INCLUSIVE SCIENCE EDUCATION

THE FRAMEWORK FOR INCLUSIVE SCIENCE EDUCATION

2ND EDITION

WORKING PAPER
IN
INCLUSIVE SCIENCE EDUCATION

1/2020

→ www.leuphana.de/inclusive-science-education

THE FRAMEWORK FOR INCLUSIVE SCIENCE EDUCATION

KATEGORIENSYSTEM INKLUSIVER NATURWISSENSCHAFTLICHER UNTERRICHT (KinU)

Sarah Brauns, Simone Abels

Authors

Sarah Brauns, M.Ed.
Science Education I Institute of Sustainable Chemistry (INSC)
Leuphana University Lüneburg
Universitätsallee 1, 21335 Lüneburg
Sarah.brauns@leuphana.de

Prof. Dr. Simone Abels
Science Education I Institute of Sustainable Chemistry (INSC)
Leuphana University Lüneburg
Universitätsallee 1, 21335 Lüneburg
Simone.abels@leuphana.de

Imprint
Prof. Dr. Simone Abels
Science Education
Leuphana University Lüneburg
Universitätsallee 1
21335 Lüneburg, Germany

ISSN 2701-3766
Published online 10th September 2020
2nd Edition published online 25th Februrary 2021
Issue 1/2020
Available at www.leuphana.de/inclusive-science-education

In the 2nd edition of the Working Paper on the Framework for Inclusive Science Education, the reading direction of the framework was mainly changed. In the previous version, the main categories were on the right side to make clear that the categories were originally derived from the literature. In order to better use the framework deductively in research and as a guide for lesson planning, the main categories are now on the left side, so that the categories become more and more concrete towards the right up to the subcode level.
In addition, the framework abstraction levels (p. 13) were no longer presented as a continuum because the levels are discrete. The listing of subcategories (p. 21 f) was adapted to include different levels of abstraction that were already present as subcategories in the framework. For the sake of structure and clarity, tables of contents, figures and tables have been added.

Citation
Brauns, S., & Abels, S. (2021). The Framework for Inclusive Science Education. *Inclusive Science Education, Working Paper No. 1/2020* (2nd ed.). Leuphana University Lüneburg, Science Education.

Collaboration and Funding

Bundesministerium
für Bildung
und Forschung

Förderkennzeichen 01NV1731
Laufzeit: 01.04.2018-31.03.2021

THE FRAMEWORK FOR INCLUSIVE SCIENCE EDUCATION

Sarah Brauns, Simone Abels

The intention of this working paper is to provide the theoretical and methodical background of the Framework for Inclusive Science Education[1]. While other publications of the authors only present selected parts of the framework, readers will find a quotable complete version of the framework in this working paper. Thereby the framework and its design are made comprehensible and transparent. Practitioners and researchers may use the framework or parts of it for their projects, but only with appropriate citation. To cite the current version of the framework: Brauns, S., & Abels, S. (2020). The Framework for Inclusive Science Education. *Inclusive Science Education, Working Paper No. 1/2020*, Leuphana University Lüneburg, Science Education. The year and version number have to be updated eventually.

ABSTRACT

In this working paper we introduce the Framework for Inclusive Science Education. For the data collection, we applied a systematic literature review. In the process, n=297 titles were generated, which empirically or theoretically address the issue of inclusive science education. The sample was analysed both qualitatively and quantitatively. In a qualitative analysis, categories that combine characteristics of science education with an inclusive implementation were inductively derived. In total, n=935 categories on different abstraction levels were derived, which represent the framework. N=16 main categories were identified, which display the characteristics of science education to be combined with inclusive pedagogies. For the quantitative analysis of the sample and the framework, descriptive statistics were performed and differences between sub-samples analysed. Over the last ten years, a significant increase in publications has been observed. Moreover, there is a minor representation in titles relating to pre- and in-service teachers working in inclusive science education. Overall, in this paper we present not only the framework itself, but also give recommendations for the application of the framework.

[1] The Framework for Inclusive Science Education is available on the pages 48-75, appendix E, and the German version on the pages 76-106, appendix F.

CONTENT

1. Introduction 1
2. Methodical approach 4
    2.1 Research question 5
    2.2 Data collection 5
    2.3 Data analysis 8
        2.3.1 Quantitative analysis of the sample 9
        2.3.2 Qualitative analysis of the sample 10
        2.3.3 Quantitative analysis of the framework 16
3. Results 16
    3.1 Descriptive presentation of the sample 17
    3.2 The Framework for Inclusive Science Education 21
    3.3 Descriptive presentation of the framework 23
4. Discussion of and implications for the application of the framework 27
5. Validation of the framework and outlook 32
Acknowledgements 34
References 34
Glossary 38
Appendix 39
    Appendix A: Search Protocol for the first search of data 39
    Appendix B: Search Protocol for the second search of data 40
    Appendix c: Search Protocol for the third search of data 41
    Appendix D: Definitions of the main categories 42
    Appendix E: The Framework for Inclusive Science Education 48
    Appendix F: KinU - Kategoriensystem inklusiver naturwissenschaftlicher Unterricht 85
    Appendix G: List of literature of the systematic review 127

FIGURES

Fig. 1. Search strategy..6
Fig. 2. Levels of abstraction of the framework..13
Fig. 3. Argumentative validation of the inductive category deriving.........................15
Fig. 4. Empirical and theoretical titles distributed over the years. Significant differences are marked with asterisks (p < .05)..18
Fig. 5. Distribution of the number of titles referring to a wide concept of inclusion and to additional educational needs...19
Fig. 6. Distribution of the focus groups in the sample universe..............................19
Fig. 7. Main categories of the framework...21
Fig. 8. Number of empirical and theoretical categories on different levels of abstraction. Significant differences are marked with asterisks (p < .05)..........24
Fig. 9. Main categories with the highest number of categories.............................25
Fig. 10. Distribution of the empirical and theoretical paraphrases on the levels of abstraction of the framework. Significant differences are marked with asterisks (p < .05)..26
Fig. 11. Main categories with the largest number of paraphrases..........................26
Fig. 12. Validation process of the framework and project procedure within Nawi-In....32

TABLES

Tab. 1. Search terms for the first, second, and third search of data..........................7
Tab. 2. Variables and scales of the quantitative analysis of the sample.....................9
Tab. 3. Deriving categories from the publications...12
Tab. 4. Clustering the paraphrases to form categories (example on subcode level)......14
Tab. 5. Development of the focus on teachers and student teachers in publications on inclusive science education...20

1. INTRODUCTION

Inclusive science education describes the connection between science education and inclusive pedagogy. One approach to present this connection is to clarify what we mean by science education and what understanding of inclusion we follow. The overarching aim of science education is to achieve "Scientific Literacy for all learners" (Bybee, 1997, p. 69). Scientific Literacy, as it is defined by the OECD (2019), is divided into three areas: content knowledge, procedural knowledge and epistemic knowledge. Firstly, content knowledge is subsumed as "explaining phenomena scientifically" and is described as "knowledge of the facts, concepts, ideas and theories about the natural world that science has established" (OECD, 2019, p. 99f). Secondly, the procedural knowledge is summarised as "evaluating and designing scientific enquiry" and is described as "concepts on which empirical enquiry is based, such as repeating measurements to minimise error and reduce uncertainty, the control of variables, and standard procedures for representing and communicating data" and "concepts of evidence" (OECD, 2019, p. 99f). Thirdly, epistemic knowledge is subsumed as "interpreting data and evidence scientifically" and is described as the "understanding of the role of specific constructs and defining features essential to the process of building scientific knowledge [...] [and the] understanding of the function that questions, observations, theories, hypotheses, models and arguments play in science; a recognition of the variety of forms of scientific enquiry; and understanding the role that peer review plays in establishing knowledge that can be trusted" (OECD, 2019, p. 99f). To take all learners into account, we follow a wide concept of inclusion that manifests the participation of all students in education independent of their diversity characteristics in abilities, age, ethnicity, gender, sexual orientation, religion and other (UNESCO, 2005; Werning, 2014). In practical implementation, however, a narrow concept of inclusion is often pursued, i.e., that mostly differentiation measures are designed for students with additional educational needs or migration background instead of providing learning opportunities which, due to their open design and self-determination, enable participation for all students without prior categorisation in the sense of stigmatisation (Florian & Black-Hawkins, 2011). The narrow understanding of inclusion is also reflected when only diversity dimensions of difference such as achievement potential or disability are used as labels (e.g., Scruggs, Mastropieri, & Okolo, 2008; Therrien, Taylor, Watt, & Kaldenberg, 2014). Particularly critical are publications that state a wide understanding of inclusion, but do not redeem this understanding in terms of empirical and/or practical application. Furthermore, there can be a third understanding necessary concerning all learners, but especially concerning vulnerable groups (Lindmeier & Lütje-Klose, 2015).

Thereby the focus is widened on marginalised learners, not only students with disabilities, but on all vulnerable groups who run the risk of being discriminated (ibid.) following the idea of "Education for all, and especially for some" (UNESCO, 2005).

"Up to now, a scheme which systematizes and combines aspects of inclusive pedagogy and science education is still missing" (Stinken-Rösner et al., 2020, p. 30). With this introductory statement, Stinken-Rösner et al. (2020) draw attention to the fact that although discussions about the implementation of inclusion in schools have increased in recent years, the combination of inclusive teaching in subject matter education disciplines is still not fully established. Therefore, they developed a theoretical scheme for linking the perspective of inclusive pedagogy (acknowledging diversity, minimizing barriers and enabling participation, cf. Booth & Ainscow, 2016; UNESCO, 2005) and the perspective of science education (reasoning about scientific issues, learning science content, doing science, learning about science; cf. Hodson, 2014). This theoretical scheme shows possible connections between the demands of inclusion and science education. In this way, the scheme is intended to advance the basis of future research and lesson planning with the two perspectives thought together.

The approach of combining science education with inclusive pedagogy is also evident in the definition of inclusive science education by the members of the Network of Inclusive Science Education (German: Netzwerk inklusiver naturwissenschaftlicher Unterricht (NinU)):

> "Science education fosters inclusion by facilitating participation in science specific learning processes for all learners. By appreciating the diversity and individual prerequisites, science education involves individual and joint teaching and learning processes to promote scientific literacy" (Walkowiak, Rott, Abels & Nehring, 2018, p. 269).

On the one hand, this definition implies that inclusive approaches are compatible with science education. On the other hand, science educators would argue that on this general level the relation between inclusion and science education is not concrete enough. The understanding of science specific learning processes and the concretisation on how these connect to inclusive pedagogy are not explicit. Therefore, what is required are indications for action, specifying how science education can be implemented in an inclusive way. For this reason, we conducted a systematic literature review in order to establish a framework that can provide evidence how to implement inclusive science education. It is called the Framework for Inclusive Science Education (German: Kategoriensystem inklusiver naturwissenschaftlicher Unterricht (KinU)).

This framework provides the basis for our research in the Nawi-In project (Teaching Science Education inclusively (German: Naturwissenschaftlichen Unterricht inklusiv gestalten (Nawi-In)), which is funded by the German Ministry of Education and Research (no. 01NV1731). In this project, we investigate the development of student teachers competency profiles with regard to inclusive science education. More precisely, we evaluate student teachers' competency development in the first two of three semesters during a Master's program, which includes a research-oriented seminar. They gain theoretical foundations for inclusive science education and practice their noticing and reasoning skills through the analysis of science lesson video clips in the first of the three semesters (Sherin, 2007; Seidel, Stürmer & Schäfer, 2013; Stürmer, Seidel & Schäfer, 2013). Afterwards in the second semester the student teachers complete an school internship and conduct their own science lessons, which they also videotape and reflect. During the third semester, they analyse their own videos and present their results. The competency profiles of the student teachers are established by combining questionnaires on self-esteemed knowledge, attitudes and self-efficacy, the use of video reflections from other teachers' lessons, video analyses of classroom activities and video reflections of their own lessons (see outlook) (Brauns, Egger & Abels, 2020; Egger, Brauns, Sellin, Barth & Abels, 2019).

The concept of competency represents a complex construct, which can be related to larger factors and is conditioned by various influences such as bias, self-efficacy and motivation. For this reason, this study investigates the professional competency in the classroom and analysis competency on practice in the context of inclusive science education. In a very brief summary, competency is „the personal capacity to cope with specific situational demands" (Kunter et al., 2013, p. 27). In our project, we focus on two areas of competency: Professional Vision (e.g. Sherin, 2007) and Professional Knowledge (e.g. Baumert & Kunter, 2006; Baumert & Kunter, 2011). Professional Vision describes a component of teaching expertise and can serve as an indicator for conceptual knowledge (Stürmer et al., 2013). For this purpose, teaching videos are often used for reflection (ibid.). Overall, Professional Vision is divided into two areas: Noticing and Knowledge-Based Reasoning (Seidel et al., 2011). The first evaluates what the student teachers see and observe, i.e., notice in teaching videos and the second evaluates how they interpret what they have noticed before (ibid.). With the help of the framework, it is among other possible to analyse the student teachers' noticing abilities when reflecting on teaching videos. More specifically, inclusive scientific characteristics, which are noticed by the becoming teachers in their own and other teachers' lessons, can be analysed. In addition, the framework will be applied to the videotaped teaching of the student teachers in order to analyse the teaching activities in school practice. Regarding professional knowledge, the framework can be used to enrich the idea of pedagogical content knowledge with inclusive science characteristics which can be applied in practice.

