
21	FW_SEB_8_Bew_B	
22	FW_SEB_9_Bew	
23	FW_SEB_10_Bew	
24	FW_SEB_11_ges_A_Bew_weich	
25	FW_SEB_12_Bew	
26	FW_SEB_13_Bew	
27	FW_SEB_14_Bew	
28	FW_SEB_28_Bew	
29	FW_SEB_29_Bew	
30	FW_SEB_30_Bew_B	
31	FW_SEB_31_Bew	
32	FW_SEB_32_ges_B_Bew_hart	
33*	FW_SEB_33_Bew	
34	FW_SEB_34_ges_Bew	
35	FW_SEB_35_ges_Bew	
36*	FW_SEB_36_Bew	
37	FW_SEB_Z1_Bew	
38	FW_SEB_Z2_ges_Bew	
39	FW_SEB_Z3_Bew	
40	FDW1a	PCK _{CRep}
41	FDW1b	
42	FDW2	

10 APPENDICES

43	FDW3	PCK _{SPR}
44	FDW4	
45	FDW5a	
46	FDW5b	
47	FDW5c	
48	FDW6a	
49	FDW6b	
50	FDW7	
51	FDW8	
52	FDW9	
53	FDW10	
54	FDW11	
55	FDW12	
56	FDW13	
57*	FDW14	
58	FDW15	
59	FDW16	
60	FDW17	
61	FDW18	
62	FDW19	
63	FDW20	
64	FDW21	
65	FDW22	
66	FDW23	
67*	FDW24	
68	FDW25	
69	FDW26	

*Items were excluded due to statistical considerations

Lebenslauf [CV]

Persönliche Daten [personal data]

Stefanie Herzog, geb. König

Olshausenstr. 79

24106 Kiel

geb. am 1. Juni 1980 in Güstrow

Staatsangehörigkeit: Deutsch

Schulbildung und Studium [education]

- 2008 – 2010 Studium Master of Education für Gymnasium Chemie/Englisch an der Carl von Ossietzky (CvO) Universität Oldenburg
Abschluss: Master of Education
Masterarbeit: „Analyse von Einstellungen, Erwartungen und fachdidaktischen Kompetenzen bei Tutorinnen und Tutoren“
- 2005 – 2008 Studium 2-Fächer-Bachelor Chemie/Anglistik an der CvO Universität Oldenburg
Abschluss: Bachelor of Arts
Bachelorarbeit: „Ziele, Wissen und Einstellungen von Lehrkräften der Fächer Chemie und Politik zum Thema Klimawandel“
- 2002 – 2005 Pikes Peak Community College, Colorado Springs, USA
- 1997 – 1998 High School Aufenthalt in Minnesota, USA
- 1993 – 2001 Max-Planck-Gymnasium Delmenhorst
Abschluss: Allgemeine Hochschulreife
- 1991 – 1993 Orientierungsstufe Bookholzberg, Ganderkesee
- 1990 – 1991 Grundschule Schierbrok, Ganderkesee
- 1987 – 1990 Polytechnische Oberschule Wilhelm Pieck, Güstrow

Beruflicher Werdegang [work experience]

- 2016 – 2018 Wissenschaftliche Mitarbeiterin am IPN, Abteilung Didaktik der Chemie
- 2014 – 2016 Referendariat am Gymnasium an der Theodor-Storm-Schule in Husum (Fächer: Chemie und Englisch)
- 2010 – 2014 Wissenschaftliche Mitarbeiterin am IPN, Abteilung Didaktik der Chemie (inkl. Forschungsaufenthalt in Dublin u Limerick, Irland, und Durham, Großbritannien)

2002 – 2005 Arbeit in der Stadtbücherei, Colorado Springs, USA

Ehrenamtliche Tätigkeiten [volunteer work]

2017 – 2018 Vorsitzende des Personalrats am IPN

2008 – 2010 gewähltes studentisches Mitglied im Institutsrat der Chemie an der
CvO Universität Oldenburg

Publikationen [publications]

Zeitschriftenartikel [journal articles]

