



## Assessing students' epistemic beliefs related to professional and school science

Frauke Voitle, Benedikt Heuckmann, Nele Kampa & Kerstin Kremer

**To cite this article:** Frauke Voitle, Benedikt Heuckmann, Nele Kampa & Kerstin Kremer (2022) Assessing students' epistemic beliefs related to professional and school science, *International Journal of Science Education*, 44:6, 1000-1020, DOI: [10.1080/09500693.2022.2059821](https://doi.org/10.1080/09500693.2022.2059821)

**To link to this article:** <https://doi.org/10.1080/09500693.2022.2059821>



© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



[View supplementary material](#)



Published online: 12 Apr 2022.



[Submit your article to this journal](#)



Article views: 1879



[View related articles](#)



[View Crossmark data](#)

# Assessing students' epistemic beliefs related to professional and school science

Frauke Voitle <sup>a,b</sup>, Benedikt Heuckmann <sup>a,c</sup>, Nele Kampa <sup>d</sup> and Kerstin Kremer <sup>a,b,e</sup>

<sup>a</sup>Institute for Science Education, Leibniz University Hannover, Hannover, Germany; <sup>b</sup>Department of Biology Education, Leibniz Institute for Science and Mathematics Education (IPN) at the University of Kiel, Kiel, Germany; <sup>c</sup>Centre for Biology Education, University of Münster, Münster, Germany; <sup>d</sup>Department of Research and Development in Teacher Education, University College of Teacher Education Tyrol, Innsbruck, Austria; <sup>e</sup>Institute for Biology Education, Justus Liebig University, Giessen, Germany

## ABSTRACT

The school context, like other influences, affects the development of students' epistemic beliefs (EBs), with theoretical conceptualizations suggesting a distinction between EBs related to professional and school science. However, empirical studies rarely address this distinction, and instruments separately assessing these EBs are not available. We address this research gap by providing a new questionnaire that is based on the established instrument by Conley et al. (2004) and which uses pairs of analogous items referring to the contexts of professional and school science. We succeeded in simultaneously assessing students' EBs with reliable scales that allow for direct comparisons. Cross-sectional data is reported from a cohort study ( $N = 405$ ) in grades 6 ( $n = 132$ ), 8 ( $n = 167$ ) and 10 ( $n = 106$ ) that showed higher agreement with the scales related to professional science as well as higher agreement in higher grade levels. We further discuss methodological aspects of scale-formation, and highlight promising implementations of our questionnaire that future research could benefit from.

## ARTICLE HISTORY


Received 30 January 2021  
Accepted 27 March 2022

## KEYWORDS

Assessment; epistemology;  
secondary/high school

## Introduction

Epistemic beliefs (EBs) have repeatedly been linked to students' learning and academic achievement (Hofer, 2016; Lee et al., 2016; Lin & Tsai, 2017) and continue to be a meaningful research interest in international science education (Lee et al. 2021). The inclusion of EBs in large-scale assessments, such as the Programme for International Student Assessment (PISA) (OECD, 2019) has helped to identify EBs as an important predictor for science learning (Karakolidis et al., 2019; She et al., 2019). Like views about the nature

**CONTACT** Frauke Voitle  [voitle@idn.uni-hannover.de](mailto:voitle@idn.uni-hannover.de)  Institute for Science Education, Leibniz University Hannover, Hannover, Germany; Department of Biology Education, Leibniz Institute for Science and Mathematics Education (IPN) at the University of Kiel, Kiel, Germany

 Supplemental data for this article can be accessed at <https://doi.org/10.1080/09500693.2022.2059821>

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

of science (Khishfe & Lederman, 2007), EBs have proven to be sensitive to science instruction (Bernholt et al., 2021; Conley et al., 2004; Schiefer et al., 2020) and differ between grade levels, with older students holding beliefs that are less absolute (Kremer & Mayer, 2013). Furthermore, in addition to their discipline- and domain-specific nature (Hofer, 2000; Muis et al., 2006), EBs are also situated and contextualised (Mason, 2016).

An important but often underrated aspect of EB research is concerned with dependencies between the specific context of school science and students' EBs about science. Researchers have argued that students may develop 'parallel ways of knowing' (Edmondson & Novak, 1993, p. 550), with students' views about professional and school science mostly remaining separate from each other. However, while students' perceptions of professional and school science are likely to differ, this explicit distinction remains commonly unacknowledged in the analyses of empirical studies (Conley et al., 2004; Solomon et al., 1996; Villanueva et al., 2019) and questionnaires separately assessing the two contexts are not available.

In our paper, we aim to (a) develop an instrument that allows for simultaneous assessment of students' EBs related to professional and school science and (b) investigate these EBs for students at different grade levels. Our newly-developed Likert-scale questionnaire is based on the most frequently used instrument for assessing students' EBs about science by Conley et al. (2004), which follows the widely-agreed four-dimensional structure of EBs (Hofer, 2000). In our adaption, pairs of analogous items simultaneously assess students' EBs related to professional and school science. We report a comparison of these EBs for students from grades 6, 8 and 10. Our study contributes to the literature by systematically describing students' EBs about both professional and school science, rather than focusing on only one context, and by providing researchers with a questionnaire that allows for simultaneous elicitation and direct comparison of these EBs.

## Theoretical background

### *The role of epistemic beliefs in science education*

The *Handbook of Epistemic Cognition* (Greene et al., 2016) impressively summarises the great variety and ramification of 'epistemic "something" terms' (p. 506, ff.) currently used in the literature. Notwithstanding the diversity of conceptualizations and intricacies, we refer to EBs as a multidimensional construct representing a set of personal theories and beliefs that one holds about the nature of knowledge and the process of knowing (Hofer & Pintrich, 1997). In this context, we use the term *epistemic beliefs* (e.g. Karakolidis et al., 2019; Lin & Tsai, 2017; Muis et al., 2006) instead of *epistemological beliefs* (e.g. Conley et al., 2004; Hofer & Pintrich, 1997; Schommer et al., 1992) in order to emphasise the focus on knowledge rather than the theory of knowledge (Kitchener, 2002).

With regard to science education, EBs have an important role for science learning, both as a prerequisite for learning and as a learning goal (Bromme et al., 2010; Kampa et al., 2016). For instance, students' EBs are related to their use of learning strategies (Lee et al., 2016; Schommer et al., 1992; Urhahne & Hopf, 2004), their construction of arguments and performance in scientific reasoning (Baytelman et al., 2020; Yang et al., 2019) and their academic achievement (Schommer et al., 1992; Stathopoulou &

Vosniadou, 2007). In line with the view of EBs as learning goals and related research on nature of science (e.g. Lederman et al., 2002), the view of knowledge, for example, as tentative, constructed and subjective is commonly considered sophisticated in EB assessment (Conley et al., 2004; Lin & Tsai, 2017; Nehring, 2020). However, this view of sophistication aligned with the interpretation of Likert-scales has been criticised as too simplistic (Elby et al., 2016; Hofer, 2016). Regarding the school science context and its particular demands, consideration of EBs in terms of productivity is suggested (Elby et al., 2016).

### ***Separating professional and school science in assessment of epistemic beliefs***

Researchers have repeatedly argued that the school context contributes to the development and change of EBs (Hogan, 2000; Muis et al., 2006; Sandoval, 2005). Theoretical support for the impact of school science experiences on EBs is provided by Hogan (2000) and Sandoval (2005). The authors propose to distinguish ‘distal and proximal knowledge of the nature of science’ (Hogan, 2000, p. 52) and students’ ‘formal epistemologies’ as well as ‘practical epistemologies’ (Sandoval, 2005, p. 635). Hogan (2000) bases her distinction on the perceived distance of personal experiences. Sandoval (2005) focuses on the context in which science is conducted, as well as on students’ active part in doing science. Further support for the impact of school science experiences on EBs comes from the Theories of Integrated Domains in Epistemology (TIDE) framework (Muis et al., 2006). The framework illustrates the impact of different contexts, including the school or academic context, in shaping and changing EBs. To empirically test the framework, Muis et al. (2016) analysed students’ rationales for their responses to an EB questionnaire. They found that school experiences shaped and influenced students’ explanations when referring to professional disciplines. Another study explored students’ beliefs about which sources inspired themselves or scientists to do science experiments (Elder, 2002). The author found that students more often reported scientists as actively constructing knowledge. Even though these studies show the importance of distinguishing and systematically assessing students’ EBs in professional and school science, studies explicitly comparing these two contexts are rare.