On this basis, we set up the Framework for Inclusive Science Education that has shown to be very extensive. On the one hand, the methodological approach has been carried out comprehensively and, on the other hand, the framework itself is very large with a total of 935 categories[2]. The framework increases in quality by presenting the methodical procedure of literature search and selection in detail. This working paper is intended to provide a guide that leads through the structure of the entire framework and gives references on its application in practice and research. The advantage of the working paper is the possibility for continuous update of the framework and a full picture of the status quo.

2. METHODICAL APPROACH

> "A research literature review is a systematic, explicit, and reproducible method for identifying, evaluating, and synthesizing the existing body of completed and recorded work produced by researchers, scholars, and practitioners" (Fink, 2009, p. 3).

With this definition, Fink (2009) summarises the characteristics of the systematic procedure of a literature review as a method for data collection. Following a strict methodical procedure in a systematic literature review and discussing the selection of literature are meant to lead to reduce the bias of the researcher (Feak & Swales, 2009). Moreover, the procedure is made transparent for other researchers. The procedure of Fink (2009), which is followed by this systematic literature review, contains seven steps:

1. Selecting a research question
2. Selecting bibliographic or article databases
3. Choosing search terms
4. Applying practical screening criteria
5. Applying methodological screening criteria
6. Doing the review
7. Synthesising the results

[2] For the use of certain terms a glossary is available on page 38.

The first step, selecting the research question, is particularly important in order to set a research focus for the review. Our systematic literature review focuses on the identification of characteristics of science education that are implemented inclusively in the classroom.

For the data analysis, categories are constructed inductively from the sample of the systematic literature review with the qualitative content analysis via focused summary according to Kuckartz (2016). The MAXQDA (version 20.0.7) software is used for technical support. The goal is to design a framework that structures the whole sample into categories of inclusive science education.

2.1 RESEARCH QUESTION

With the systematic literature review we answered the following question: What are the characteristics of inclusive science education. In order to show the connection between science education and inclusive pedagogy, we structure examples of inclusive science teaching following the characteristics of science education. The answer to the research question will be fundamental for the analysis of the subquestion, which characteristics of inclusive science education student teachers reflect on and show in teaching. To answer this subquestion the framework will be applied to the data as analysis tool.

2.2 DATA COLLECTION

In order to make the data search procedure comprehensible, it is presented here in detail (Fig. 1). The structure results from the fact that the search has been repeatedly revised. After the first search (left strand in Fig. 1), it became apparent that important publications by authors from the field of inclusive science education were not present in the sample, therefore the search terms were refined (middle strand). The third search was carried out at the beginning of 2020 in order to keep the literature as up-to-date as possible.

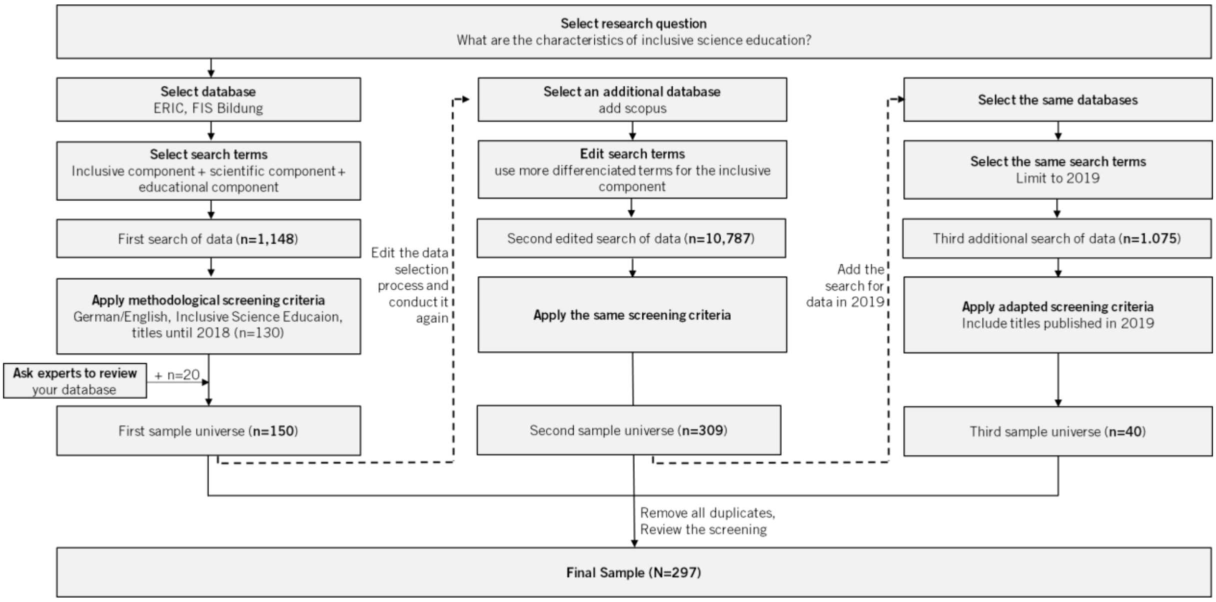

Fig. 1. Search strategy

The data collection relates to the research question what the characteristics of inclusive science education are. The databases "ERIC" and the German equivalent "FIS Bildung" were used in the first search. Both databases contain specific literature from the pedagogical field. In the second search, the database "Scopus" was added to further increase the search radius and to include respective authors of the field. All three databases were used in the second and third search.

In the next step, the search terms were selected (Tab. 1). The search terms always consisted of an inclusive, a scientific AND an educational component. Initially, in the first search only the term "inclus*" was used as the inclusive component, consequently, a substantial part of the inclusive scientific literature was not listed as those often operate with terms like heterogeneity or integration. The term "exclusion" was added to list the opposite and to lead to literature in an inclusive context. The search terms should be chosen carefully, because they essentially determine whether titles are listed in the literature search or not. In this case, it can be reflected that despite the revision of the search terms, words like diversity or equity are still missing, which could especially increase the international hits. The reason why, despite these missing search terms, some titles with the word "diversity" (e.g., Markic & Abels, 2014; Watt, Therrien & Kaldenberg, 2014; Nawarathne, 2019) appeared in the sample is probably, because keywords were specified for these titles, which were taken into account in the data search. The science and educational components were slightly adapted in the second and third search (Tab. 1).

Tab. 1. Search terms for the first, second, and third search of data

Language	Inclusive Component	Scientific Component	Educational Component
English	Inclus*[1, 2, 3] Heterogen*[2, 3] Integrat*[2, 3] Exclus*[2, 3]	Science[1, 2, 3] Natural science[1] Early science education[1] Chemistry[1, 2, 3] Biology[1, 2, 3] Physics[1, 2, 3]	Learning[1, 2, 3] Class[1, 2, 3] School[1, 2, 3] Primary[1, 2, 3] Secondary [1, 2, 3]
German	Inklus*[1, 2, 3] Heterogen*[2, 3] Integrat*[2, 3] Exklus*[2, 3]	Naturw*[1, 2, 3] Sachunterricht[1, 2, 3] Chemie[1, 2, 3] Biologie[1, 2, 3] Physik[1, 2, 3]	Unterricht[1, 2, 3] Primar*[1, 2, 3] Grundschule[1, 2, 3] Sek*[1, 2, 3] Weiterführende Schule[1]

([1]=first search, [2]=second search, [3]=third search)

When using the search terms, the combinations of the three components were formulated according to the search procedure of the databases, e.g., ERIC needs search terms like follows: "Inclusion AND science AND (learning OR class OR school OR primary OR secondary)".[3] In order to increase the quality of the data generation, only peer reviewed publications were considered in ERIC and Scopus. Furthermore, the ERIC and Scopus databases allow for searching in specific journals. We chose journals that were focused on science education or inclusive education so that the amount of data was feasible with the purpose of our inquiry. While the first two searches for publications were conducted until 2018, the third search was carried out once only for 2019 due to a turn of the year.

If we compare the publications found in the three searches, we count n=1,148 titles for the first data search, n=10,787 titles for the second search and n=1,075 titles for the third search. From the first to the second search, there is an enormous increase in the number of hits. This can result from different reasons. Especially, the terms of the inclusive component are often used in other contexts. The terms of this component may refer to processes at the molecular level, e.g, the integration of atoms into molecule structures, or to several meanings in the school context (e.g. Abramova, Shilova, Varankina & Rubanova, 2019; Abdella, Walczak, Kandl, & Schwinefus, 2011; Bardeche et al., 1980). In German, the word "integrative" also means the combination of chemistry, physics and biology into one subject called natural sciences. In this way, the revision of the search terms for the second search not only collected literature of the target group, but also literature that was eliminated again in the next step.

The search criteria were essentially the same in all three searches: time frame, language, school type, focus on inclusive science education. The only change that was made is that the first two searches include all titles until 2018 while the third search was restricted to 2019. A targeted search was conducted for publications in English

[3] The search strings and the exact search history is available on the pages 39ff, appendix A, B, C.

and German. All titles that explicitly refer to inclusive science education were included. Titles that explicitly refer to school types other than primary and secondary education were excluded. Similarly, when it came to the inclusion of teacher students at the university level into a scientific subject, the titles were sorted out (diversity sensitive teaching, e.g., Godovnikova, Gerasimova, Galchun, & Shitikova, 2019; Ghanbari, 2015; Alheit, 2009; Fraser, Giddings & McRobbie, 1992). When implementing inclusion, we refer to science education on school level and the inclusion of students in school science education. This is equally the case, when student teachers are taught at university level how to implement inclusion in science education at school (e.g., Brauns, Egger, Abels & Barth, 2019; Benny & Blonder, 2018; Kahn, Pigman & Ottley, 2017; Abels & Koliander, 2014). As a result, titles were retained in the sample if, for example, they dealt with teaching concepts or materials without specific reference to school levels.

In the first step of the selection process, the samples were selected according to the titles, then the abstracts were read and selected. A total of n=130 titles were identified in the first search, the sample was reviewed by experts from the network of inclusive science education (NinU) and extended by n=20 titles. Overall, the selection process resulted in n=150 titles for the first sample, n=309 titles for the second sample and n=40 titles for the third sample. Since all search strands were carried out separately, they were then merged. All duplicates were removed and the screening was once more reviewed. Finally, all publications were read completely for the later qualitative content analysis, which resulted in further eliminations. This complete data search and selection resulted in a final sample of n=297 titles[4].

2.3 DATA ANALYSIS

All titles in the sample were evaluated both quantitatively and qualitatively. The quantitative analysis not only describes the composition of the sample, but also shows the current state of research, which is represented by the systematic literature review. The aim of the qualitative analysis, which is the main focus of the literature analysis, was to derive categories from the literature to summarise characteristics of inclusive science education. All titles were evaluated with the qualitative content analysis via focused summary according to Kuckartz (2016) by inductively constructing categories from the data material. The basic approach is that paraphrases were derived from text passages, which were gradually abstracted more and more until they formed categories (Kuckartz, 2016). After the qualitative analysis of the literature resulting in the framework, a further quantitative analysis was carried out. The results of this can, for example, provide

[4] The literature list is available on the pages 107-126, appendix G.

information on the main focus of the literature and which titles are most frequently cited in the framework.

2.3.1 QUANTITATIVE ANALYSIS OF THE SAMPLE

For the quantitative analysis of the sample, data was collected and evaluated using the statistics software SPSS (version 25) with the following variables: publication language, publication type, conception, school type, focus group(s), year of publication and diversity dimension(s) (Tab. 2). All measurement levels were nominally distributed.

Tab. 2. Variables and scales of the quantitative analysis of the sample

Variables	**Scale**
Publication language	1=German, 2=English
Publication type	1=journal article, 2=chapter, 3=monograph
Conception	1=empirical, 2=theoretical
Type of school	1=primary school, 2=secondary school, 3=indefinable
Focus groups	1=students, 2=teachers, 3=student teachers, 4=indefinable
Year of publication	1=before 1990, 2=1990-1994, 3=1995-1999, 4=2000-2004, 5=2005-2009, 6=2010-2014, 7=2015-2019
Diversity dimensions	1=wide concept of inclusion, 2=ethnicity (culture), 3=socio-economic status, 4=gender, 5=organisational role, 6=additional educational needs, 7=age, 8=religion, 9=sexual orientation, 10=language, 11=highly gifted

Concerning the type of school, it should be noted that the scale element 'indefinable' was used when the school form could not be clearly assigned. The fact that these titles were included in the sample is due to the fact that titles were only excluded in the selection process if they explicitly referred to a school form other than primary and secondary school. If no clear assignment to a school level was possible, the titles were retained in the sample. For the variable focus groups, an assignment to 'indefinable' is also possible. There were publications in which the protagonists of instruction were the central theme. However, it could also happen that educational models, teaching materials or teaching concepts were presented in the literature. In these cases, the titles were assigned to 'indefinable'. The diversity dimensions are essentially based on the "Big 8" (Krell, Riedmüller, Sieben & Vinz, 2007) and have been expanded by the wide concept of inclusion, language and high talent. The wide concept was used for titles that refer to the inclusion of all students rather than to individual diversity dimensions. This variable sets a contrast to the narrow concept of inclusion associated with additional educational needs (see introduction). The understanding of vulnerable groups was not explicitly considered as it is not yet common in science education. Language has been considered separately from ethnicity because linguistic diversity is not always the cause of ethnical differences and in literature it is usually considered as a single factor. Whereas in the "Big 8" ability is listed as a facet, we have distinguished between the additional education focus and giftedness, because in the literature of our sample both are dealt with differently and even imply a contradiction.