- Parchmann, I., Herzog, S., Terada, M. (2018). Formation of Basic Concepts of Chemistry Education in Germany. *Journal of Science Education in Japan*, 42(2), 65-72. DOI: <https://doi.org/10.14935/jssej.42.65>
- Lüthjohann, F., Herzog, S. & Parchmann, I. (2018). Metalle: Vielfältige Einsatzgebiete dank vielfältiger Berufe: Struktur-Eigenschaftsbeziehungen erarbeiten - Berufsorientierung im Fachunterricht ermöglichen. *Naturwissenschaften im Unterricht - Chemie* 164, 32-37.
- Herzog, S., Lüthjohann, F. & Parchmann, I. (2018). Wertstoffe aus Müll: Re-Funktionalisierung von erwünschten Eigenschaften. *Naturwissenschaften im Unterricht - Chemie* 164, 37 - 41.
- Herzog, S., Niebuhr, B., Heinze, A., Kampschulte, L., Lindmeier, A. & Parchmann, I. (2016). Mit PANaMa für Beruf und Studium werben. *Schule aktuell* 2016(8), 10.
- Parchmann, I., Ledwig, G. & Herzog, S. (2016). Der Energie auf der Spur: Energetische Phänomene rund ums Wasser. *Naturwissenschaften im Unterricht - Chemie* 27(153), 38-41.
- Herzog, S. & Parchmann, I. (2016). Wie bleibt das Wasser in der Windel? Superabsorbierende Polymere als Modellsubstanz zur Erarbeitung von Struktur-Eigenschafts-Beziehungen. *Naturwissenschaften im Unterricht - Chemie* 27(153), 24-27.
- Herzog, S., Finlayson, O., Hickmann, K. & Parchmann, I. (2015). Poren nach Maß – Struktur-Eigenschafts-Beziehungen unter der chemiedidaktischen Lupe. *CHEMKON - Chemie konkret* 22(1), 29-36.
- Kleickmann, T., Großschedl, J., Harms, U., Heinze, A., Herzog, S., Hohenstein, F., Köller, O., Kröger, J., Lindmeier, A., Loch, C., Mahler, D., Möller, J., Neumann, K., Parchmann, I., Steffensky, M., Taskin, V. & Zimmermann, F. (2014). Professionswissen von Lehramtsstudierenden der mathematisch-naturwissenschaftlichen Fächer – Testentwicklung im Rahmen des Projekts KiL. *Unterrichtswissenschaft* 42, 280-288.
- Parchmann, I. & Herzog, S. (2013). Chemische Haushaltsdetektive – Eine Einführung in chemische Denk- und Arbeitsweisen am Beispiel von Reinigungs- und Pflegemitteln. *Naturwissenschaften im Unterricht – Chemie* 24(137), 42-45.

Busker, M., Klostermann, M., Herzog, S., Huber, A. & Parchmann, I. (2011). Nicht nur Schulwissen auffrischen: Vorkurse in Chemie. *Nachrichten aus der Chemie* 59(6), 684-687.

Tagungsbandbeiträge [conference proceedings]

Lüthjohann, F., Herzog, S., Parchmann, I., Niebuhr, B., Heinze, A., Lindmeier, A., Kampschulte, L. & Wilken, M. (2018). Neue Ansätze zur Berufsorientierung im naturwissenschaftlichen Fachunterricht. In C. Maurer (Hrsg.): *Qualitätvoller Chemie - und Physikunterricht – normative und empirische Dimensionen*. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Regensburg 2017 (S. 663-666). Regensburg: Universität Regensburg.

Herzog, S., Taskin, V., Bernholt, S., Steffensky, M. & Parchmann, I. (2013). Entwicklung eines Testinstruments zur Erfassung des fachdidaktischen Wissens von Chemie-Lehramtsstudierenden. In S. Bernholt (Hrsg.): *Inquiry-Based Learning – Forschendes Lernen*. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Hannover 2012 (S. 536-538). Kiel: IPN.

Herzog, S. & Parchmann, I. (2013). Fachwissen von Lehramtsstudierenden zum Struktur-Eigenschafts-Konzept. In S. Bernholt (Hrsg.): *Inquiry-Based Learning – Forschendes Lernen*. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Hannover 2012 (S. 230-232). Kiel: IPN.

Klostermann, M., Busker, M., Herzog, S. & Parchmann, I. (2012). Vorkurse als Schnittstelle zwischen Schule und Universität. In S. Bernholt (Hrsg.): *Konzepte fachdidaktischer Strukturierung für den Unterricht*. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Oldenburg 2011 (S. 482-484). Münster: LIT-Verlag.

Herzog, S., Busker, M. & Parchmann, I. (2012). Untersuchung des Konzeptverständnisses am Beispiel von Struktur-Eigenschafts-Beziehungen. In S. Bernholt (Hrsg.): *Konzepte fachdidaktischer Strukturierung für den Unterricht*. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Oldenburg 2011 (S. 610-612). Münster: LIT Verlag.

Herzog, S., Busker, M. & Parchmann, I. (2011). Analyse von Einstellungen, Erwartungen und fachdidaktischen Kompetenzen bei Tutorinnen und Tutoren. In D. Höttecke (Hrsg.): *Naturwissenschaftliche Bildung als Beitrag partizipativer Demokratie*. Gesellschaft für Didaktik der Chemie und Physik, Jahrestagung in Potsdam 2010 (S. 543-545). Berlin, Münster: LIT Verlag.