In our paper, we build on the conceptualizations by Hogan (2000) and Sandoval (2005) in terms of distinguishing between two contexts but restrict the scope of the considered constructs to the study of EBs. Therefore, we use the terms ‘EBs related to professional science’ and ‘EBs related to school science’ in order to emphasise the respective contexts of conducting science. Nevertheless, our notion of EBs related to professional science is very similar to Hogan’s (2000) notion of distal knowledge and Sandoval’s (2005) notion of formal epistemologies. However, our notion of EBs related to school science entails an important difference. This difference becomes evident in the study conducted by Burgin and Sadler (2013), who investigated the relation between students’ practical and formal epistemologies. During the highly authentic experience of completing an internship in a laboratory, students acquired first-hand experience of professional science. The study found that students’ practical and formal epistemologies were partially consistent. However, while the experiences acquired during the internship are considered as part of students’ practical epistemologies (or proximal knowledge), we would not assume them to be part of students’ EBs related to school science. In line with Muis

et al. (2006) and Sandoval (2016), we limit our definition of school science to science conducted in the educational context of institutionalised, formal science education in K12. Building on the use of the term by Sandoval (2005), we define professional science as science conducted by professional researchers in the context of for example, universities or private companies. In this regard, we understand and use the term ‘science’ as an umbrella term referring to various forms of science, including for example science conducted at school (‘school science’) or in a professional setting (‘professional science’). In particular, instead of ‘science’ the term ‘professional science’ is used to stress its difference from the school subject and emphasise its performance by scientists.

### **The assessment of epistemic beliefs about science**

Several instruments assess students’ EBs about science (Elby et al., 2016; Lee et al., 2021; Sandoval, 2016), with self-reported Likert-scale questionnaires being commonly used for the assessment of multidimensional EBs (Mason, 2016). Despite the limitations of Likert-scales, they facilitate the use of sophisticated inferential statistics and representative large-scale assessments to investigate relationships between EBs and students’ achievement, as well as other constructs (Hofer, 2016; Mason, 2016). Among the most cited instruments referring to the context of professional science (see Table 1) are the ones by Conley et al. (2004), Stathopoulou and Vosniadou (2007), and Tsai and Liu (2005). These instruments differ mainly in terms of the number of items, the examined EB dimensions, and their application to different study populations. The 26-item instrument by Conley et al. (2004) is the most established. It has frequently been implemented in various age groups (grade 5 and upwards) to assess EBs in the dimensions of Source, Certainty, Development, and Justification of knowledge.

**Table 1.** Most frequently cited papers that include instruments on students’ EBs about (professional) science.

| Authors (Year) & (Web of Science Citations)         | Description   |
|---|---|
| Conley et al. (2004)<br>(240 citations)             | <ul style="list-style-type: none"> <li>• 26 items</li> <li>• 5-point Likert-scale</li> <li>• 4 dimensions: source, certainty, development, and justification</li> <li>• initially administered to fifth-grade students</li> <li>• used in various age groups</li> <li>• various translations and adaptations available.</li> </ul>  |
| Stathopoulou and Vosniadou (2007)<br>(74 citations) | <ul style="list-style-type: none"> <li>• Greek Epistemological Beliefs Evaluation Instrument for Physics (GEBEP)</li> <li>• 40 statement items and 9 debate items</li> <li>• 4 dimensions: structure, stability, source, and justification</li> <li>• validated with tenth-grade students.</li> </ul>   |
| Tsai and Liu (2005)<br>(68 citations)               | <ul style="list-style-type: none"> <li>• 19 items</li> <li>• 5-point Likert-scale</li> <li>• 5 dimensions: role of social negotiation, invented and creative nature of science, theory-laden exploration, cultural impacts, and changing and tentative feature of science knowledge</li> <li>• validated for high school students (and their teachers)</li> <li>• 25 item version available for college students (Liu &amp; Tsai, 2008).</li> </ul> |

Note. Citations last updated: 03/23/2022.

Research on EBs related to school science, however, is scarce, with mainly qualitative data available, such as students' journals and laboratory write-ups (e.g. Schalk, 2012), interviews (e.g. Burgin & Sadler, 2013; Wu & Wu, 2011) and videotaped lessons (e.g. Hamza & Wickman, 2013; Wu & Wu, 2011). To the best of our knowledge, the Practical Epistemology in Science Survey (PESS) by Villanueva et al. (2019) is the first and only questionnaire explicitly assessing EBs related to school science in larger sample sizes. Similar to instruments available for professional science, the PESS is a self-report tool which consists of 26 Likert-scale items in the dimensions of immersive and non-immersive inquiry, and which measures students' views about their participation in school science. The items display high specificity and exclusiveness for the school-science context.

## Research questions

Despite efforts to account for the differences between the school-science and professional-science contexts in theoretical conceptualizations (Hogan, 2000; Sandoval, 2005), empirical studies and instruments systematically assessing this difference are rare. On the one hand, qualitative methods such as interviews or videotaped lessons are not designed for comparing students' EBs related to professional and school science across grade levels within larger samples (e.g. Burgin & Sadler, 2013; Wu & Wu, 2011). On the other hand, existing instruments with closed-ended items that allow for quantitative assessment of larger samples (e.g. Conley et al., 2004; Stathopoulou & Vosniadou, 2007; Tsai & Liu, 2005; Villanueva et al., 2019) do not address differences between students' EBs related to professional and school science, and presume different dimensionalities of EBs (Conley et al., 2004; Villanueva et al., 2019). Therefore, comparisons based on these questionnaires would be seriously impaired. In summary, science education research specifically lacks an instrument that allows for a quantitative assessment of EBs in larger samples and a simultaneous separation between students' EBs related to professional and school science. To address this gap, we developed a questionnaire that allows the separation and systematic comparison of students' EBs related to professional and school science.

The following two research questions (RQ) guided our research:

- (1) To what extent does the newly developed questionnaire measuring EBs related to school science and the adapted questionnaire measuring EBs related to professional science allow for forming reliable scales?
- (2) Do students hold differing EBs with regard to professional and school science at different grade levels?

## Methods

### *Study design and procedures*

We conducted our cohort study in German academic-track schools (*Gymnasium*) between 2017 and 2019 to investigate students' EBs related to professional and school science. The sample comprised 405 secondary school students (female = 221, male =

178, diverse = 3,  $M_{\text{age}} = 13.4$  years,  $SD_{\text{age}} = 1.60$ ). In our sample, 9.9% of the students initially learned a first language other than German and 20.7% were raised multi-lingual. The socioeconomic status within our sample was rather high, with 66.2% of the students having at least one parent who had attained the highest German educational degree A level (*Abitur*) and 53.6% of the students reporting that at least one parent studied at a university. In Germany, science education in academic track schools commonly teaches biology, chemistry, and physics separately, but not all subjects are being taught simultaneously in each school year. To account for different experience levels, we surveyed students from grade levels 6 ( $n = 132$ ), 8 ( $n = 167$ ) and 10 ( $n = 106$ ). Data collection took place in German during regular science classes. We obtained informed consent from all participants and their parents. No incentives were given. The Federal Ministry of Education, Science and Culture in Schleswig-Holstein, Germany, and the heads of schools approved the study protocol.

### **Development of the questionnaire**

We developed a questionnaire with two analogous item sets which assess students' EBs related to professional and school science and allow for direct comparison. Since it is the most frequently used instrument for assessing EBs about science (Lee et al., 2021) and first analytical checks for the use of a German version by Urhahne and Hopf (2004) are available, we used the questionnaire by Conley et al. (2004) as the basis for this process. The questionnaire assesses students' EBs about science in professional contexts. We extended its scope by adding parallel items that target analogous situations in school science.

Our questionnaire development followed three steps: First, we constructed general descriptions of the EB dimensions Source, Certainty, Development, and Justification for the context of professional science and school science. We summarised descriptions for the context of professional science provided by Conley et al. (2004) as well as related articles and formulated detailed descriptions related to professional science (see Supplemental Material 1). In order to form school-related descriptions of the EB dimensions (see Table 2), entities or situations exclusive to the context of professional science were systematically exchanged with phrases explicitly referring to school science.

Second, we constructed 26 items for the school science context that were analogous to the existing 26 items for the professional science context by Conley et al. (2004).