Several means were used to assure the quality of the framework. Starting, the first 10 % of the literature were evaluated by a second trained coder. At that time the intercoder reliability in terms of Cronbachs alpha was 0.67. This dissatisfactory agreement may result from the fact that for the first part of the sample only the abstracts were analysed quantitatively. Both coders discussed each coding together along the data material. Afterwards, the whole sample was quantitatively analysed by one researcher. This led to the decision to read the texts completely for the analysis. All n=297 titles of the sample were read and coded according to the manual (Tab. 2). Finally, 10 % of the literature was again randomly selected and evaluated by the same second trained coder as above. The renewed intercoder reliability was 0.84. This result can be rated as good, but was not entirely satisfying for us. It became apparent that the variable of diversity dimensions caused this difference in the analysis. Therefore, all codes of this variable were reviewed and revised again. The difficulty in assigning diversity dimensions often lies in the fact that authors focus on one dimension but theoretically justify it with another dimension. As an example, a study investigates students with ethnic background. However, the theory in such an article describes the understanding of inclusion with a focus on special educational needs (e.g., Koomen, 2016). The reference to special educational needs is also sometimes made, although the authors explicitly mention that they refer to a broad understanding of inclusion. In these cases, it was necessary to discuss which diversity dimensions were actually addressed in the article.

2.3.2 QUALITATIVE ANALYSIS OF THE SAMPLE

The focus of the qualitative analysis is identification of categories for the Framework for Inclusive Science Education. We used the qualitative content analysis via focused summary according to Kuckartz (2016). Altogether, six guidelines serve as an orientation frame for the inductive category formation (Kuckartz, 2016):

1. Determine the goal of category formation on the basis of the research question
2. Determine category type and level of abstraction
3. Familiarise yourself with the data and determine the type of encoding unit
4. Edit the text sequentially and create categories directly on the text; Assign existing categories or create new ones

5. Systematising and organising the category system
6. Define the category system

In the following, the methodical procedure of the inductive category formation will be described in detail.

(1.) It is important to phrase a clear research question in order to be able to set a focus in the analysis. In our case, the research question was which characteristics of inclusive science education are suggested in literature. It is also relevant that a certain amount of prior knowledge of the subject area is available to the researcher so that passages in the text are considered relevant to the research question (Kuckartz, 2016). Based on the research goal, inductive categories should be formed that represent an aspect of science that has been combined with an aspect of inclusion. One can also ask the question of how a characteristic of science education is implemented in an inclusive way. In this case, the coders needed to be aware of what is specific about science education. It was necessary to know the scientific subject and to be able to distinguish it from other school subjects. When authors describe in the text how a characteristic of science education is implemented in the classroom, we first assumed that the implementation is meant to be inclusive, when the publication generally refers to inclusive education. However, at a later stage we will need to validate which categories actually contribute to inclusive science education.

(2.) The next step is to determine the level of abstraction of the categories. The goal of category formation was to construct categories that allow for precise statements about how exactly science education should be implemented in an inclusive manner. We wanted to obtain information that answer the question of what exactly can be done by (becoming) teachers in inclusive science education. First of all, codings were marked in the text and adopted as quotations of the text passages. From these codings, the paraphrases were made, which were as close as possible to the wording of the original text. Already the paraphrases were formed in such a way that they consisted of a characteristic of science education in combination with an infinitive representing means of inclusive pedagogy. This kind of construction ran not only through all paraphrases, but also through all categories. The example coding (Tab. 3) shows a coded section of text dealing with a thermometer that emits sounds and vibration pulses. The use of the thermometer represents the characteristic of science education, a scientific investigation method, which is adapted here inclusively. This example also shows that two aspects can be addressed in one sentence or with reference to the same characteristic of science education. In these cases, the individual aspects are each listed in their own paraphrase in order to be able to classify them later in the framework. An italic coding in the framework means that the code is from empirical findings in the literature. Empirical findings have been tested as effective for inclusive science education. The paraphrases derived from this were then also marked in italics, so that it is clear which paraphrases originate from empirical results and which originally come from a theoretically formulated text passage.

Tab. 3. Deriving categories from the publications

Coding	**Paraphrase**
„[…] The device is suitable for temperature measurements in degrees Celsius, emitting sounds and vibration pulses similar to morse code, with a measuring scale ranging from -15°C up to 115°C. This thermometer will give the opportunity to participate actively in the acquiring knowledge process […]“ (Vitoriano et al., 2016, p. B)	*Providing a device for temperature measurements emitting sounds similar to morse code* (Vitoriano et al., 2016, p. B) *Providing a device for temperature measurements emitting vibration pulses similar to morse code* (Vitoriano et al., 2016, p. B)

(Colour code: characteristics of science education, aspects of inclusive pedagogy, paraphrases with the focus on secondary education; *italic paraphrases are derived from empirical studies)*

(3.) The coding units were determined by one aspect occurring in a paraphrase of a coded text passage. The first objective is to choose a coding unit that is as small as possible to allow a paraphrase to be derived from it. This goal is always coupled with the condition that the coded passage must be coherently understandable for the coders and readers. This means that at least one-half sentence must be coded. This rule results from the fact that MAXQDA (version 20.0.7) is used as analysis software. After coding, this program lists all coded passages in a table separate from the original text. The coders must decide during the coding process whether the meaning of the coded unit can be understood without context. Therefore, a section with one aspect is to be coded at most. All parts except the methodological part were coded in each publication and of the final sample all titles were coded.

(4.) The coding procedure will now be determined. In the citation program Citavi, the literature in the sample was arranged by year, starting with 2019, and within years alphabetically. Due to the fact that the sample for 2019 was generated later, the first analyses were performed starting with 2018 and according to the alphabetical order. According to Mayring (2000), it is recommended to first analyse 10 % to 50 % of the sample using the inductive procedure. We started coding text passages with 30 % of the sample. Kuckartz (2016) states that the categories from the coding of the first part of the sample are usually applied to the further data in a deductive way. In our case, it turned out that we had to deviate from Kuckartz's (2016) methodical approach. Already after the analysis of 30 % of the sample, the preliminary framework was already very comprehensive, so that it was not practicable to apply it to further data material. For this reason, the remaining part of the sample was then analysed using the same inductive procedure. The inductive procedure had further advantages for the later quantitative analysis of the framework as well as for the intersubjective transparency of the methodical procedure. The focus was also on being able to trace a reference to each source even later. MAXQDA then listed a table with all coded text sections. Paraphrases were formed manually from each coding (Tab. 3). In this way, the wording of the original text could be preserved. This information is important as it demonstrates how abstract or how concrete inclusive science education is explicated in the literature. Our assumption was that the empirical works represent the connection between science education and inclusion more concretely than the theoretical works, but also that inclusive science education is described rather superficially and not very concretely at all. The consequence of the diversity of the literature was that a saturation did not occur in the analysed 30 % of the sample. In order to minimise the gaps in the framework and to be able to evaluate the contents of the literature quantitatively at the end, all other titles in the sample were analysed in full.

(5.) When the categories formed become gradually unclear and when hardly any more categories are found, the derived categories should be structured in a framework (Kuckartz, 2016). In our case, the first structure of the framework was formed after analysing the first 30 % of the sample. First, all paraphrases formed were clustered according to the superordinate characteristics of science education. In this way, 16 characteristics of science education were identified, which serve as main categories in the framework. In the next step, the paraphrases were sorted according to their degree of abstraction within each main category. In this way, the four levels of abstraction of the framework were created: subcode, code, subcategory, main category (Fig. 2). On the main category level, the characteristics of science education are presented and written together with "implementing inclusively" or similar. On the subcategory level, the type of the inclusive implementation of the characteristics of science education is summarized. Up to the subcode level these implementation suggestions become more and more concrete, but on the subcategory and code level they leave open the question of the concrete implementation. It is only at the subcode level that concrete instructions for teaching inclusively are given, leaving no questions unanswered.

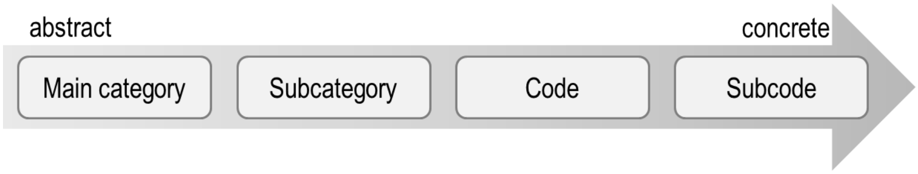

Fig. 2. Levels of abstraction of the framework

Within the levels of abstraction, clusters were formed. In the example above in Table 3, a paraphrase from the text to a device for temperature measurements emitting sounds was derived from the coding. Similar paraphrases could have been formed from codings of other sources. These similar paraphrases were then combined into groups. In the process a cluster of different paraphrases was formed. In Table 4, all paraphrases refer to thermometers that make sounds. Each cluster has then been given a heading. In this example it is "Enabling the application of scientific investigation methods with acoustic thermometers". This heading is later adopted for the framework and forms a category. In this case, the category is concretely formulated and is therefore at the subcode level. In this way, a first structure of the framework was created. All categories were assigned to the four different levels of abstraction depending on how concrete the implementation suggestion was (Fig. 2). When assigning the categories, it was noticeable that a similar structure was created at the subcategory level throughout the framework. Consequently, care was taken to ensure that the structure of the framework is as uniform as possible at the subcategory level. After this first procedure, Kuckartz (2016) suggests asking oneself how many categories are reasonably needed for the analysis, and to include economic factors as well as the goal of the research when answering the research question. Due to the large number of categories formed, the use of the framework for deductive analysis would have hardly ever been possible. One solution could have been to combine the categories to such an extent that the scope of the framework would have been reduced. Even if a maximum of ten main categories is recommended (Kuckartz, 2016), it was not expedient and not in line with the theoretical background to further summarise the 16 identified characteristics of science education. Furthermore, we found no concrete guidelines for inclusive science education in the literature sample. Definitions that were written at a general level tended to leave open the question of how exactly inclusive science education could be implemented. For this reason, the economic factor was rejected as the provision of concrete recommendations for action outweighed it. Therefore, inductive paraphrases were formed from the complete remaining sample, which were either added to existing categories or formed new categories.

Tab. 4. Clustering the paraphrases to form categories (example on subcode level)

Enabling the application of scientific investigation methods with acoustic thermometers
Providing a device for temperature measurements emitting sounds similar to Morse code, with a measuring scale ranging from -15 °C up to 115 °C (Vitoriano et al., 2016), Providing talking thermometers (Koehler & Wild, 2019), Providing alteration of common laboratory measurement devices for successful independent use by the visually impaired such as the substitution of talking thermometers for traditional visual thermometers (Watson & Johnston, 2007), *Providing a thermometer that provides information through beep sounds (Vitoriano et al., 2016),* Providing audible electronic to understand temperature thermometers (Teke & Sozbilir, 2019)

(Colour code: sources with the focus on secondary education; *italic paraphrases are derived from empirical studies)*

(6.) After arranging the categories within the first draft of the framework, definitions of the categories were established.

In a repeating review procedure by groups of researchers and revision by the authors, the structure of the framework was developed further until the current framework was created. It cannot be ruled out that, with advancing application of the framework and because of the vivid research field, it will be further revised. Reviewing and revising in the communicative process is a typical procedure in inductive category building to ensure the quality of the methodical approach. When it comes to the quality of the framework, Kuckartz (2016) points out that a distinction must be made between the creation of the framework and the application of the framework. When creating the framework, it is not possible to generate a perfect match between the coders. Therefore, all substeps of the inductive category deriving were optimised in ways of argumentative validation (Bortz & Döring, 2016) (Fig. 3).