Erklärung

Hiermit versichere ich, dass ich die vorliegende Dissertation – abgesehen von der Beratung durch die betreuenden Personen – nach Inhalt und Form selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe. Die Arbeit ist unter Einhaltung der Regeln guter wissenschaftlicher Praxis der Deutschen Forschungsgemeinschaft entstanden. Sie hat weder im Ganzen noch zum Teil einer anderen Stelle im Rahmen eines Promotionsverfahrens vorgelegen. Der dem Kapitel 8 zugrundeliegende Artikel ist in der Zeitschrift CHEMKON veröffentlicht. Ein Teil der Ergebnisse dieser Arbeit wurde bereits in Form von Tagungsbandbeiträgen (s. Publikationen) publiziert.

Kiel, den _____

Stefanie Herzog

Danksagung [Expression of thanks]

Wie immer an dieser Stelle gebührt der Dank eines Prüflings vielen Leuten, die ihn im Lauf seiner Arbeit maßgeblich unterstützt haben. So bedanke auch ich mich bei den folgenden Menschen:

Die erste Person, bei der ich mich von Herzen bedanke, ist meine Betreuerin und Erstgutachterin Prof. Dr. Ilka Parchmann. Ohne sie wäre ich vermutlich mehrmals in eine andere Richtung marschiert, was meine Lebensplanung betrifft: Wahrscheinlich wäre ich ohne sie weder von Oldenburg nach Kiel (zur Aufnahme einer Promotion) gezogen, noch hätte ich nach dem Referndariat einen so guten Grund gehabt, wieder nach Kiel (zur Beendigung meiner Promotion) zurückzukehren. Sie hat mich immer unterstützt und motiviert, oft aber auch gefordert und mir an den richtigen Stellen Freiräume gelassen.

My cordial thanks go to Dr. Vanessa Kind for not only making my first big international talk at a conference so much less stressful by telling me that nobody in the audience knows more about what I am going to say than I do. She also invited me not only to her institution for a research visit but also into her home allowing me to find out that research is great but that there is also a life out there with so many different skills to be mastered. I am grateful that she accepted to be my second supervisor.

Ebenfalls bedanke ich mich herzlich bei Prof. Dr. Mirjam Steffensky für fruchtbare Diskussionen zum fachdidaktischen Wissen wo immer nötig und pragmatische Entscheidungen wo immer möglich.

Ich bedanke mich bei meinen Mentoren Dr. Sascha Bernholt und J-Prof. Dr. Jörg Großschedl, die mir u.a. bei allen Fragen rund um die Statistik jederzeit qualifizierte Ansprechpartner waren.

Den Beginn meiner Promotionszeit haben insbesondere meine Mit-Doktorandinnen und inzwischen promovierten Freundinnen Dr. Vahide Taskin und Dr. Mareike Klostermann maßgeblich mitgestaltet. Ihnen danke ich für viele anregende fachliche Gespräche im Büro und auf den zahlreichen gemeinsamen Tagungen sowie zahlreiche Kinoabende und Kochrunden. Dr. Margot Janzen und Dr. habil. Sabine

Nick danke ich für fachliche Unterstützung, für die Bereitstellung von Aufgaben und ebenso für privaten motivierenden Austausch; Gerda Ledwig für unzählige chemische Gespräche und geistigen Ausgleich durch die Aufnahme in einen Chor; Frank Lüthjohann, Wilfried Wentorf und Dr. Pay Ove Dierks für ein offenes Ohr bei allen Fragen rund um die unterrichtliche Praxis. Ich bedanke mich bei der gesamten Abteilung der Chemiedidaktik in wechselnder Besetzung, allen Mitarbeiterinnen und Mitarbeitern des *KiZ*-Projekts, allen nationalen sowie internationalen Experten und Expertinnen zur Item-Bewertung und Diskussion sowie allen Studienteilnehmerinnen und -teilnehmern.

Abschließend bedanke ich mich natürlich bei meiner Familie: bei meiner Mutter, die stets für mich da war, mir den Umzug nach Kiel nicht übel genommen hat und bei der ich immer meine Batterien wieder aufladen konnte, und bei meinem Freund, der mir gezeigt hat, dass das Leben zu zweit signifikant mehr Freude bereitet und der durch konstante Motivation und Unterstützung in nicht unerheblichem Maß dazu beigetragen hat, dass diese Arbeit fertig wurde.

Herzlichen Dank!