**Table 2.** Descriptions of the school-related EB dimensions.

| Dimension     | Description   |
|---------------|---|
| Source        | Comprises a learner's personal relation to authorities who are relevant and important in science class (teachers and textbooks). The source of knowledge in science class is either located as external or internal.  |
| Certainty     | Refers to personal views as to whether questions in science class have always one correct and absolute answer.  |
| Development   | Comprises the consideration of one's own knowledge in science class as either static or dynamic and subject to change. Progression and advancement of knowledge can be related to the (additive) extension, reorganisation or qualitative, content-wise changes of knowledge. |
| Justification | Comprises answers to the question of how knowledge can be achieved and justified in science class. Focuses on the role of experiments for scientific inquiry in science class.  |

However, the questionnaire by Conley et al. (2004) is not completely consistent in its use of terminology referring to the context of professional science. Some items in the dimension Source target school science rather than professional science. For these items we developed analogous item pairs with item siblings referring to one of the contexts each. We transformed the items by systematically exchanging phrases related to professional science with phrases referring to school science (see Supplemental Materials 2 and 3 for a table of analogous phrases). For example, the Certainty item *In scientific research, scientists always agree about what is true* that refers to professional science was transformed to the analogous school-related item *In science class, we always agree about what is true*. Analogous phrases resulted from a recursive process during which we checked the suitability of analogous phrases within a pair of items as well as across all items until we reached consistency.

Third, we pre-piloted the newly developed items for the school-science context in a paper-and-pencil assignment, as well as in short interviews with a focus group of eighth-graders. In the pre-piloting of the school-related items we checked whether referring to integrated school science (rather than separate school subjects) irritated students. Students commented on the questionnaire and provided feedback on the item formulations. Responses to several items showed ceiling effects or distinct peaks for example on the fourth response option and some students asked for a higher number of response categories ('I would like to have more possible options for answers'). Therefore, in addition to minor changes of item formulations, pre-piloting led to usage of a 7-point Likert-scale in order to increase the sensitivity of the scale. Finally, a panel of experts from science education ( $n = 3$ ) reviewed the revised items. All items were discussed and modified until consensus among experts and authors was reached.

## Measures

The questionnaire administered in the present study consisted of 26 analogous item pairs (five item pairs on Source, six on Certainty, six on Development, and nine on Justification) assessed with a 7-point Likert-scale ranging from 1 *do not agree at all* to 7 *completely agree*. It is common to interpret students' responses in terms of sophistication with higher agreement indicating higher sophistication (Conley et al., 2004). In contrast to Development and Justification, items on Source and Certainty were stated from a perspective referred to as less sophisticated by Conley et al. (2004). We randomised items on the four dimensions within the item sets related to professional and school science, with the school-related item set presented first (see full list of items in German and English in the Supplemental Materials 2 and 3).

## Psychometric properties of the questionnaire

First, we performed confirmatory factor analyses (CFA) to assess the questionnaire's dimensional structure in terms of model fit and factor loadings as well as the discriminatory power of the items. Separately specifying EBs related to professional and school science, we ran two CFAs with four factors and 26 items each. Benchmarks indicating a good model fit are values of at least .90 (or better .95) for Comparative Fit Index



(CFI) and Tucker-Lewis Index (TLI) and values of .08 (or better .06) for Root Mean Square Error of Approximation (RMSEA) and Standardised Root Mean Square Residual (SRMR) (Schreiber et al., 2006). We further analysed modification indices to identify possible sources of measurement error due to item formulation and considered standardised factor loadings greater .40 as acceptable. Items combined in a scale were selected to have a discriminatory power, that is, item-total correlation,  $r_{it} \geq 0.3$ . We calculated Cronbach's  $\alpha$  to measure scales' internal consistency and aimed for values above  $\alpha \geq .70$ . Since reliability coefficients of most self-report measures range between  $\alpha = .50$  and  $.70$  (Muis et al., 2006), we evaluated these coefficients as acceptable. In addition, we calculated the composite reliability (CR) and the average variance extracted (AVE). In line with Huang et al. (2013), we aimed for values above .70 for CR and higher than .40 or .50 for AVE (Fornell & Larcker, 1981). All three coefficients taken together give insights into the internal consistency as well as explained variance due to the construct. To maintain analogous item pairs, and therefore allow a direct comparison between EBs related to professional and school science, we excluded items and their respective siblings from the other context in further analyses, if they did not meet the aforementioned criteria (see Supplemental Material 4 for omission criteria applying to each item). We reviewed each excluded item pair critically for contextual support of its exclusion to ensure that scale formation did not impair the scope of the instrument. In total, 13 item pairs remained in the final instrument. An overview of the final scales' psychometric properties is provided in Table 3 (see Supplemental Material 5 for intercorrelations of final scales).

Second, we assessed whether EBs related to professional and school science should be considered as two separate constructs. For this purpose, we used CFAs to compare different theoretical models and analysed the extent to which EBs related to professional and school science can be distinguished empirically. We used both data sets for the respective CFAs and due to the parallel item formulations correlated the errors of analogous items for each item pair.

Based on our conceptualisation of EBs related to professional as well as school science, we assumed that the scales of our instrument would be best represented with a model in which the EB dimensions as well as the context were considered in the factor structure. Therefore, we theoretically defined a CFA model of eight factors, each representing one of the EB dimensions related either to the context of professional or school science (Model A). In Model A, item siblings load onto different factors.

Alternative models were analysed in order to compare the fit of competing models and rule out other theoretical conceptualisations (see Supplemental Material 6 for an overview of the competing models). In order to compare the competing models, we

**Table 3.** Overview of the EB scales.

| Scales measuring the EB dimensions | Number of items |       | Cronbach's $\alpha$ |       | Item-total correlation ( $r_{it}$ ) |           | Composite reliability |       | Average variance extracted |       |
|------------------------------------|-----------------|-------|---------------------|-------|-------------------------------------|-----------|-----------------------|-------|----------------------------|-------|
|                                    | School          | Prof. | School              | Prof. | School                              | Prof.     | School                | Prof. | School                     | Prof. |
|                                    | Source          | 3     | 3                   | .62   | .71                                 | .39 – .46 | .40 – .64             | .36   | .48                        | .36   |
| Certainty                          | 3               | 3     | .58                 | .68   | .38 – .39                           | .43 – .56 | .31                   | .42   | .31                        | .42   |
| Development                        | 3               | 3     | .55                 | .65   | .29 – .44                           | .38 – .51 | .32                   | .39   | .34                        | .40   |
| Justification                      | 4               | 4     | .60                 | .69   | .30 – .47                           | .37 – .52 | .35                   | .44   | .30                        | .37   |

Note. School = School science context; Prof. = Professional science context.

considered the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), that allow for direct comparison of the two models with a lower AIC and BIC representing a better model fit (Kline, 2016). Model B was a four-factor model with each factor representing one of the four EB dimensions and no distinction between the contexts of professional science and school science. Therefore, for each item pair, both analogous item siblings load onto the same factor. If the distinction of the professional and school science context was negligible, Model B should fit better than Model A. Assuming a distinction of the contexts, we also considered three second-order models (Models C, D, and E). Item siblings loaded on different factors within each of these models. Model C had four second-order factors, which each represent one of the four EB dimensions and determines two first-order factors representing context-dependent manifestations of the respective EB dimension. Model D considered the two contexts of professional and school science as higher-order factors with each higher-order factor determining four first-order factors representing context-dependent EB dimensions. Model E had two second-order factors with four related first-order factors each based on the correlation matrix of the final scales (see Supplemental Material 5). One second-order factor was defined to determine factors representing the context-dependent EB dimensions Source and Certainty, while the other second-order factor determined factors representing the context-dependent EB dimensions Development and Justification. Finally, we calculated two third-order models in which the latent variables of a second-order model correlated (Models F and G). Model F added two third-order factors to Model C with one of these factors determining the second-order factors representing the EB dimensions Source and Certainty and the other one determining the second-order factors representing the EB dimensions Development and Justification. Model G added a third-order factor to Model D that determined the two second-order factors of this model.