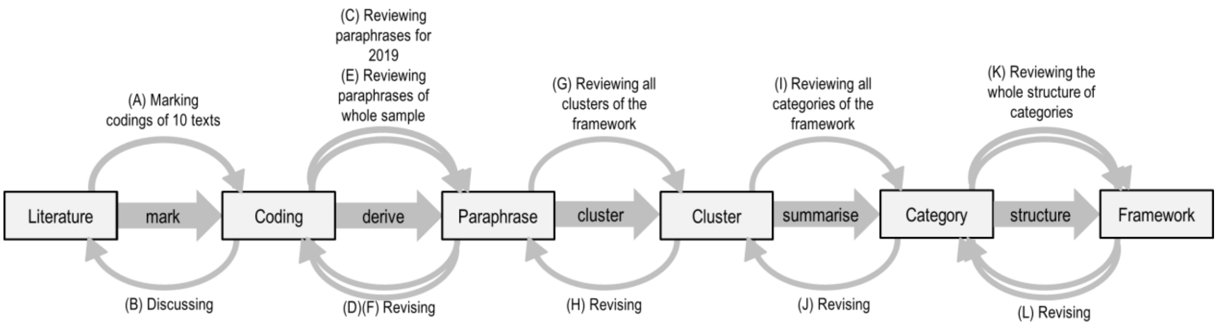

Fig. 3. Argumentative validation of the inductive category deriving

The validation process of the category deriving is divided into twelve steps. (A) Ten randomly selected publications were coded by two coders. This means that the two coders inductively marked codings in the text that relate to inclusive science education. The aim was to find similarities and differences in the identification of the characteristics of science education and to make sure that an inclusive implementation was marked for these characteristics of science education. Afterwards, the codes were discussed, the coding units were defined particularly clearly and it was also brought to mind once again that the emphasis in the coding should be on the characteristics of science education in order not to code generally inclusive statements and not to lose the subject-specific focus. (B) After the paraphrases had been formed for the 2019 sample, (C) the process of summarising a coding to a paraphrase was reviewed by the second researcher. Overall, these paraphrases were used to review the first 25 % of all paraphrases derived from the literature of the entire sample. The aim was to ensure that the paraphrases were specific to the subject and particularly close to the wording of the original text. (D) All critical points were discussed and revised. These discussions were taken into account in the further procedure. (E) Then the paraphrases from all codings were derived, and this whole process was again reviewed by the second researcher. (F)

All irregularities were subsequently revised. (G) A similar approach was followed for allocating the paraphrases to the clusters. An argumentative process was chosen again, and this time again the second researcher reviewed all clusters of the complete framework with the 1627 paraphrases. (H) Changes resulting from the discussions were incorporated into the clusters. This means that paraphrases that did not belong to one group were added to other groups or new groups were created. (I) Each cluster received its own headline and these category names were also reviewed, discussed and (J) revised in dialogue. (K) In the final step, all categories and their assignments to the respective subcodes, codes, subcategories and main categories were discussed and (L) revised by experts in this field and within the Nawi-In project. Some of the revision processes were carried out in cycles in order to continually optimise the framework. In this final revision process the inductive procedure overlapped with a deductive procedure. In order to make the framework transparent, we created a structure that is recurrent. Therefore, there is an interaction between the inductive and the deductive procedure. For this purpose, the wording was deductively standardized on the subcategory and code level. On the subcategory level, a uniform structure was chosen, which was adopted for all main categories, and on the code level, the wording can also be found under the various subcategories.

2.3.3 QUANTITATIVE ANALYSIS OF THE FRAMEWORK

After the framework was created, the categories and the paraphrases were each quantitatively analysed. The categories were evaluated by counting how many categories are listed in the respective abstraction level and in the different main categories. By counting the paraphrases, it can be determined which titles of the sample were cited most often in the framework and which categories or characteristics of science education are most often addressed in the literature. The categories were counted manually, while the citations were both counted through Citavi and checked manually. Using Citavi has the advantage that the program displays the number of citations per literature source automatically. The significances were calculated here with the Qhi-squared test and an alpha $\alpha = .05$ as in the quantitative analysis of the sample.

3. RESULTS

The presentation of the results is divided into three areas. First, the distributions of the sample are presented descriptively. Then, the presentation of the qualitative results, on which the focus of this paper is, are described. In this part, the Framework for Inclusive Science Education is not only presented, but also references for its application and adaptation in further research are given. Finally, the results of the quantitative analysis of the framework are presented.

3.1 DESCRIPTIVE PRESENTATION OF THE SAMPLE

The quantitative analysis of the final sample is presented along the variables publication language, publication type, conception, type of school, focus groups, year of publication and diversity dimensions. Here a selection of the variables, which we surveyed, is shown descriptively and described in more detail, both individually and as cross-tables. The total sample consists of n=297 publications. Approximately one third of them are written in English and two thirds in German. Although the majority of the collected literature was found in English databases, a large part of the English literature was excluded in the selection process. This is due to the fact that the articles mostly focused on scientific subjects and did not show any connection to inclusive education. Next time, the search terms should be revised, for example, terms such as diversity should be included in the search terms, i.e., the search strings in the databases should be adapted again for a more international perspective.

If we look at the distribution of empirical and theoretical publications over the years, we can see that the number of titles has risen sharply, especially in the last ten years (Fig. 4). This is probably a consequence of the UN Convention of 2006, which established an agreement for the rights of people with disabilities and other vulnerable groups and which was gradually signed and ratified by the countries (United Nations, 2006). Germany ratified the agreement in 2009, which was a factor that stimulated the discussions on inclusive education. Accordingly, Figure 4 shows that the number of publications has increased along with the theoretical discussions. Particularly in the last five years, the empirical papers have risen sharply and exceeded the number of theoretical papers today. Initially, the difference of titles from 2005-2009 to 2010-2014 ($X^2(1, N = 114) = 5.053, p = .025^*$) is significant. While in the period from 2010 to 2014 n=31 empirical and n=54 theoretical titles were published, in the following period from 2015 to 2019 there have already been n=79 empirical and n=69 theoretical titles. The number of empirical and theoretical publications differ significantly within the periods 2000-2004 ($X^2(1, N = 18) = 5.556, p = .018^*$) and 2010-2014 ($X^2(1, N =

$85) = 6.224, p = .013^{*}$). This is a welcome development, since inclusive science education can be further developed, especially if empirical studies can demonstrate how inclusive practice can be effectively implemented.

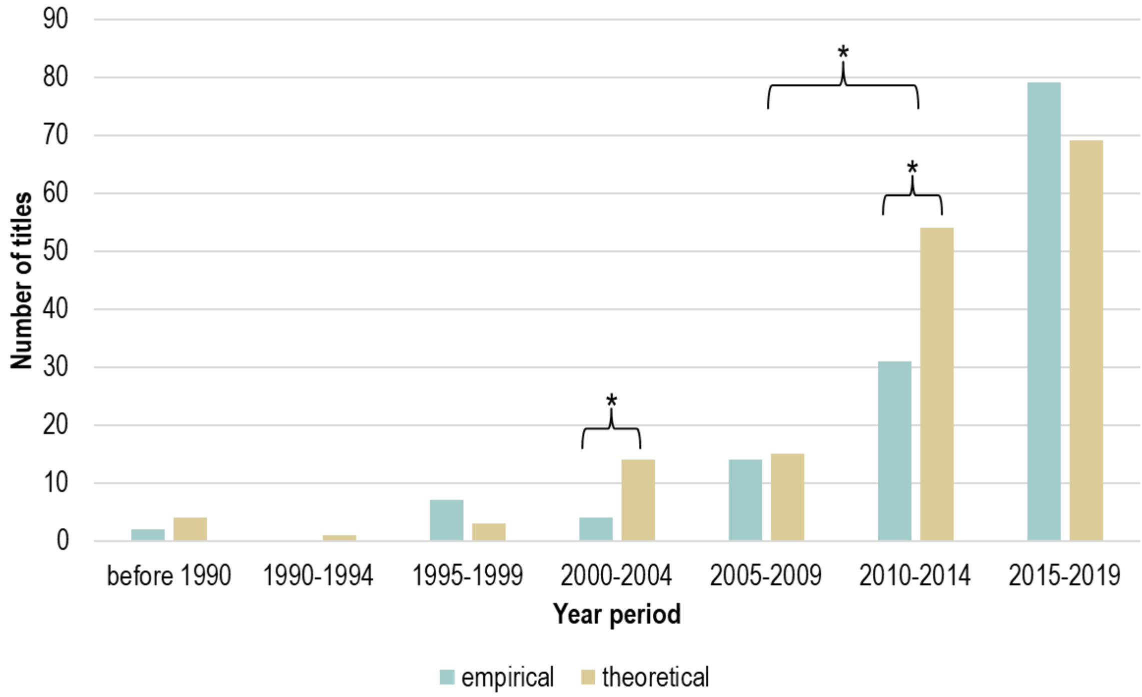

Fig. 4. Empirical and theoretical titles distributed over the years. Significant differences are marked with asterisks (p < .05).

Although the UN Convention refers to people with disabilities, which implies at first sight a narrow understanding of inclusion, the distribution of the concepts of inclusion over the years (Fig. 5) in the dataset shows that not a certain but both the wide and the narrow concept of inclusion are addressed more and more The wide concept of inclusion includes various dimensions of diversity. All titles, which do not focus on a single diversity dimension, but include all individualities of the students, are summarized in this category. Essentially, it means that all students can participate in science lessons. No additional educational needs are labelled in a deficit-oriented manner, but all students with their individual abilities are taken into account. The wide concept of inclusion is contrasted by the additional educational needs concept, which follows a narrow understanding of inclusion. To compare the distributions of both concepts, the graph shows that the number of titles with the wide concept of inclusion increases more steeply than the number of titles with the narrow concept. For comparison, the number of papers which focus on a wide concept of inclusion increase from n=37 (2010-2014) to n=76 (2015-2019) and for additional educational needs as the narrow concept of inclusion from n=34 (2010-2014) to n=57 (2015-2019). Nevertheless, this difference is not significant ($X^2(1, N = 204) = 2.373, p = .123$).

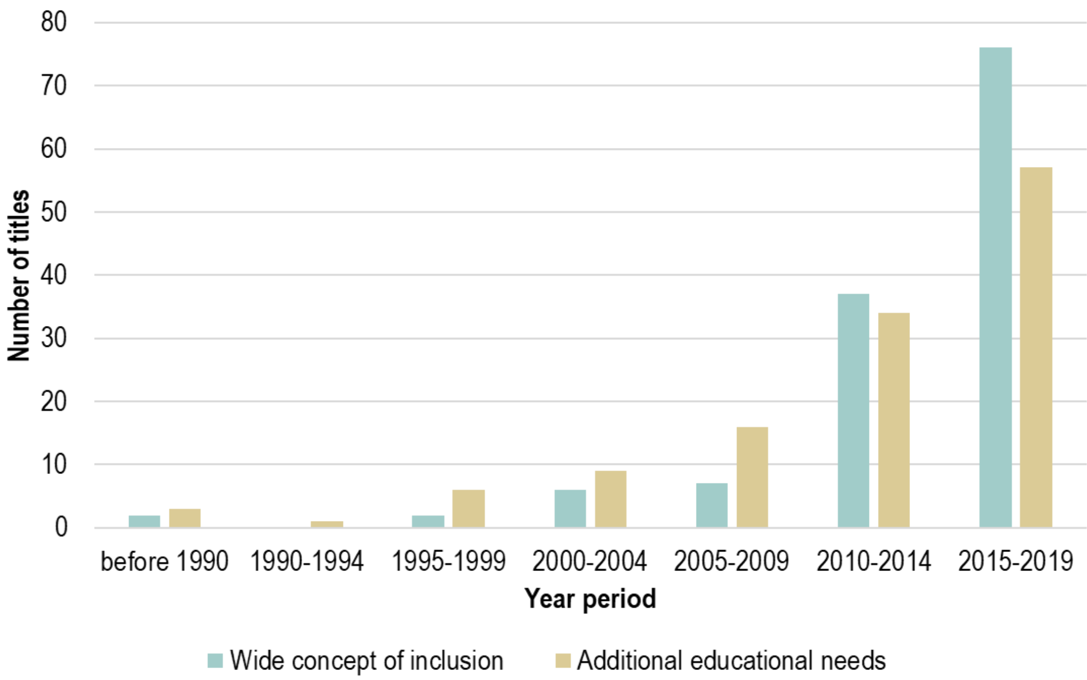

Fig. 5. Distribution of the number of titles referring to a wide concept of inclusion and to additional educational needs

If we look at the distribution of focus groups in the final sample universe (n=297), which illustrates the emphasis in the publications, two groups stand out: the students with 40.7 % and the indefinable group with 39.7 % (Fig. 6). 'Indefinable' includes all titles that have no specific reference to protagonists of the publication. This means that, for example, teaching concepts, educational models or teaching materials are discussed.

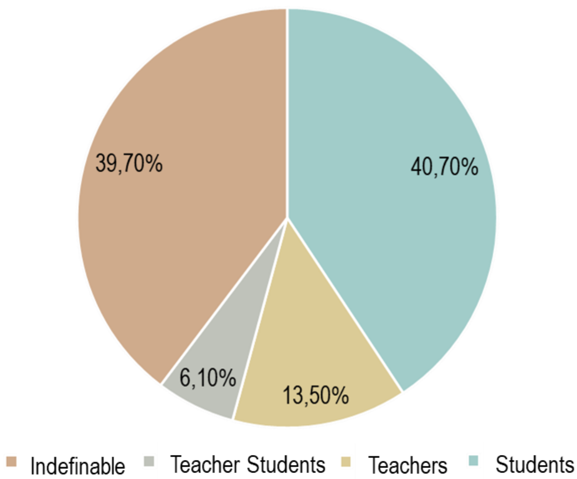

Fig. 6. Distribution of the focus groups in the sample universe

In only 13.5 % of the publications teachers are researched or thematised. It is particularly noticeable that only 6.1 % of the titles focus on student teachers. This shows the strong need for pre- and in-service research. Focusing on student teachers is a prerequisite for advancing and further developing teacher education with regard to inclusive science teaching. It should not be neglected that student teachers are the ones who can take the new findings on inclusive science education into school.

With a deeper look into the two focus groups teachers and student teachers, the distribution of publications over the years shows that individual publications of both focus groups have already appeared between 1995 and 1999 (Tab. 5). Regarding teachers, an increasing trend began between 2005 and 2009. For student teachers, there was one publication between 2010 and 2014, but the actual work in this field has only begun in the current period from 2015 to 2019.

Tab. 5. Development of the focus on teachers and student teachers in publications on inclusive science education

	Before 1990	1990-1994	1995-1999	2000-2004	2005-2009	2010-2014	2015-2019
Teachers	0	0	3	0	8	7	22
Student teachers	0	0	1	0	0	1	16

In summary, especially the last decade (since 2010) a steep increase in publications in the field of inclusive science education is noticeable. A similar development can be observed for empirical publications as well as for publications with a focus on teachers and becoming teachers.

3.2 THE FRAMEWORK FOR INCLUSIVE SCIENCE EDUCATION

The Framework for Inclusive Science Education is presented in full in the annex D (Tab. 8). It consists of 16 main categories[5] as listed in Figure 7. In green are the characteristics of science education marked which run through all levels of the framework with the same name. This means, for example, that under the category inquiry-based learning the term is used for all categories at sub-category level, code level and subcode level. The degree of abstraction level is determined by the red-coloured addition to the characteristic of science education. Here at the level of the main categories terms such as "developing inclusive...", "adapting ... for inclusive education", "teaching ... inclusively", "creating inclusive ..." are used. This highest abstraction level is intended to list the characteristics of science education and connect them to a phrase of inclusion. This very general level of the main categories does not answer in any way how the characteristics of science education can be implemented in inclusive practice.

<table>
<tr><td>Creating inclusive science learning environments</td><td>Adapting security for inclusive education</td><td>(Developing inclusive) diagnostics for scientific specifics</td><td>Teaching scientific concepts inclusively</td><td>Creating inclusive scientific contexts</td></tr>
<tr><td>Teaching the understanding of nature of science inclusively</td><td colspan="3" rowspan="3">THE FRAMEWORK FOR INCLUSIVE SCIENCE EDUCATION</td><td>Enabling the development of scientific terminology inclusively</td></tr>
<tr><td>Creating inclusive data evaluation and result presentation</td><td>Creating inclusive inquiry-based learning</td></tr>
<tr><td>Developing students' science conceptions inclusively</td><td>Teaching scientific phenomena inclusively</td></tr>
<tr><td>Creating inclusive application of scientific research methods</td><td>Creating inclusive scientific documentation</td><td>Developing inclusive scientific information media</td><td>Creating inclusive generation of hypotheses and research questions</td><td>Teaching scientific models inclusively</td></tr>
</table>

Fig. 7. Main categories of the framework

To specify the connection between the characteristics of science education and the inclusive implementation, the further levels of the framework are necessary. At the subcategory level, a recurring pattern is used in the additions to the characteristics of science education regarding the terms of inclusive pedagogy. Essentially, the order of the subcategories is as follows for each main category, whereby it should be noted that the omission always mean a characteristic of science education.

1. Enabling ... materially guided
2. Enabling ... action-oriented
3. Supporting ... linguistically
4. Enabling ... digitally
5. Supporting ... cognitively

[5] The definitions of the main categories are available on the pages 42-47, appendix D.

6. Supporting… communicatively
7. Enabling … through various degrees of openness
8. Creating … on different levels of requirements
9. Creating … on different levels of abstraction
10. Enabling … reflectively
11. Pre-teaching ...
12. Enabling … at certain learning locations
13. Enabling ... in a constructive learning atmosphere

(1.) Materially guided refers to all things that can be perceived by the students with their senses. This includes visualisations, auditory materials, help cards etc. Categories are called for example "Enabling inquiry-based learning materially guided" or "Enabling scientific concepts materially guided". (2.) Activity-based includes actions such as experiments or exploratory learning that the students carry out. (3.) Linguistic support is used, for example, in connection with adaptations in easy language. (4.) Technology-based includes materials and equipment used which are meant to be digital implementations to foster inclusion in science teaching. (5.) Learning strategies are cognitive supports that are given to the students so that they can apply strategies for learning as independently as possible. (6.) Communicative support includes offers that are given orally, for example, by peer-support, by a learning group, and by the teacher as a learning companion. Work in multi-professional teams is also included under communicative support. (7.) Different degrees of openness describes the degree of guidance. Students can have strict guidelines for learning science or be more freely involved with a higher degree of self-activity. The different degrees of openness should not be confused with different levels of requirements. For example, if a new scientific method is introduced, it may be more teacher-led, but still place high demands on the learners' cognitive or practical skills (cf. Abels, 2015). (8.) The different levels of requirements describe how to address different levels of student competencies. (9.) Different levels of abstraction can occur in science teaching. Accordingly, the level of abstraction can be on a concrete phenomenal level or abstract thought processes can be on a molecular imaginary level (cf. Johnstone, 2000). (10.) Reflecting on a specific scientific feature is in some parts of the framework less and in other parts more superficially presented. It means, for example, that characteristics of science education are conveyed in distinction to something else. For example, when models are reflected, they are distinguished from reality or the existence of different models is justified. (11.) Pre-teaching is used when teachers prepare students in school for the actual science lesson. (12.) Different places of learning can be attended in or out of school. These include, for example, school laboratories, school gardens or museums. (13.) Finally, enabling a constructive learning atmosphere means, for example, that the students and their potential, but also any mistakes that occur, are respected and valued.

Only subcategories, which occurred in the literature are included in the framework. Therefore, not all main categories contain all theoretically possible subcategories. Although a comprehensive system of categories has been established based on the literature, there are gaps which we did not want to fill theoretically or arbitrarily, but which will be filled by further research and application. The Framework for Inclusive Science Education is a helpful tool to make these gaps visible.

With regard to the subcategories, it should also be noted that the first main category 'Developing inclusive science learning environments' has a different structure than the other main categories. This results from the fact that a room is adapted here that "behaves" differently from, for example, materials that the students use or actions that the students perform. Overall, on the subcategory level it can be seen in which directions adaptations can go in order to make the characteristics of science education more inclusive. However, the question of what exactly a teacher can do to make science lessons inclusive cannot be answered at this level either.

The next specific level in the framework is the code level. The code level also contains recurring terms. This level already gives concrete instructions on how the characteristics of science education can be implemented in an inclusive way. The instructions for action are given in even more detail at the subcode level. These categories leave no question of implementation unanswered. At this level, the distinctive feature is that the structure of the categories differs from the categories of the other levels of abstraction. While at all other levels a category consists of the characteristic of science education with an inclusive infinitive, at subcode level the category is formed with three dots and the inclusive implementation as a modal adverbial (e.g. "... by short sentences", "... with glossaries"). With the three dots, a link is made to the higher-level code whose wording is specified at this level. This means that the questions "what?", "with what?", "how?" are answered in a reduced form in relation to the higher-level code. The reductions were made in order to provide clarity and to simplify the reading of the framework.

3.3 DESCRIPTIVE PRESENTATION OF THE FRAMEWORK

The framework was analysed in terms of the distribution of the paraphrases and the categories. To repeat the terms, the paraphrases were initially clustered. Each paraphrase originate from a single citation. The clusters of paraphrases were then combined into headings, each of which represents a category. Overall, the categories or the paraphrases enclosed can be located on the four levels of abstraction of the framework.

The framework contains in total of n=935 categories. These are divided into n=421 subcodes, n=368 codes, n=130 subcategories and n=16 main categories (Fig. 8). We recall that categories were marked as empirically tested if they contain at least one paraphrase derived from empirical evidence. The distribution of empirical and theoretical categories shows that at subcode and code level over 40 % of the categories are empirical. At the subcategory level, over 20 % are empirical, while at the main category level almost 70 % are considered as empirical. In summary, the ratio of empirical categories to the total quantity is 0.41, which is lower than the ratio of theoretical categories to the total quantity. The difference between empirical and theoretical categories is significant at the subcategory level $(X^2(1, N = 130) = 15.934, p = .007^*)$ and the main category level $(X^2(1, N = 16) = 5.065, p = .024^*)$. Note that each category considered as empirical can also contain paraphrases originating from theoretical papers.

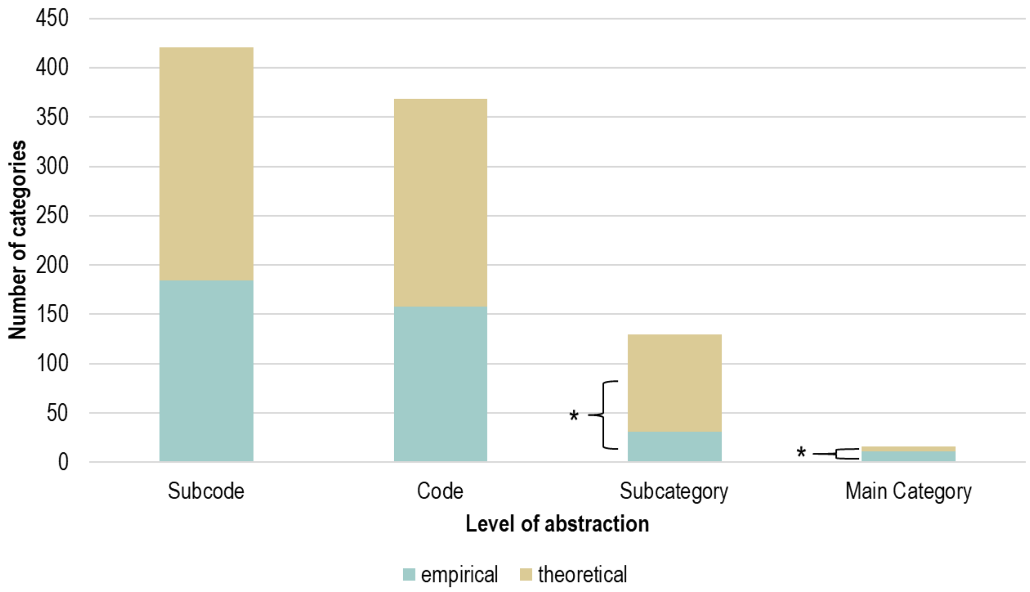

Fig. 8. Number of empirical and theoretical categories on different levels of abstraction. Significant differences are marked with asterisks (p < .05).

The quantitative analysis of the categories also shows which main categories are the largest regarding the number of all underlying categories (subcategories, codes and subcodes). This provides information about which characteristics of science education are described extensively in the literature. The five largest main categories include 'Teaching scientific concepts inclusively', 'Developing inclusive scientific information media', 'Teaching scientific terminology inclusively' and 'Creating inclusive inquiry-based learning' (Fig. 9). Among these main categories, between 89 and 126 categories are listed. The other main categories, which are not listed in Figure 9, have between 20 and 55 categories.