### **Data analysis**

Since they were stated from a perspective referred to as less sophisticated, we reverse-coded students' responses to items in the Source and Certainty scales and calculated descriptive statistics. We used one-sided *t*-tests to analyse differences from the mid-points of the 7-point Likert-scales as well as differences between students' EBs related to professional and school science. The former was calculated using aggregated data from all grade levels, the latter separately for each grade level. To measure the effect sizes, Cohen's *d* was calculated, with  $d = .20$  presenting a low,  $d = .50$  a medium, and  $d = .80$  a large effect (Cohen, 1988). A one-way analysis of variances (ANOVA) with post hoc comparisons using pairwise *t*-tests with pooled standard deviation and Bonferroni adjustment were used to identify age-dependent differences between students' EBs related to professional as well as school science. Eta-squared was used to report the ANOVA's effect size, with  $\eta^2 = .01$  presenting a low,  $\eta^2 = .06$  a medium, and  $\eta^2 = .14$  a large effect (Cohen, 1988). For statistical computing, we used the software R Studio 1.3.959 (RStudio Team, 2020). For the calculation of the CFA, we utilised the 'lavaan' 0.6–7 (Rosseel, 2012) package in R with robust maximum-likelihood estimators.

**Table 4.** Overview of CFA figures before and after scale formation.

| Model                         | $\chi^2$   | df  | CFI  | TLI  | RMSEA | SRMR | AIC       | BIC       |
|-------------------------------|------------|-----|------|------|-------|------|-----------|-----------|
| Model <sub>prof.pre</sub>     | 676.576*** | 293 | .848 | .832 | .065  | .076 | 32080.530 | 32416.441 |
| Model <sub>school.pre</sub>   | 684.990*** | 293 | .774 | .750 | .063  | .072 | 34040.789 | 34377.116 |
| Model <sub>prof.final</sub>   | 92.267**   | 59  | .965 | .954 | .043  | .040 | 16242.544 | 16422.496 |
| Model <sub>school.final</sub> | 71.670     | 59  | .978 | .971 | .025  | .039 | 17244.246 | 17424.421 |

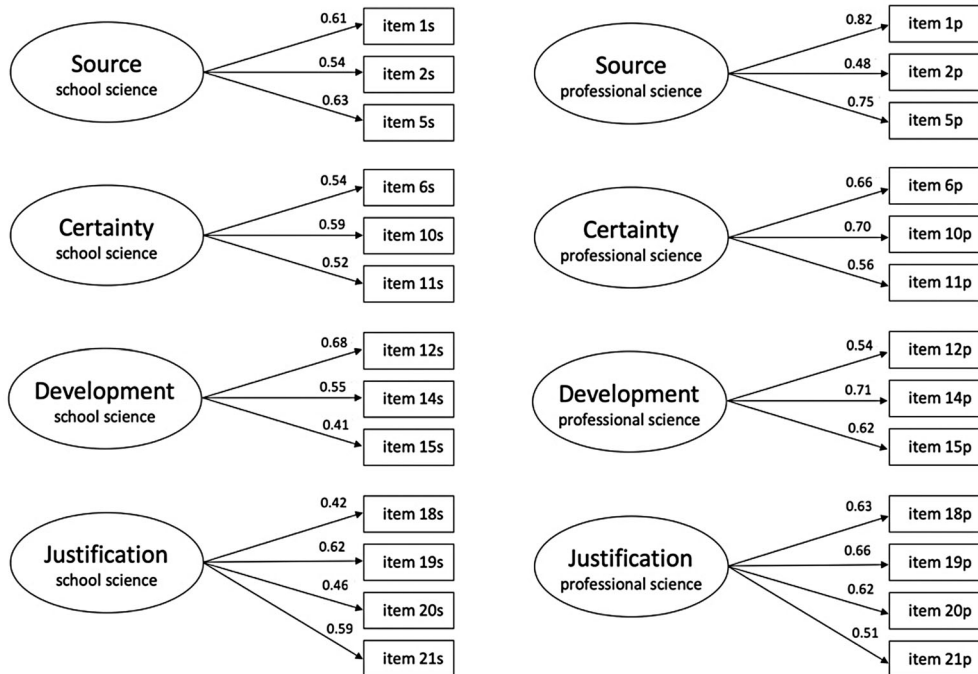
Note. CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square error of approximation; SRMR = standardised root mean square residual; AIC = Akaike information criterion; BIC = Bayesian information criterion; prof. = EBs related to professional science; school. = EBs related to school science; pre: full item set with 26 items before scale formation; final: item set after scale formation with 13 items.

\*\* $p < .01$ ; \*\*\* $p < .001$ .

## Results

### Scale formation and instrument development

In order to form the scales, we separately ran two CFAs each specifying four factors for the EB dimensions and using one of the item sets. Table 4 shows the respective CFA model fit indices for the two CFAs prior to and after scale formation. Factor loadings for the final scales are shown in Figure 1. Initially, the CFAs with the complete item sets containing 26 items showed poor model-fit for EBs related to professional as well as school science. The CFAs with the final scales formed from the remaining 13 analogous item pairs showed good model fit values.



**Figure 1.** CFA models with factor loadings for EBs related to school science (left) and EBs related to professional science (right). Note. Error terms and variances of the latent variables are not shown to avoid overloading the figures. s = school science context, p = professional science context.

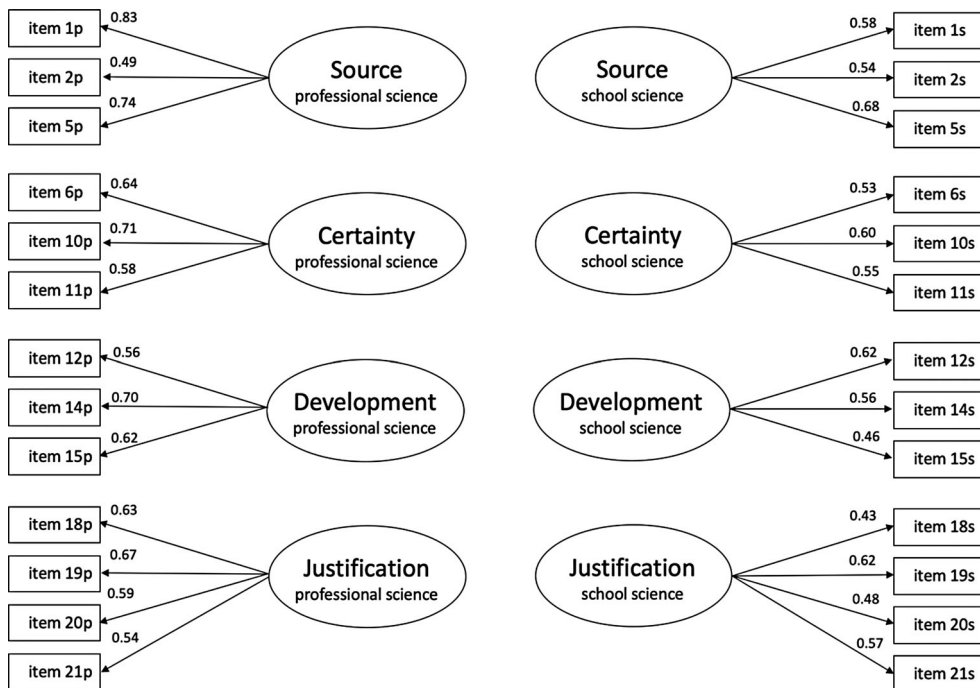
**Table 5.** Overview of CFA figures for competing models.

| Model   | $\chi^2$   | df  | CFI  | TLI  | RMSEA | SRMR | AIC       | BIC       |
|---------|------------|-----|------|------|-------|------|-----------|-----------|
| Model A | 355.308*** | 258 | .954 | .943 | .033  | .045 | 33021.222 | 33497.685 |
| Model B | 555.724*** | 280 | .872 | .851 | .054  | .059 | 33211.996 | 33600.373 |
| Model C | 441.039*** | 272 | .921 | .906 | .043  | .057 | 33091.805 | 33512.213 |
| Model D | 641.247*** | 277 | .830 | .801 | .062  | .099 | 33320.900 | 33721.288 |
| Model E | 467.786*** | 277 | .911 | .896 | .045  | .061 | 33113.838 | 33514.227 |
| Model F | NA         | NA  | NA   | NA   | NA    | NA   | NA        | NA        |
| Model G | 638.933*** | 276 | .830 | .800 | .062  | .099 | 33322.900 | 33727.292 |

Note. Model F could not be calculated after 989 iterations. CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square error of approximation; SRMR = standardised root mean square residual; AIC = Akaike information criterion; BIC = Bayesian information criterion; NA = not applicable.

\*\*\* $p < .001$ .