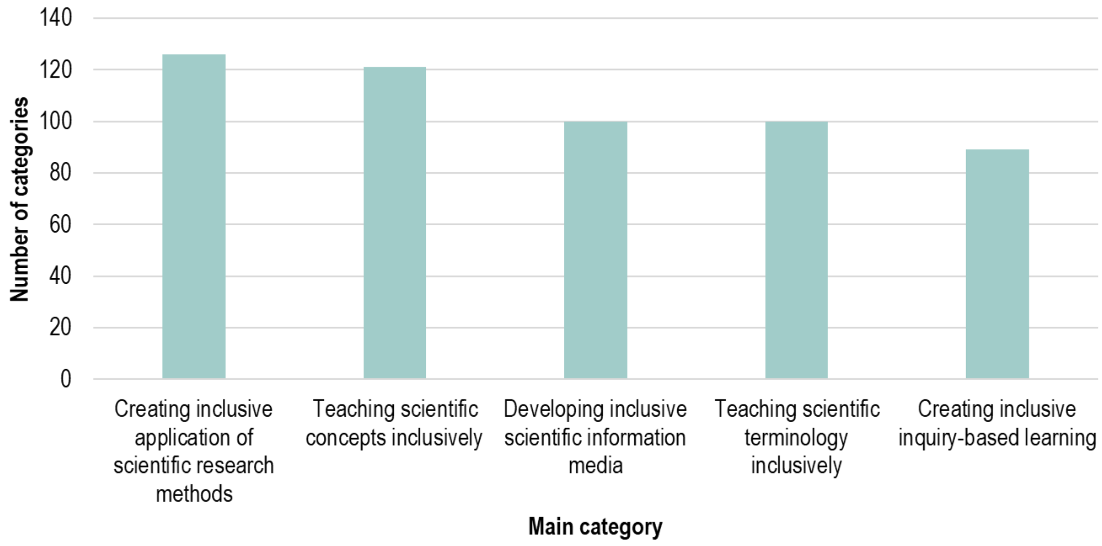

Fig. 9. Main categories with the highest number of categories

The paraphrases were derived directly from the coded text passages. Each of the passages contains a single aspect and was quoted individually via Citavi. In this way, their quantitative analysis can establish a link from the framework to the literature and can provide conclusions about the content structure of the literature in the sample. Altogether n=1627 paraphrases were generated from the literature, which were later combined into categories. Of these, n=1023 were derived from theoretical and n=604 from empirical publications. More precisely, the difference between the theoretical and empirical paraphrases becomes apparent in the distribution across the abstraction levels of the framework (Fig. 10). N=603 paraphrases at subcode level, n=633 paraphrases at code level, n=206 paraphrases at subcategory level and n=185 paraphrases at main category level were derived from the text. The distribution of empirical paraphrases in the total amount of paraphrases per level of abstraction increases from main category level to subcode level. At the main category level, the percentage of empirical paraphrases in the total number of paraphrases per abstraction level is 25 %, at the subcategory level the percentage is 33 %, at the code level the percentage is 39 % and at the code level 40 %. The difference between the empirical and theoretical paraphrases is significant at the main category level ($X^2(1, N = 185) = 10.883, p = .001^*$). As we previously outlined the gaps in inclusive science education (see section 1.), this comparison shows that the parts in the literature that bring together the connection between science education and inclusion are mostly theoretical, but have not yet been empirically tested.

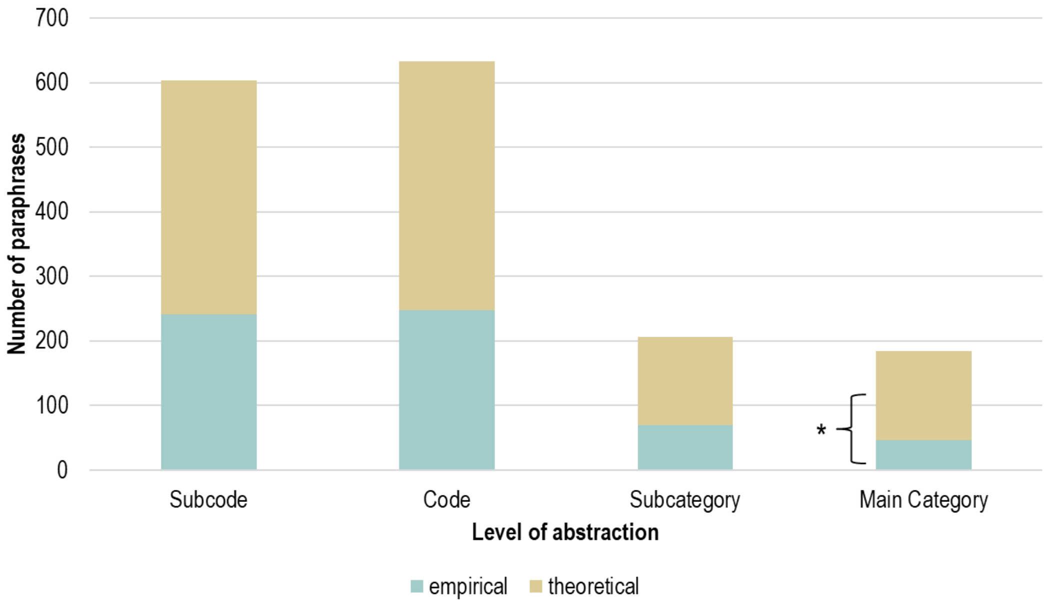

Fig. 10. Distribution of the empirical and theoretical paraphrases on the levels of abstraction of the frame-work. Significant differences are marked with asterisks (p < .05).

In Figure 11, the distribution of paraphrases regarding the main categories is shown. This describes on which level of abstraction inclusive science education is depicted in the literature. It can be seen, that the paraphrases are most frequently found in the main categories 'Teaching scientific concepts inclusively', 'Creating inclusive inquiry-based learning' and 'Creating inclusive application of scientific research methods' with a number of paraphrases from n=238 to n=252. These main categories are closely followed by 'Developing inclusive information media' and 'Teaching scientific terminology inclusively' with a number of paraphrases of n=195 and n=182.

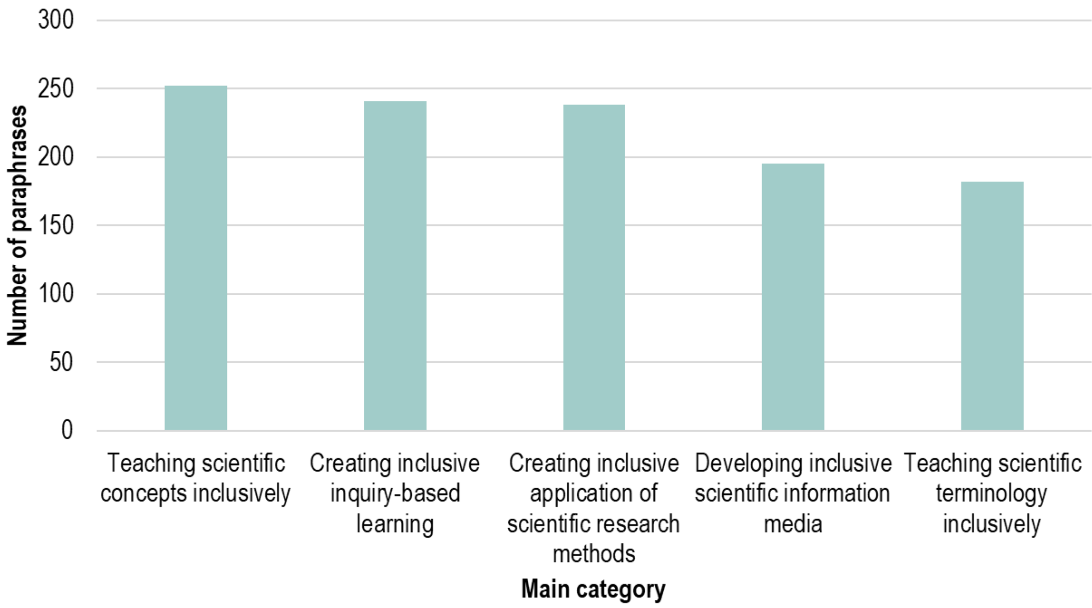

Fig. 11. Main categories with the largest number of paraphrases

If we compare the main categories with the most categories and paraphrases, the same five main categories are listed in each case. This shows that the category deriving adequately summarises the data. There are differences in the number of paraphrases, i.e., how often content is taken from the literature, and the number of categories derived from it. This difference is particularly evident in the main category for inquiry-based learning. Considering the number of paraphrases it is in second place and considering the number of categories in fifth place. Here, more paraphrases were combined in clusters than, for example, in the main category of scientific concepts. The difference between the empirical and theoretical elements is greater for the paraphrases than for the categories. This means that the share of empirical elements is smaller for paraphrases than for categories.

In summary, it can be stated that the framework itself is very large due to the categories, summarised by the methodical derivation of the categories through the paraphrases. In relation to each other, from subcode to subcategory level, on average between 1.4 and 1.7 paraphrases are combined into one category. With this number the framework is rather delicate. We will discuss later why a higher degree of summary is not appropriate for our purposes.

4. DISCUSSION OF AND IMPLICATIONS FOR THE APPLICATION OF THE FRAMEWORK

Overall, the systematic literature review allows understanding the development of the state of research regarding inclusive science education. The quantitative analyses show the extent to which research in this area has developed and provide information on the quantitative structure and extent of the framework. The categories of the framework were generated by the qualitative analyses. These results do not only provide clues for the implementation of inclusive science education, but also indicate which areas or characteristics of science education are (not) addressed in literature up to now.

As the data search forms the basis of a systematic literature review, it has to be reflected upon first. The use of the search terms is decisive for the hits that are later available in the sample. During our data search, it became clear that despite the revision of the search terms, relevant terms such as “diversity” did not occur. Furthermore, we did not consider that in an international context the term “equity” is used for discussions on inclusion rather extensively. This limits the sample and the results generated from it. Another option for revising the data search is to use other than the listed databases. That means that a manual search in Google Scholar or on platforms like ResearchGate can be done as well. In our case it became clear that some titles were not listed during the data analyses, e.g., because the journals were not listed in the databases (e.g., Abels, 2019).

Also, it can be noted that the development of research in inclusive science education is rapidly increasing. Starting around 2010 there is a significant increase in publications. The overall increase is the strongest of all increases in the published literature since 1975 (Fig. 4). This development seems to indicate that the demand for work in inclusive science education is being addressed. A similarly positive development can be observed in relation to the wide concept of inclusion. The publications show a strong increase. Due to the fact that in our sample only a small number of titles refer to specific diversity dimensions such as ethnicity, gender or language, there are two possible reasons for this small number. Either the data search has to be specifically targeted to specific diversity dimensions, or little has been published in these fields of research so far – which is definitely not true for gender. This could mean that publications on certain diversity dimensions do not use keywords related to inclusion.

In the quantitative analysis, we did not present the comparison of publications with the focus on primary and secondary school. Nevertheless, this comparison can be very interesting, since science education differs between the primary and secondary level, for example, especially in teaching and learning the scientific concepts. While the concepts of science in primary school are mainly at the phenomenal level, the level of abstraction (molecular or atomic level) increases with grade. At the phenomenal level, concepts can still be perceived with human senses. At the more abstract molecular level, for example, a certain amount of abstraction ability is required of the students. In terms of teaching scientific concepts, the implementation of inclusion can reach its limits at this point as a difficulty of the access to the abstract concepts may occur (Abels, 2020). Although we have not coded the different subjects (chemistry, biology and physics) in the quantitative analysis yet, it would be interesting to distinguish in which subjects the inclusive implementation of characteristics of science education is primarily discussed.

The quantitative analysis of the focus groups shows the relatively small proportion of literature in the sample that is devoted to student teachers, but also to teachers. However, in order to be able to make science education inclusive, these are the protagonists responsible for the implementation of inclusive practice. For a long time teachers were not prepared for inclusive science education (Abels, 2019; Kahn, Pigman & Ottley, 2017). They were apparently left out in the process of implementing inclusion. This may also be a reason for the low level of research in this area.

All in all, the overview of the quantitative analysis of the framework shows how extensive the framework is. With a total of n=935 categories the framework is rather uncommon in practical use. For this reason, we give information on how the framework can be applied later in this section. Although the framework has a large scope, it does not claim to be complete. If we compare the main categories, which essentially represent the characteristics of science education, with the goals of science education displayed by the OECD (2018), the characteristics of science education derived from the literature reflect the content of scientific literacy. The fact that the framework has gaps is obvious by looking at the subcategory level. A total of twelve different ways of implementing inclusion in the characteristics of science education were found at this level of abstraction (Section 3.2). However, not every main category is filled with all twelve different subcategories. This results from the fact that categories were only formed if they could be derived from literature. The structure of the subcategories essentially reflects the areas of the inclusive implementation of scientific characteristics. The extent to which these would need to be differentiated in order to emphasize important aspects of inclusive science education needs to be discussed. An example of this are the subcategories for science concepts and students' scientific conceptions, whether the enabling of different levels of abstraction should be given its own subcategory instead of falling under the subcategory of different levels of requirements. Nevertheless, for the consistent structure of the framework we have deductively modified the subcategories. Further publications are needed to fill these gaps. In process of developing the framework, it has to be taken into account that the categories or the framework are the result of an interaction of the English and German language. While the paraphrases were still derived in the original language of literature, the paraphrases were combined into categories in German. For this publication the German categories were then translated into English. By matching the English paraphrases, an attempt was made to stick to the original wording whenever possible. However, due to the summarization and translation processes, deviations may have occurred. It should be noted that this framework is to be understood as dynamic. This means that it will change in the future and will be expanded or reduced in some places. In order to represent the first stage of the framework and to be able to illustrate explicitly that the results from the literature up to 2019 are presented here, we decided not to fill the subcategories abductively at this stage. On the other hand, we have added a category "other" on the subcategory level, which is meant to emphasize that the framework is extensible. By using the framework in our Nawi-In project, we will provide more updated versions of the framework in the future. Furthermore, further publications on the use of the framework are expected. This working paper has been prepared to ensure that sufficient consideration is given to the complex methodological approach of the systematic literature and data analysis. We will refer to this working paper in further publications on the contents of the framework and in which the main categories and characteristics of science education will be explained in detail.