We compared the competing models in order to assess the underlying dimensional structure of EBs related to professional and school science, respectively (see Section *Psychometric properties of the questionnaire* and Supplemental Material 6 for an overview of compositional differences of the competing models). In support of the distinction between professional and school science, the suggested eight factor-model (Model A) obtained the lowest values for AIC and BIC compared to the alternative models (cf. Table 5). Thus, Model A obtained the best fit with the data. In addition, Model A showed an acceptable model fit (CFI = .95; TLI = .94; RMSEA = .03; SRMR = .05),



**Figure 2.** CFA model with factor loadings for Model A. Note. Correlations between errors, error terms and variances of the latent variables are not shown to avoid overloading the figures. s = school science context, p = professional science context.

**Table 6.** Findings from the ANOVA for students' EB related to professional and school science at different grade levels.

| Scales        | Grade 6                  |                         | Grade 8                  |                         | Grade 10                 |                         | ANOVA               |                    |
|---------------|--------------------------|-------------------------|--------------------------|-------------------------|--------------------------|-------------------------|---------------------|--------------------|
|               | $M_{\text{School}}$ (SD) | $M_{\text{Prof.}}$ (SD) | $M_{\text{School}}$ (SD) | $M_{\text{Prof.}}$ (SD) | $M_{\text{School}}$ (SD) | $M_{\text{Prof.}}$ (SD) | $F_{\text{School}}$ | $F_{\text{Prof.}}$ |
| Source        | 4.20 (1.09)              | 4.43 (1.19)             | 4.09 (1.14)              | 4.37 (1.19)             | 4.23 (1.16)              | 4.47 (1.12)             | 4.41*               | 2.30               |
| Certainty     | 4.18 (1.19)              | 4.45 (1.21)             | 4.36 (1.08)              | 4.96 (1.12)             | 4.76 (1.02)              | 5.50 (0.99)             | 8.35***             | 26.03***           |
| Development   | 5.25 (0.98)              | 5.18 (1.02)             | 5.20 (0.94)              | 5.50 (0.90)             | 5.35 (0.89)              | 5.82 (0.76)             | 0.81                | 14.71***           |
| Justification | 5.90 (0.73)              | 5.56 (0.95)             | 5.57 (0.80)              | 5.62 (0.81)             | 5.41 (0.81)              | 5.70 (0.72)             | 12.22***            | 0.80               |

Note.  $M$  = mean;  $SD$  = standard deviation;  $F$  = F-Test statistic from analysis of variance (ANOVA); School = School science context; Prof. = Professional science context.

\* $p < .05$ ; \*\* $p < .01$ ; \*\*\* $p < .001$ .

whereas the other models did not or only barely meet the benchmarks for good model fit. Factor loadings for Model A are displayed in [Figure 2](#).

### ***Students' epistemic beliefs related to professional and school science***

[Table 6](#) provides an overview of the descriptive statistics for the four EB scales with respect to the contexts of professional and school science. Across the four EB dimensions Source, Certainty, Development, and Justification mean values were higher than the midpoint of the 7-point Likert-scales for EBs related to school science,  $T_{\text{source}}(402) = 21.75$ ,  $p < .001$ ;  $T_{\text{certainty}}(404) = 25.18$ ,  $p < .001$ ;  $T_{\text{development}}(403) = 48.33$ ,  $p < .001$ ;  $T_{\text{justification}}(403) = 65.76$ ,  $p < .001$ , and EBs related to professional science,  $T_{\text{source}}(398) = 24.92$ ,  $p < .001$ ;  $T_{\text{certainty}}(401) = 32.75$ ,  $p < .001$ ;  $T_{\text{development}}(399) = 53.02$ ,  $p < .001$ ;  $T_{\text{justification}}(398) = 62.77$ ,  $p < .001$ . Thus, students beliefs are rather similar to views commonly accepted as more sophisticated (Conley et al., 2004; Lederman et al., 2002). Therefore, higher scores on the scales indicate that students tend to believe more strongly in knowledge originating from sources other than authorities (Source), in the existence of multiple solutions for a given issue (Certainty), in knowledge to change over time (Development), and in experiments as an appropriate way of scientific inquiry (Justification). This can be observed for both contexts.

We also found significantly different mean values for students' EBs related to professional science and school science. For Source, students expressed stronger agreement with items in the scale referring to professional science compared to school science in all grades,  $T_{\text{grade6}}(125) = -2.86$ ,  $p < .05$ ;  $T_{\text{grade8}}(164) = -3.64$ ,  $p < .001$ ;  $T_{\text{grade10}}(105) = -1.80$ ,  $p < .05$ . The corresponding small-to-medium effect sizes ranged between  $d_{\text{grade6}} = 0.21$ ,  $d_{\text{grade8}} = 0.24$ , and  $d_{\text{grade10}} = 0.15$ . For Certainty, again higher agreement with items related to professional science existed,  $T_{\text{grade6}}(129) = -3.02$ ,  $p < .05$ ;  $T_{\text{grade8}}(165) = -7.32$ ,  $p < .001$ ;  $T_{\text{grade10}}(105) = -8.85$ ,  $p < .001$ , and effects were of medium and large sizes,  $d_{\text{grade6}} = 0.22$ ;  $d_{\text{grade8}} = 0.55$ ;  $d_{\text{grade10}} = 1.00$ . For Development, we found no differences between the mean values of students' EBs related to professional and school science at grade 6,  $T(126) = 0.52$ ,  $p > .05$ . However, at grades 8 and 10 students again expressed higher agreement with items referring to professional science,  $T_{\text{grade8}}(165) = -3.89$ ,  $p < .001$ ;  $T_{\text{grade10}}(105) = -4.62$ ,  $p < .001$ , with corresponding effects of medium sizes,  $d_{\text{grade8}} = 0.32$ ;  $d_{\text{grade10}} = 0.24$ . For Justification, higher agreement with items referring to professional science was found for tenth-graders,  $T_{\text{grade10}}(105) = -3.42$ ,  $p < .001$ , with medium effect size,  $d_{\text{grade10}} = 0.37$ . No significant differences between the mean values of EBs related to professional science and school science were found for eighth-graders,  $T_{\text{grade8}}(164) = -.52$ ,  $p > .05$ . In contrast to the previous findings, we found sixth-graders were in higher agreement with items in the school-related scale,  $T_{\text{grade6}}(125) = 4.51$ ,  $p < .001$ , with a corresponding medium-sized effect  $d_{\text{grade6}} = 0.40$ .

### ***Differences in students' epistemic beliefs related to professional and school science at different grade levels***

To analyse differences between grade levels, we performed two separate ANOVAs, with EBs related to professional and school science as dependent variables (see [Table 6](#)). For EBs related to professional science, we revealed significant main effects for grade levels

regarding the Certainty and Development dimensions,  $F_{\text{certainty}}(2, 399) = 26.03, p < .001$ ;  $F_{\text{development}}(2, 397) = 14.71, p < .001$ , with corresponding medium-sized effects,  $\eta^2_{\text{certainty}} = .116$ ;  $\eta^2_{\text{development}} = .069$ . Post hoc analyses revealed that students were in significantly higher agreement with the EB scales at higher grade levels. Differences for the Justification and Source dimensions were not statistically significant,  $F_{\text{justification}}(2, 396) = 0.80, p = .450$ ;  $F_{\text{source}}(2, 396) = 2.30, p = .101$ .

For EBs related to school science, we identified a significant main effect for grade levels regarding Source,  $F_{\text{source}}(2, 400) = 4.41, p = .013$ , Certainty,  $F_{\text{certainty}}(2, 402) = 8.35, p < .001$ , and Justification,  $F_{\text{justification}}(2, 401) = 12.22, p < .001$ . The corresponding medium effect sizes ranged between  $\eta^2_{\text{source}} = .022$ ,  $\eta^2_{\text{certainty}} = .040$ , and  $\eta^2_{\text{justification}} = .057$ . For Source, post hoc analyses revealed that significant differences in students' agreement occurred solely between eighth- and tenth-graders. For Certainty, post hoc analyses revealed significantly different mean values between sixth- and tenth-graders as well as eighth- and tenth-graders. For Justification, in contrast, post hoc analyses revealed decreasing agreement with higher grade levels between grades 6 and 8 as well as grades 6 and 10. For Development, we found no significant main effect for grade levels,  $F_{\text{development}}(2, 401) = 0.81, p = .447$ , indicating that students' agreement with the scales did not differ significantly between grade levels (see Table 6).