The framework illustrates which areas of inclusive science education have been addressed in a more differentiated manner in theory and research and where a need for further theoretical discussions and empirical studies is. The OECD (2019) shows that the specificity of science education is already very clear, nevertheless the inclusive implementation of teaching and especially of subject teaching is not always distinct. The challenge lies in the fact that an inclusive implementation cannot only be presented theoretically, but that it must be empirically tested whether the implications of the framework actually lead to the implementation of inclusion in science education. The question here is which of the categories constitute truly inclusive science education and which categories just represent “good” science education. It is not always possible to determine from a single category whether this category is in fact inclusive. One example is to teach scientific concepts through technical language. Whether this implementation is just part of an ordinary science class or whether this example leads to an inclusive implementation is not clear. The reason for this is that the contents of the categories for this status were adopted from the literature without being discussed and selected theoretically with regard to the reference to inclusive implementation, and without being applied in practice. It is questionable in which form our project will be able to clarify the question of the actual inclusive implementation, if for its verification not only the teachers’ perspective is to be considered, but also the students’ and their well-being in class (Brauns, 2020). We can validate the extent to which the framework can be applied to practice using video analysis. It can be assumed that the implementation of inclusion in science education using the framework depends on the individuality of the learning group. This means that not all categories cannot and do not necessarily have to be applied in inclusive science teaching. The number of codes and sub-codes that have to be applied to a learning group in order to make science education inclusive depends on the learning group itself, the teaching objectives and the resources. We will empirically investigate which categories and how many categories lead to the implementation of inclusive science education in the future. Nevertheless, the quantitative presentation of the framework shows that some of the categories have so far only been formulated theoretically in the literature. In order to ensure practical efficacy of the framework in teaching further use is required. For this reason, the framework is to be applied and validated in practice in a next step of the research project (Section 5.). By examining the teachers’ perspective of inclusive science education in our project, we will not be able to determine with the framework whether the teaching is in fact inclusive. For this, we would have to take the students’ side into account. With the framework, we can interpret whether participation is facilitated in science lessons. Further research would have to clarify whether students feel recognized and accepted and actually develop scientific skills.

The framework created by the systematic literature review reflects the current theory and research on inclusive science education and can be used both as (1) an analysis grid for researching (student) teachers and as (2) a handbook for teaching. (1) For example, if the focus in a research project is set on the inclusive design of experiments, the main category 'Creating inclusive applications of scientific research methods' can be used with its subordinate categories as a stand-alone analysis grid to analyse the data of a project. The framework can also be used in research on inclusive science education, for example, with the main and/or sub-categories. In this case, the more abstract categories can be used to analyse the data material first. One way is to stay on this level of abstraction and summarising the results on a more general level. The challenge here, however, is that the more abstract the level is, the more difficult it is to understand the concrete implementation and actual description of the relationship between science education and inclusion. Therefore, after coding with the main or subcategories, it is recommendable to insert the results of the analysis at the code and subcode level. During this process, comparisons can be made with the existing codes and subcodes of the framework. In this step, either the results of the analysis are inserted into already existing codes and subcodes of the framework or they supplement the framework with new codes and subcodes that do not yet exist. Overall, the adaptation of the use of the framework must always be adapted to the research question and the research objectives.

(2) These approaches are also possible in teacher education. Becoming teachers can, for example, be given the main category 'Creating inclusive inquiry-based learning' to plan, design and implement inquiry-based science education during a school internship. Such a category can also serve as a stand-alone grid for student teachers to reflect on their own teaching or on teaching of experienced teachers. With the grid, becoming teachers can systematically be introduced to the idea of inclusive science teaching in instructional videos (Brauns et al., 2020). The other way to apply the framework is to first blend out the lower levels of abstraction. This means that the framework is considered as a whole, but the details are omitted. In teacher education, the advantage is that the student teachers are made aware of the specifics of science teaching. The challenge of being able to describe or analyse inclusive science education often lies in the fact that the inclusive implementation of teaching is not thought of in a subject-specific way (Egger, Brauns & Abels, in prep.). Egger et al. (in prep.) show that thinking science teaching and inclusion together requires high analytical skills, which is why novices need to be fostered specifically in this area. Only when the specifics of the natural sciences can be identified, it is possible, starting from this, to address inclusive and science specific implementation.

To make the framework practical for further application and validation, there are in summary two main ways to make the framework practical. On the one hand, fragments of the framework can be put into focus. This means that individual main categories can be considered separately from the framework. On the other hand, the framework can be applied to data on a more abstract level (e.g. only using main or subcategories).

5. VALIDATION OF THE FRAMEWORK AND OUTLOOK

The framework is used in our Nawi-In project for research of student teachers. For the validation of the framework we use the data of our project. The framework is applied to four different types of data: (1) student teachers' videotaped action in science classes, (2) student teachers' video-based and audiotaped self-reflection, (3) student teachers' audiotaped reflection on teaching videos of experienced teachers and (4) teaching videos of experienced teachers.

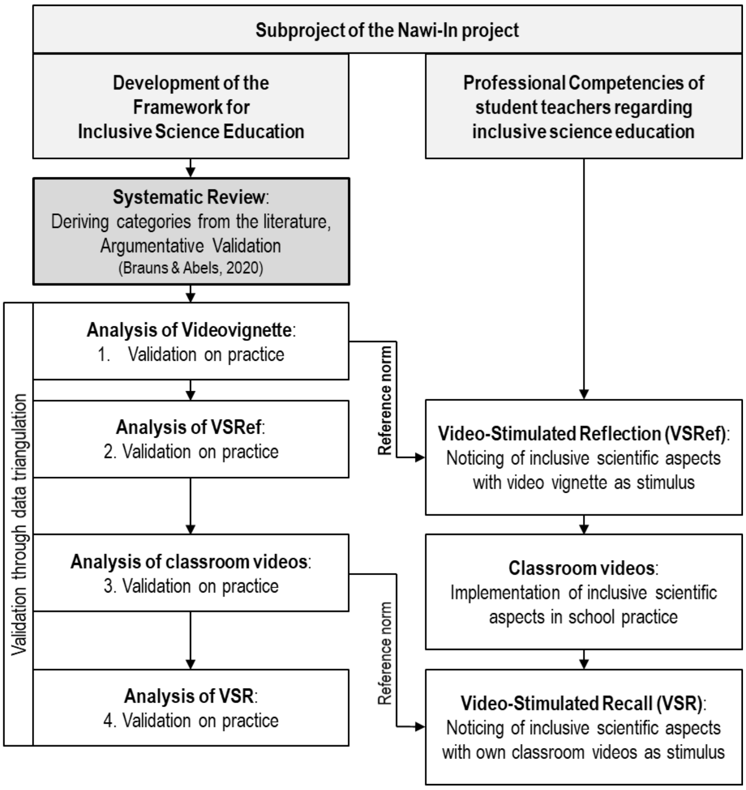

Fig. 12. Validation process of the framework and project procedure within Nawi-In (adapted after Brauns & Abels, 2021, p. 72)

In qualitative research, validation requires development from within the project logic. Exactly from this logic, the validation process of the Framework for Inclusive Science Education is designed (Fig. 12). Therefore, the framework is validated several times on practice by analysing different data from the Nawi-In project (Brauns et al., 2020). In the end, the data will be triangulated and in this way a final validation of the framework will be conducted.

In order to collect the data for the validation of the framework, we accompany the student teachers for two semesters during a project seminar at master level offered by the second author (Brauns et al., 2020). In this seminar, the student teachers are theoretically prepared for the inclusive science lessons and practice their noticing with teaching videos in the first semester (Egger et al., 2019). In the second semester, the student teachers complete an internship in school. While they are at school, they plan and teach their own science lessons. They also conduct their own minor research project in which they use their videos to analyse their lessons in terms of inclusive science education (Brauns et al., 2020). These videos serve us as database. On the one hand, we can analyse which characteristics of inclusive science education the student teachers show in their own lessons. On the other hand, we can validate the framework itself as an analysis tool by deductively coding the teaching videos with the framework using qualitative content analysis (Kuckartz, 2016).

Moreover, these teaching videos are reflected by the student teachers. These self-reflections are recorded on audio and are analysed with the framework. Here, we consider the question of which characteristics of inclusive science education are noticed by the student teachers in their own videos. Furthermore, before and after the first seminar and after the internship, we record external reflections of the student teachers on audio at three times of data collection. They reflect other teachers' implementation of inclusive science education. These external reflections are also evaluated under the same question as the teacher students' self-reflections. All reflections get transcribed and the transcripts analysed with qualitative content analysis using the framework.

Despite to all data that we collect and evaluate from student teachers through our Nawi-In project we have access to videos of lessons by experienced teachers. These teachers teach inquiry-based learning in inclusive science lessons from primary and secondary schools. To these videos we also apply the framework.

In summary, we have a large amount of data with teaching videos of student teachers and experienced teachers as well as self- and external reflections of the student teachers as audio recordings available to validate the framework. With these data, a connection between the theoretical development of the framework and the practical application in the school context will be established. Finally, we will also address the question of which categories are indeed inclusive and which ones represent only "good" teaching. In this context, discussions with experts, for example from the NinU network, will be useful. To what extent this question can be answered will be a challenge.

ACKNOWLEDGEMENTS

This work was sponsored by the Federal Ministry of Education and Research (no. 01NV1731). We are thankful to Samantha Wöbcke, Lisa Stinken-Rösner, Daniela Egger and Lea Heyden for supporting the work by reviewing and methodical support.

REFERENCES

Abdella, B. R. J., Walczak, M. M., Kandl, K. A., & Schwinefus, J. J. (2011). Integrated Chemistry and Biology for First-Year College Students. *Journal of Chemical Education, 88*(9), 1257-1263.

Abels, S. (2015). Scaffolding inquiry-based science and chemistry education in inclusive classrooms. In N. L. Yates (Ed.), *New developments in science education research* (pp. 77–96). New York City: Nova.

Abels, S. (2019). Science Teacher Professional Development for Inclusive Practice. *International Journal of Physics and Chemistry Education, 11*(1), 19–29.

Abels, S. (2020). Naturwissenschaftliche Kompetenzen und Inklusion – Inklusion durch Kompetenzorientierung? In S. Habig (Hg.), *Naturwissenschaftliche Kompetenzen in der Gesellschaft von morgen*. Gesellschaft für Didaktik der Chemie und Physik. 40. Jahrestagung in Wien 2019 (pp. 20-30). Essen: Universität Duisburg-Essen.

Abels, S., & Koliander, B. (2014). „Forschendes Lernen in der Schule" – ein hochschuldidaktisches Konzept. In S. Bernholt (Hg.), *Naturwissenschaftliche Bildung zwischen Science- und Fachunterricht.* Gesellschaft für Didaktik der Chemie und Physik. 34. Jahrestagung 2013 (pp. 79–81). Kiel: IPN.

Abramova, I. V., Shilova, Z. V., Varankina, V. I., & Rubanova, I. V. (2019). Pedagogical Model of Integrative-Modular Training in Professional Preparation of Students. *European Journal of Contemporary Education, 8*(1), 187–200.

Alheit, P. (2009). The symbolic power of knowledge: exclusion mechanisms of the 'University Habitus' in the German HE System. In B. Merrill (Hd.*), Learning to change? The role of identity and learning careers in adult education* (pp. 161-171). Frankfurt, Main: European studies in lifelong learning and adult learning research 5.

American Chemical Society (2016). *Guidlines for Chemical Laboratory Safety in Secondary Schools*. Washington, DC: ACS.

Bardeche, G., Belargent, C., Dardelin , M., Garbagnati, G., Guichard, M., Guichon, D., Hardin, J., Jouffroy, G., Risset, C. A., Rueff, P., & Valentin, J. P. (1980). Integrated Physics at University Level. *European Journal of Science Education, 2*(2), 139–144.

Barke, H.-D., Harsch, G., Kröger, S., & Marohn, A. (2018). *Chemiedidaktik kompakt*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Baumert, J., & Kunter, M. (2006). Stichwort: Professionelle Kompetenz von Lehrkräften. *Zeitschrift für Erziehungswissenschaft, 9*(4), 469–520.

Baumert, J., & Kunter, M. (2011). Das Kompetenzmodell von COACTIV. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Hg.), *Professionelle Kompetenz von Lehrkräften. Ergebnisse des Forschungsprogramms COACTIV* (pp. 29–54). Münster: Waxmann.

Benny, N., & Blonder, R. (2018): Interactions of Chemistry Teachers with Gifted Students in a Regular High-School Chemistry Classroom. *Chemistry Education Research and Practice, 19*(1), 122-134.

Blanchard, M. R., Southerland, S. A., Osborne, J. W., Sampson, V. D., Annetta, L. A., & Granger, E. M. (2010). Is Inquiry Possible in Light of Accountability?: A Quantitative Comparison of the Relative Effectiveness of Guided Inquiry and Verification Laboratory Instruction. *Science Education, 94*(4), 577–616.

Booth, T., & Ainscow, M. (2016). *The index for inclusion: A guide to school development led by inclusive values* (Fourth edition). Cambridge: Index for Inclusion Network (IfIN).

Brauns, S (2020). The Framework for inclusive science education. In ESERA (Ed.), *Book of Synopses. European Science Education Research Association. Virtual Doctoral Network 2020* (pp. 333-341). Oxford: Oxford University.