## Discussion

In our study, we explored students' EBs related to professional and school science in grades 6, 8, and 10. For this purpose, we developed a questionnaire based on the instrument by Conley et al. (2004) that allows for simultaneous assessment of students' EBs related to professional and school science. To our best knowledge, and despite recent calls within the debate on science EBs, our study is the first quantitative attempt to compare these two contexts, rather than explore only one of the contexts. First, we succeeded in forming reliable scales of EBs related to professional and school science. Second, data analyses revealed that the EBs can empirically be separated as two distinct constructs. We discuss our findings with respect to the research questions focusing on the comparison of EBs related to professional and school science (RQ 2) and the formation of EB scales (RQ 1).

### *Comparison between epistemic beliefs related to professional and school science*

As one key finding of our study, we showed that EBs related to school science show similar characteristics to EBs related to professional science in terms of level of agreement within, as well as between, grade levels.

For the four EB dimensions, the mean values were above the midpoint of the 7-point Likert-scales. Because prior empirical studies applying the questionnaire by Conley et al. (2004) reported similar findings (Conley et al., 2004; Osborne et al., 2013; Urhahne & Hopf, 2004), we expected these findings for EBs related to professional science. Studies investigating students' EBs related to school science applied different instruments but also reported relatively high agreement (e.g. Villanueva et al., 2019). By applying our newly-developed instrument, we could show

that school-related EBs seem to have similar properties to EBs related to professional science but nevertheless are significantly different in direct comparisons of the scales.

Furthermore, for both contexts, we found differences in the expressed agreement with EB scales between grade levels for some of the dimensions. For EBs related to professional science, the differences were consistent, with significantly higher agreement occurring for higher grade levels. This is in line with existing related research on students' views about science and scientific inquiry (Reith & Nehring, 2020; Solomon et al., 1996) and studies on EBs that mostly reported higher scores at higher grade levels, although these studies varied regarding the dimensions and grade levels in which significant differences occurred (Kremer & Mayer, 2013; Osborne et al., 2013; Urhahne et al., 2008). For EBs related to school science, our findings were less consistent. In particular, for the school-related Justification scale, we found significantly lower scores for higher grade levels compared to the score for grade 6. Previously, this phenomenon of younger students expressing higher agreement with Justification items had also occurred in German samples (Kremer & Mayer, 2013; Urhahne et al., 2008). The authors' explanation referred to the German curriculum, in which considerable time and effort are applied, especially in early secondary education, to teaching with experimental practices. However, we cannot rule out other influences, as our sample was, for example, heterogeneous in terms of different science teachers. The growing content knowledge in the area of science for the older students could be a factor that shapes the students' view to become more sceptical about science epistemology. However, research indicates that there is not necessarily a linear relationship between the increase of scientific knowledge and a more sceptical stance towards science epistemology (Bromme et al., 2008).

Finally, the empirical results of our study support our central assumption of substantial differences between EBs related to school science and EBs related to professional science. Our analysis within grade levels allowed us to compare students' responses to the EB scales related to professional and school science. With two exceptions (Development: grade 6; Justification: grade 8), we found significant differences between the school and professional contexts for all scales and grade levels in favour of professional science. The only exception from this pattern occurred in grade 6 for Justification, where higher mean values occurred for EBs related to school science. These findings can be interpreted with respect to the dual nature of EBs as learning goals as well as prerequisites for learning (Bromme et al., 2010; Kampa et al., 2016). If EBs are primarily considered as learning goals, stronger agreement with the scales of the questionnaire might be referred to as sophisticated (Kampa et al., 2016; Lederman et al., 2002). This interpretation in terms of sophistication is commonly applied for EBs related to professional science (Conley et al., 2004; Kizilgunes et al., 2009). However, researchers have pointed out that interpretations of EBs in terms of sophistication fail to acknowledge the particular demands inherent to educational environments (Elby et al., 2016; Hofer, 2016). The separate consideration of EBs related to school science suggests a middle ground in this ongoing debate. While EBs, when considered as learning goals, might be interpreted in terms of sophistication (Conley et al., 2004), EBs primarily considered as prerequisites for learning might be viewed through the lens of productivity (Elby et al., 2016).



### **Formation of epistemic beliefs scales**

In EB research, one of the main challenges is the formation of EB scales that allow for a reliable assessment (Mason, 2016; Muis et al., 2006). Studies showed that high levels of Cronbach's  $\alpha$  cannot necessarily be expected and that the reported levels of internal consistency vary greatly between samples (Kizilgunes et al., 2009; Kremer & Mayer, 2013; Urhahne & Hopf, 2004). In a review article, Muis et al. (2006) found that reliability coefficients of self-report measures on students' EBs commonly range between  $\alpha = .50$  and  $.70$ . In line with these former results, reliability measured in our study ranged between  $\alpha = .55$  and  $.62$  for EBs related to school science and  $\alpha = .65$  and  $.71$  for EBs related to professional science. The AVE ranges between  $.30$  and  $.36$  for EBs related to school science and between  $.37$  and  $.48$  for EBs related to professional science. While according to the boundaries suggested by Huang et al. (2013) AVE values may be acceptable for three out of four scales for EBs related to professional science, AVE values were not yet satisfactory for the school science item set. For both item sets CR was below the aspired values ranging between  $CR = .31$  and  $.36$  for EBs related to school science and  $CR = .39$  and  $.48$  for EBs related to professional science. However, as these criteria were not reported in the publications providing the original English and German version of the EB questionnaire by Conley et al. (2004) and Urhahne and Hopf (2004), we are not able to compare our results to former studies. We therefore need to focus on Cronbach's  $\alpha$ . The reliabilities found in our study for Cronbach's  $\alpha$  are within the range commonly found and reported in other studies on EBs (Muis et al., 2006). At the same time, findings from our study on the fit of the model as well as the factor loadings are well within the cut-off criteria and thus provide support for the overall structure of the eight factor model (Hair et al., 2010). Regarding the scales related to professional science, we found reliabilities equal to or above the estimates for the German translation (Urhahne & Hopf, 2004) and the questionnaire by Conley et al. (2004). Regarding the scales related to school science, we obtained lower reliability in comparison to the two scales of the PESS (Villanueva et al., 2019). Although Cronbach's  $\alpha$  were satisfactory for all scales, consideration of AVE and CR values as well as the reliabilities obtained for the PESS promote the desire for further improvement of the questionnaire as well as future research in this regard. Also considering the amount of item pairs that were excluded during scale formation, it might be beneficial to develop additional item pairs, in order to further improve the reliability of the instrument and especially the scales related to school science.

During scale formation, we had to exclude 13 of the initial 26 item pairs. The number of excluded items resulted primarily from the pairwise omission of items which we performed in favour of parallelism of the two item sets. Thus, the scale formation left us with scales that allow for direct comparisons between the contexts, while, at the same time, the diminishing number of items per dimension might have resulted in lower reliabilities. However, if we had refrained from pairwise item exclusion and only omitted individual items, we would have produced heterogeneous scales. These scales would then not have been comparable. However, our procedure affected the scope of the Justification dimension. In contrast to the original conceptualisation by Conley et al. (2004), the Justification scales of our study do not cover the use of multiple methods in science, and solely focus on the role of experiments. However, this emphasis may be acceptable with respect to the

important role of experiments in school science practice (Arnold et al., 2014; Osborne et al., 2013).

The comparison of competing models revealed that an eight-factor model accounting for both, the context and dimension of science EBs, best represented the data. This is in line with our theoretical conceptualisation of EBs related to school science separately from EBs related to professional science. Our study therefore provides empirical evidence that favours the distinction of EBs related to different contexts of science over their joined consideration. Furthermore, comparison of competing models suggests a relationship between the EB dimensions and the considered context that is not dominated by either of these factors. Finally, this finding provides empirical support for the claims of other authors, who argued for a distinction between the contexts of professional and school science (e.g. Hogan, 2000; Sandoval, 2005).

### **Limitations**

Even though our study sheds light on a very important topic within the discussion on science EBs, we need to outline some limitations of our study. Firstly, differences in the reported degrees of freedom occurred in the reported t-tests and ANOVAs. These differences resulted from missing data for individual students. Secondly, our results are based on cross-sectional data. Therefore, our cohort study can only give first hints on longitudinal inferences regarding the development of EBs during secondary education. We cannot rule out bias across cohorts nor effects from students first answering EB scales related to school science and subsequently answering scales related to professional science. Therefore, higher agreement with scales for professional science might be due to sequential or positional effects. Future studies may address this limitation by longitudinally investigating EBs related to school science and EBs related to professional science simultaneously, using a balanced questionnaire design. Finally, due to the missing baseline data for the German and English versions of the original questionnaire regarding the AVE and CR values, reasons for our low values are difficult to determine. Prospectively, investigating and reporting external psychometric qualities such as the consideration of relations between the assessed EBs and learning related variables might be a promising approach to foster understanding about the relationship between EBs related to professional and school science.

### **Directions for further research and conclusion**

The results and the questionnaire presented in this study lay the groundwork for future research on the relationship between students' science learning and their EBs related to professional and school science. To our knowledge, our study differentiates for the first time empirically between EB concerning school and professional science. Further research has to elaborate specifically the interplay between instructional and learning-related effects on the formation of these beliefs. First, correlational studies in large and preferably representative samples may be conducted to assess whether students' EBs related to professional and school science possess differing relevance for students' learning. Second, since we developed parallel item sets for EBs related to professional and school science, differential manifestations of the respective EBs can be investigated.

Researchers should consider if future interventional studies designed to foster EBs might benefit from using our questionnaire in addition to qualitative, procedural methods. Finally, a supplemental assessment using a mixed methods approach seems particularly fruitful in the light of interpreting EBs related to school science beyond common views of sophistication and instead through the lens of productivity as suggested for example, by Elby et al. (2016). In particular, using the developed instrument in a mixed methods approach might allow to link assessments of ‘practical epistemologies’ (Sandoval, 2005, p. 635) with correlational analyses linked to prior research on ‘epistemological beliefs in science’ (Conley et al., 2004, p. 187). In this regard, our questionnaire provides a promising tool for researchers to address and hopefully bridge the gap between ‘students’ epistemic beliefs and their practices in school science’ (Lee et al., 2021, p.899). However, with regard to science teaching, continued efforts are needed to gain a comprehensive understanding of students’ EBs related to professional and school science that will provide teachers with powerful knowledge to bridge students’ ‘parallel ways of knowing’ (Edmondson & Novak, 1993, p. 550) and efficiently foster students’ EBs about science.

In our study, we successfully assessed students’ EBs related to professional and school science. Initially, we faced the question of how best to assess these EBs. We had the two options of either using existing but structurally and conceptually different instruments, or developing a new questionnaire that allows for parallel assessment and ensures comparability. We chose the latter and believe that our results affirm our decision. Our results indicate that we succeeded in separately measuring EBs related to professional and school science and provide empirical support for their distinction, which allows for inferences with respect to both contexts.

## Acknowledgements

We would like to thank Irene Neumann (IPN Kiel) and Julia Schwanewedel (Universität Hamburg) for the fruitful discussions during the development of the research idea and the measuring instrument.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## ORCID

Frauke Voitle  <http://orcid.org/0000-0002-1795-7170>

Benedikt Heuckmann  <http://orcid.org/0000-0001-5008-6031>

Nele Kampa  <http://orcid.org/0000-0001-5633-9335>

Kerstin Kremer  <http://orcid.org/0000-0003-2305-3650>

## References

- Arnold, J. C., Kremer, K., & Mayer, J. (2014). Understanding students’ experiments—what kind of support do they need in inquiry tasks? *International Journal of Science Education*, 36(16), 2719–2749. <https://doi.org/10.1080/09500693.2014.930209>
- Baytelman, A., Iordanou, K., & Constantinou, C. P. (2020). Epistemic beliefs and prior knowledge as predictors of the construction of different types of arguments on socioscientific issues. *Journal of Research in Science Teaching*, 57(8), 1199–1227. <https://doi.org/10.1002/tea.21627>

- Bernholt, A., Lindfors, M., & Winberg, M. (2021). Students' epistemic beliefs in Sweden and Germany and their interrelations with classroom characteristics. *Scandinavian Journal of Educational Research*, 65(1), 54–70. <https://doi.org/10.1080/00313831.2019.1651763>
- Bromme, R., Kienhues, D., & Stahl, E. (2008). Knowledge and epistemological beliefs: An intimate but complicate relationship. In M. S. Khine (Ed.), *Knowing, knowledge and beliefs* (pp. 423–441). Springer. [https://doi.org/10.1007/978-1-4020-6596-5\\_20](https://doi.org/10.1007/978-1-4020-6596-5_20)
- Bromme, R., Pieschl, S., & Stahl, E. (2010). Epistemological beliefs are standards for adaptive learning: A functional theory about epistemological beliefs and metacognition. *Metacognition and Learning*, 5(1), 7–26. <https://doi.org/10.1007/s11409-009-9053-5>
- Burgin, S. R., & Sadler, T. D. (2013). Consistency of practical and formal epistemologies of science held by participants of a research apprenticeship. *Research in Science Education*, 43(6), 2179–2206. <https://doi.org/10.1007/s11165-013-9351-4>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*, 2nd edn. Á/L. Erlbaum Press.
- Conley, A. M., Pintrich, P. R., Vekiri, I., & Harrison, D. (2004). Changes in epistemological beliefs in elementary science students. *Contemporary Educational Psychology*, 29(2), 186–204. <https://doi.org/10.1016/j.cedpsych.2004.01.004>
- Edmondson, K. M., & Novak, J. D. (1993). The interplay of scientific epistemological views, learning strategies, and attitudes of college students. *Journal of Research in Science Teaching*, 30(6), 547–559. <https://doi.org/10.1002/tea.3660300604>
- Elby, A., Macrander, C., & Hammer, D. (2016). Epistemic cognition in science. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Educational psychology handbook. Handbook of epistemic cognition* (pp. 113–127). Taylor and Francis.
- Elder, A. D. (2002). Characterizing fifth grade students' epistemological beliefs in science. In B. K. Hofer, & P. R. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 347–363). Routledge.
- Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50. <https://doi.org/10.1177/002224378101800104>
- Greene, J. A., Sandoval, W. A., & Bråten, I. (Eds.). (2016). *Educational psychology handbook. Handbook of epistemic cognition*. Taylor and Francis.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis (7th ed.)*. Prentice-Hall, Inc.
- Hamza, K. M., & Wickman, P.-O. (2013). Student engagement with artefacts and scientific ideas in a laboratory and a concept-mapping activity. *International Journal of Science Education*, 35(13), 2254–2277. <https://doi.org/10.1080/09500693.2012.743696>
- Hofer, B. K. (2000). Dimensionality and disciplinary differences in personal epistemology. *Contemporary Educational Psychology*, 25(4), 378–405. <https://doi.org/10.1006/ceps.1999.1026>
- Hofer, B. K. (2016). Epistemic cognition as a psychological construct: Advancements and challenges. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Educational psychology handbook. Handbook of epistemic cognition* (pp. 31–50). Taylor and Francis.
- Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of Educational Research*, 67(1), 88–140. <https://doi.org/10.3102/00346543067001088>
- Hogan, K. (2000). Exploring a process view of students' knowledge about the nature of science. *Science Education*, 84(1), 51–70. [https://doi.org/10.1002/\(SICI\)1098-237X\(200001\)84:1<51::AID-SCE5>3.0.CO;2-H](https://doi.org/10.1002/(SICI)1098-237X(200001)84:1<51::AID-SCE5>3.0.CO;2-H)
- Huang, C. C., Wang, Y. M., Wu, T. W., & Wang, P. A. (2013). An empirical analysis of the antecedents and performance consequences of using the moodle platform. *International Journal of Information and Education Technology*, 3(2), 217. <https://doi.org/10.7763/IJJET.2013.V3.267>
- Kampa, N., Neumann, I., Heitmann, P., & Kremer, K. (2016). Epistemological beliefs in science—a person-centered approach to investigate high school students' profiles. *Contemporary Educational Psychology*, 46, 81–93. <https://doi.org/10.1016/j.cedpsych.2016.04.007>