Brauns, S., & Abels, S. (2021). Videoanalyse mit dem Kategoriensystem inklusiver naturwissenschaftlicher Unterricht (KinU), *4*(2), 71–84. https://doi.org/10.25321/prise.2021.1146

Brauns, S., Egger, D. & Abels, S. (2020). Forschendes Lernen auf Hochschul- und Unterrichtsebene beforschen. In K. Mayr-Keiler & I. Pichler (Hrsg.), *Forschendes Lernen, Transfer Forschung – Schule, Heft 6* (S. 201-211). Bad Heilbrunn: Klinkhardt.

Brauns, S., Egger, D., Abels, S., & Barth, M. (2019). Naturwissenschaftlichen Unterricht inklusiv gestalten (Nawi-In) – Ein Vergleich der Primar- und Sekundarstufe I. In C. Maurer (Hg.), *Naturwissenschaftliche Bildung als Grundlage für berufliche und gesellschaftliche Teilhabe.* Gesellschaft für Didaktik der Chemie und Physik. 39. Jahrestagung 2018 (pp. 675-678). Regensburg: Universität Regensburg.

Bybee, R. W. (1997). Toward an understanding of scientific literacy. In W. Gräber, & C. Bolte (Ed.), *Scientific literacy: An international symposium* (pp. 37–69). Kiel: IPN-Leibniz Institute for Science and Mathematics Education.

Bybee, R. W., Taylor, J. A., Gardner, A., van Scotter, P., Carlson Powell, J., Westbrook, A., & Landes, N. (2006). *The BSCS 5E Instructional Model: Origins, Effectiveness, and Applications.* Colorado Springs: BSCS Science Learning.

Döring, N., & Bortz, J. (2016). *Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Egger, D., Brauns, S., & Abels, S. (in prep.). Competency Development of Pre-service Teachers in Inclusive Science Education.

Egger, D., Brauns, S., Sellin, K., Barth, M., & Abels, S. (2019). Professionalisierung von Lehramtsstudierenden für inklusiven naturwissenschaftlichen Unterricht. *Journal Für Psychologie*, *27*(2), 50–70.

Feak, C. B., & Swales, J. M. (2011). *Telling a research story. Writing a literature review. Rev. and expanded ed.,* [Repr.]. Ann Arbor: Univ. of Michigan Press.

Fink, A. (2009). *Conducting Research Literature Reviews: From the Internet to Paper*. Los Angeles: CA: Sage.

Florian, L., & Black-Hawkins, K. (2011). Exploring inclusive pedagogy. *British Educational Research Journal, 37*(5), 813–828.

Fraser, B. J., Giddings, G. J., & McRobbie, C. J. (1992). Assessment of the Psychosocial Environment of University Science Laboratory Classrooms: A Cross-National Study. *Higher education, 24*(4), 431–451.

Gebhard, U., Höttecke, D., & Rehm, M. (2017). *Pädagogik der Naturwissenschaften*. Wiesbaden: Springer Fachmedien Wiesbaden.

Ghanbari, S. (2015). Learning across Disciplines: A Collective Case Study of Two University Programs That Integrate the Arts with STEM. *International Journal of Education & the Arts, 16*(7), 1-21.

Godovnikova, L. V., Gerasimova, A. S., Galchun, Y. V., & Shitikova, E. V. (2019). The Competency Levels of Disabled Students Who Study in University. *Cypriot Journal of Educational Sciences, 14*(1), 99–110.

Heran-Dörr, E. (2010). Entscheidungsrelevante Strukturelemente der Planung von Unterrichtseinheiten. In S. Tänzer (Ed.), *Sachunterricht begründet planen: Bedingungen, Entscheidungen, Modelle* (pp. 84–99). Bad Heilbrunn: Klinkhardt.

Hermanns, J., Krabbe, C., Hornung, G., Krüpper, A., & Pusch, A. (2019). Experimentieren im inklusiven naturwissenschaftlichen Unterricht. *Inklusive Lehr-Lernprozesse gestalten*, 77–93.

Hodson, D. (2014). Learning Science, Learning about Science, Doing Science: Different goals demand different learning methods. *International Journal of Science Education, 36*(15), 2534–2553.

Johnstone, A. H. (2000). Teaching of chemistry – logical or psychological? *Chemistry Education: Research and Practice in Europe, 1*(1), 9–15.

Kahlert, J. (2016). Inklusion im Sachunterricht. Eine Fachdidaktik auf dem Weg in die Individualisierungsfalle. *Sonderpädagogische Förderung heute*, (3), 234–243.

Kahn, S., Pigman, R., & Ottley, J. (2017). A Tale of Two Courses: Exploring Teacher Candidates' Translation of Science and Special Education Methods Instruction into Inclusive Science Practices. *Journal of Science Education for Students with Disabilities, 20*(1), 50–68.

Koehler, K. E., & Wild, T. A. (2019). Students with Visual Impairments Access and Participation in the Science Curriculum: Views of Teachers of Students with Visual Impairments. *Journal of Science Education for Students with Disabilities,* 22(1), 1-17.

Koomen, M. H. (2016). Inclusive Science Education: Learning from Wizard. In: *Cultural Studies of Science Education, 11*(2), 293-325 (33 Seiten).

Koska, J. & Krüger, D. (2012). Nature of Science-Perspektiven von Studierenden – Schritte zur Entwicklung eines Testinstruments. In D. Krüger, A. Upmeier zu Belzen, P. Schmiemann, A. Möller, & D. Elster (Hg.), *Erkenntnisweg Biologiedidaktik 11* (pp. 115-127), 14. Frühjahrsschule der Fachsektion Didaktik der Biologie im Verband Biologie.Biowissenschaften und Biomedizin in Deutschland.

Krell, G., Riedmüller, B., Sieben, B.,; & Vinz, D. (2007). Einleitung – Diversity Studies als integrierende Forschungsrichtung. In G. Krell, B. Riedmüller, B. Sieben, & D. Vinz (Eds.), *Diversity Studies. Grundlagen und disziplinäre Ansätze* (pp. 7–16). Frankfurt a.M., New York: Campus.

Kuckartz, U. (2018). *Qualitative Inhaltsanalyse. Methoden, Praxis, Computerunterstützung* (Grundlagentexte Methoden, 4., überarbeitete Auflage). Weinheim: Beltz Juventa.

Kunter, M., Baumert, J., Blum, W., Klusmann, U., Krauss, S., & Neubrand, M. (2013). *Cognitive Activation in the Mathematics Classroom and Professional Competence of Teachers*. Boston, MA: Springer US.

Lindmeier, C., & Lütje-Klose, B. (2015). Inklusion als Querschnittsaufgabe in der Erziehungswissenschaft. *Erziehungswissenschaft - Mitteilungen Der Deutschen Gesellschaft Für Erziehungswissenschaft, 26*(51), 7–16.

Markic, S., & Abels, S. (2014). Heterogeneity and Diversity: A Growing Challenge or Enrichment for Science Education in German Schools? *EURASIA Journal of Mathematics, Science & Technology Education, 10*(4), 271-283.

Mayring, P. (2000). Qualitative Content Analysis. Forum: *Qualitative Social Research,* 1(2), n.p.

McComas, W. F. and Olson, J. (1998). The nature of science in international science education standards documents. In W. F. McComas (Ed.) *Nature of science in science education: rationales and strategies* (pp. 41-52). Kluwer (Springer) Academic Publishers.

Nawarathne, I. N. (2019). Introducing Diversity through an Organic Approach. *Journal of Chemical Education, 96*(9), 2042-2049.

OECD (2019). *PISA 2018 assessment and analytical framework*. Paris: OECD Publishing (PISA).

Prengel, A. (2016). Didaktische Diagnostik als Element alltäglicher Lehrarbeit – „Formatives Assessment“ im inklusiven Unterricht. In B. Amrhein (Hg.), *Diagnostik im Kontext inklusiver Bildung, Theorien, Ambivalenzen, Akteure, Konzepte* (pp. 39-48). Bad Heilbrunn: Verlag Lulius Klinkhardt.

Reiners, C. S. (2017). *Chemie vermitteln*. Berlin, Heidelberg: Springer Berlin Heidelberg.

Rott, L. (2018). *Vorstellungsentwicklungen und gemeinsames Lernen im inklusiven Sachunterricht initiieren: Die Unterrichtskonzeption "choice2explore"*. Berlin: Logos.

Rott, L., & Marohn, A. (2018). choice2explore: gemeinsam lernen im inklusiven Sachunterricht. In U. Franz, H. Giest, A. Hartinger, A. Heinrich-Dönge, & B. Reinhoffer (Eds.), *Probleme und Perspektiven des Sachunterrichts. Handeln im Sachunterricht* (pp. 223–230).

Rott, L., Nowosadek, B., & Marohn, A. (2017). Warum kann man Salz im Wasser nicht sehen? Teilchenmodelle im inklusiven Unterricht. *Naturwissenschaften im Unterricht. Chemie*, *28*(162), 16–21.

Scruggs, T. E., Mastropieri, M. A., & Okolo, C. M. (2008). Science and Social Studies for Students with Disabilities. *Focus on Exceptional Children, 41*(2), 1–24.

Seidel, T., Stürmer, K., Blomberg, G., Kobarg, M., & Schwindt, K. (2011). Teacher learning from analysis of videotaped classroom situations: Does it make a difference whether teachers observe their own teaching or that of others? *Teaching and Teacher Education, 27*(2), 259–267.

Sherin, M. G. (2007). The development of teachers' professional vision in video clubs. In R. Goldman, R. Pea, B. Barron und S. J. Derry (Ed.), *Video research in the learning sciences.* Hillsdale: Erlbaum, 383–395.

Stinken-Rösner, L., Rott, L., Hundertmark, S., Baumann, Th., Menthe, J., Hoffmann, Th., Nehring, A., & Abels, S. (2020). Thinking Inclusive Science Education from two Perspectives: inclusive Pedagogy and Science Education. *RISTAL* (3), 30–45.

Stürmer, K., Seidel, T., & Schäfer, S. (2013). Changes in professional vision in the context of practice. *Gruppendyn Organisationsberat, 44*(3), 339–355.

Teke, D., & Sozbilir, M. (2019). Teaching Energy in Living Systems to a Blind Student in an Inclusive Classroom Environment. *Chemistry Education Research and Practice, 20*(4), 890-901.

Therrien, W. J., Taylor, J. C., Watt, S., & Kaldenberg, E. R. (2014). Science Instruction for Students with Emotional and Behavioral Disorders. *Remedial and Special Education, 35*(1), 15–27.

UNESCO (2005). *Guidelines for Inclusion: Ensuring Access to Education for All*. Paris: UNESCO.

United Nations (2006). Convention on the Rights of Persons with Disabilities. http://www.un.org/disabilities/documents/convention/convoptprot-e.pdf [23.07.2020]

Veber, M., & Fischer, C. (2016). Individuelle Förderung in Inklusiver Bildung – eine potenzialorientierte Verortung. In B. Amrhein (Hg.), *Diagnostik im Kontext inklusiver Bildung, Theorien, Ambivalenzen, Akteure, Konzepte* (pp. 39-48). Bad Heilbrunn: Verlag Lulius Klinkhardt.

Vitoriano, F. A., Teles, V. L. G., Rizzatti, I. M., & Pesssoa de Lima, Régia C. (2016). Promoting Inclusive Chemistry Teaching by Developing an Accessible Thermometer for Students with Visual Disabilities. *Journal of Chemical Education, 93*(12), 2046–2051.

Walkowiak, M., Rott, L., Abels, S., & Nehring, A. (2018). Network and work for inclusive science education. In I. Eilks, S. Markic and B. Ralle (Eds.): *Building bridges across disciplines* (pp. 269–274). Aachen: Shaker.

Watson, S., & Johnston, L. (2007). Assistive Technology in the Inclusive Science Classroom. *Science Teacher, 74*(3), 34–38.

Watt, S. J., Therrien, W. J., & Kaldenberg, E.R. (2014). Meeting the Diverse Needs of Students with EBD in Inclusive Science Classrooms. *Beyond Behavior, 23*(2), 14–19.

Weinert, F. E. (2000). *Lehren und Lernen für die Zukunft - Ansprüche an das Lernen in der Schule*, Pädagogisches Zentrum in Bad Kreuznach.

Werning, R. (2014). Stichwort: Schulische Inklusion. *Zeitschrift für Erziehungswissenschaft, 14,* 601–623.

Wuttke E., Seifried J. (2013) Diagnostic Competence of (Prospective) Teachers in Vocational Education. *From Diagnostics to Learning Success*. Professional and VET learning, vol 1. SensePublishers, Rotterdam.

Ziemen, K. (2016). Inklusion und diagnostisches Handeln. In B. Amrhein (Hg.), *Diagnostik im Kontext inklusiver Bildung, Theorien, Ambivalenzen, Akteure, Konzepte* (pp. 39-48). Bad Heilbrunn: Verlag Lulius Klinkhardt.