- Karakolidis, A., Pitsia, V., & Emvalotis, A. (2019). The case of high motivation and low achievement in science: What is the role of students' epistemic beliefs? *International Journal of Science Education*, 1457–1474. <https://doi.org/10.1080/09500693.2019.1612121>
- Khishfe, R., & Lederman, N. G. (2007). Relationship between instructional context and views of nature of science. *International Journal of Science Education*, 29(8), 939–961. <https://doi.org/10.1080/09500690601110947>
- Kitchener, R. F. (2002). Folk epistemology: An introduction. *New Ideas in Psychology*, 20(2-3), 89–105. [https://doi.org/10.1016/S0732-118X\(02\)00003-X](https://doi.org/10.1016/S0732-118X(02)00003-X)
- Kizilgunes, B., Tekkaya, C., & Sungur, S. (2009). Modeling the relations among students' epistemological beliefs, motivation, learning approach, and achievement. *The Journal of Educational Research*, 102(4), 243–256. <https://doi.org/10.3200/JOER.102.4.243-256>
- Kline, R. B. (2016). *Principles and Practice of Structural Equation Modeling* (4th ed.). New York, NY: The Guilford Press.
- Kremer, K., & Mayer, J. (2013). Entwicklung und Stabilität von Vorstellungen über die Natur der Naturwissenschaften [Development and stability of conceptions of the nature of science]. *Zeitschrift für Didaktik der Naturwissenschaften*, 19, 77–101.
- Lederman, N. G., Abd-El-Khalick, F., Bell, R. L., & Schwartz, R. S. (2002). Views of nature of science questionnaire: Toward valid and meaningful assessment of learners' conceptions of nature of science. *Journal of Research in Science Teaching*, 39(6), 497–521. <https://doi.org/10.1002/tea.10034>
- Lee, S.W.-Y., Liang, J.-C., & Tsai, C.-C. (2016). Do sophisticated epistemic beliefs predict meaningful learning? Findings from a structural equation model of undergraduate biology learning. *International Journal of Science Education*, 38(15), 2327–2345. <https://doi.org/10.1080/09500693.2016.1240384>
- Lee, S. W.-Y., Luan, H., Lee, M.-H., Chang, H.-Y., Liang, J.-C., Lee, Y.-H., Lin, T.-J., Wu, A.-H., Chiu, Y.-J., & Tsai, C.-C. (2021). Measuring epistemologies in science learning and teaching: A systematic review of the literature. *Science Education*, 105(5), 880–907. <https://doi.org/10.1002/sce.v105.5>
- Lin, T.-J., & Tsai, C.-C. (2017). Developing instruments concerning scientific epistemic beliefs and goal orientations in learning science: A validation study. *International Journal of Science Education*, 39(17), 2382–2401. <https://doi.org/10.1080/09500693.2017.1384593>
- Liu, S.-Y., & Tsai, C.-C. (2008). Differences in the scientific epistemological views of undergraduate students. *International Journal of Science Education*, 30(8), 1055–1073. <https://doi.org/10.1080/09500690701338901>
- Mason, L. (2016). Psychological perspectives on measuring epistemic cognition. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Educational psychology handbook. Handbook of epistemic cognition* (pp. 375–392). Taylor and Francis.
- Muis, K. R., Bendixen, L. D., & Haerle, F. C. (2006). Domain-generality and domain-specificity in personal epistemology research: Philosophical and empirical reflections in the development of a theoretical framework. *Educational Psychology Review*, 18(1), 3–54. <https://doi.org/10.1007/s10648-006-9003-6>
- Muis, K. R., Trevors, G., Duffy, M., Ranellucci, J., & Foy, M. J. (2016). Testing the TIDE: Examining the nature of students' epistemic beliefs using a multiple methods approach. *The Journal of Experimental Education*, 84(2), 264–288. <https://doi.org/10.1080/00220973.2015.1048843>
- Nehring, A. (2020). Naïve and informed views on the nature of scientific inquiry in large-scale assessments: Two sides of the same coin or different currencies? *Journal of Research in Science Teaching*, 57(4), 510–535. <https://doi.org/10.1002/tea.21598>
- OECD (Ed.). (2019). *Pisa. PISA 2018 assessment and analytical framework*. OECD. <https://doi.org/10.1787/19963777>
- Osborne, J., Simon, S., Christodoulou, A., Howell-Richardson, C., & Richardson, K. (2013). Learning to argue: A study of four schools and their attempt to develop the use of argumentation as a common instructional practice and its impact on students. *Journal of Research in Science Teaching*, 50(3), 315–347. <https://doi.org/10.1002/tea.21073>
- Reith, M., & Nehring, A. (2020). Scientific reasoning and views on the nature of scientific inquiry: Testing a new framework to understand and model epistemic cognition in science. *International Journal of Science Education*, 42(16), 2716–2741. <https://doi.org/10.1080/09500693.2020.1834168>

- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36. <https://doi.org/10.18637/jss.v048.i02>
- RStudio: Integrated development environment for R [Computer software]. (2020). RStudio, PBC. Boston, MA. <http://www.rstudio.com/>
- Sandoval, W. A. (2005). Understanding students' practical epistemologies and their influence on learning through inquiry. *Science Education*, 89(4), 634–656. <https://doi.org/10.1002/sc.20065>
- Sandoval, W. A. (2016). Disciplinary insights into the study of epistemic cognition. In J. A. Greene, W. A. Sandoval, & I. Bråten (Eds.), *Educational psychology handbook. Handbook of epistemic cognition* (pp. 184–194). Taylor and Francis.
- Schalk, K. A. (2012). A socioscientific curriculum facilitating the development of distal and proximal NOS conceptualizations. *International Journal of Science Education*, 34(1), 1–24. <https://doi.org/10.1080/09500693.2010.546895>
- Schiefer, J., Golle, J., Tibus, M., Herbein, E., Gindele, V., Trautwein, U., & Oschatz, K. (2020). Effects of an extracurricular science intervention on elementary school children's epistemic beliefs: A randomized controlled trial. *British Journal of Educational Psychology*, 90(2), 382–402. <https://doi.org/10.1111/bjep.12301>
- Schommer, M., Crouse, A., & Rhodes, N. (1992). Epistemological beliefs and mathematical text comprehension: Believing it is simple does not make it so. *Journal of Educational Psychology*, 84(4), 435. <https://doi.org/10.1037/0022-0663.84.4.435>
- Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research*, 99(6), 323–338. <https://doi.org/10.3200/JOER.99.6.323-338>
- She, H.-C., Lin, H., & Huang, L.-Y. (2019). Reflections on and implications of the programme for international student assessment 2015 (PISA 2015) performance of students in Taiwan: The role of epistemic beliefs about science in scientific literacy. *Journal of Research in Science Teaching*, 56(10), 1309–1340. <https://doi.org/10.1002/tea.21553>
- Solomon, J., Scott, L., & Duveen, J. (1996). Large-scale exploration of pupils' understanding of the nature of science. *Science Education*, 80(5), 493–508. [https://doi.org/10.1002/\(SICI\)1098-237X\(199609\)80:5<493::AID-SCE1>3.0.CO;2-6](https://doi.org/10.1002/(SICI)1098-237X(199609)80:5<493::AID-SCE1>3.0.CO;2-6)
- Stathopoulou, C., & Vosniadou, S. (2007). Exploring the relationship between physics-related epistemological beliefs and physics understanding. *Contemporary Educational Psychology*, 32(3), 255–281. <https://doi.org/10.1016/j.cedpsych.2005.12.002>
- Tsai, C.-C., & Liu, S.-Y. (2005). Developing a multi-dimensional instrument for assessing students' epistemological views toward science. *International Journal of Science Education*, 27(13), 1621–1638. <https://doi.org/10.1080/09500690500206432>
- Urhahne, D., & Hopf, M. (2004). Epistemologische Überzeugungen in den naturwissenschaften und ihre zusammenhänge mit motivation, selbstkonzept und lernstrategien [epistemological beliefs in science and their relationships with motivation, self-concept and learning-strategies]. *Zeitschrift für Didaktik der Naturwissenschaften*, 10(1), 71–87.
- Urhahne, D., Kremer, K., & Mayer, J. (2008). Welches verstandnis haben jugendliche von der natur der naturwissenschaften? Entwicklung und erste schritte zur validierung eines fragebogens [what understanding possess adolescents of nature of science? Development and first steps for validation of a questionnaire]. *Unterrichtswissenschaft*, 36(1), 71.
- Villanueva, M. G., Hand, B., Shelley, M., & Therrien, W. (2019). The conceptualization and Development of the practical epistemology in science survey (PESS). *Research in Science Education*, 49(3), 635–655. <https://doi.org/10.1007/s11165-017-9629-z>
- Wu, H.-K., & Wu, C.-L. (2011). Exploring the development of fifth graders' practical epistemologies and explanation skills in inquiry-based learning classrooms. *Research in Science Education*, 41(3), 319–340. <https://doi.org/10.1007/s11165-010-9167-4>
- Yang, F. Y., Bhagat, K. K., & Cheng, C. H. (2019). Associations of epistemic beliefs in science and scientific reasoning in university students from Taiwan and India. *International Journal of Science Education*, 41(10), 1347–1365. <https://doi.org/10.1080/09500693.2019.1606960>