

German university students' views of nature of science in the introductory phase

David Woitkowski¹,* Leonie Rochell, and Anna B. Bauer¹
 Paderborn University, Physics Education, Department of Physics,
 Warburger Straße 100, 33098 Paderborn, Germany

 (Received 11 August 2020; accepted 25 February 2021; published 15 March 2021)

Learning about the nature of science is an important aspect of becoming part of the culture of practicing physicists. While these topics have a long-standing tradition in U.S. curricula they are currently not part of German educational standards or curricula. And while there is some research concerning primary and secondary school students' beliefs in this area, there are hardly any data about German university students. In this study we query 100 physics students in their first to fifth semester at a German university for their views of nature of science. We then compare them to a sample of physics professors at German universities from an earlier study. The students exhibit quite adequate views in some areas but could improve on their knowledge of the variability of scientific methods and the tentativeness and limited certainty of scientific knowledge—two issues they have in common with the professors. If we interpret the sample longitudinally, we see positive development in the first but worsening in the latter area.

DOI: [10.1103/PhysRevPhysEducRes.17.010118](https://doi.org/10.1103/PhysRevPhysEducRes.17.010118)

I. INTRODUCTION

During their first semesters at German universities, students undergo a major process of socialization and identity building [1], which, over time, transforms a *student* to a *physics student* or merely a *physicist*—someone who has found her or his place in a complex community of practice, of shared values, ideas, questions, and methods—someone who thinks about and solves problems like physicists do [2–5]. While typical university curricula include intended learning outcomes such as content knowledge or experimental methods, one may also identify a “hidden curriculum” with aspects of epistemology, ontology, and discourse that is taught merely implicitly [6]. An important part of physics education research (PER) is to investigate the students' learning of these hidden curriculum as well as the open curriculum, especially the students' attitudes and beliefs [7].

One important part of this hidden curriculum in physics might be the students' views of *nature of science* (NOS), of how science works and what characterizes scientific knowledge. While there is a plenitude of research concerning students at the K–12 level (especially in the U.S.), the university level is relatively under-researched. This is even more so the case in Germany where school curricula more concentrate on physics' contribution to “Bildung” [8]

which traditionally does not entail NOS (as the American tradition does). However, studies suggest that adequate views of the nature of science support the learning of physics content knowledge [9].

Viewing the study introductory phase as a process of enculturation, we see the social processes between students, peers, and professors therein as a major source for the development of undergrad students' views of nature of science. While we investigated professors' views in the past [10], we now ask three cohorts of physics students in their first semesters for their views. We want to trace their development and then investigate similarities and differences to the professors' views.

This publication is part of a broader research program where the whole process of acquisition of views from professors and other instructing personnel, peers, and other sources will be analyzed. From the first step of surveying the professors' views, we here present the results of the second step where the students' views are analyzed.

II. THEORY

When discussing nature of science, we deal with a person's perception of how physics—and, similarly, the sciences as a whole—works, what methods they employ, and how one can characterize the knowledge they generate. Lederman characterizes NOS as “epistemology of science, science as a way of knowing, or the values and beliefs inherent to scientific knowledge or the development of scientific knowledge” [11] (p. 303).

In this section, we will first discuss more detailed categorizations of the various aspects of NOS and then analyze what we already know about the views of high

*david.woitkowski@uni-paderborn.de

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

school students and professors on the matter. Here, our goal is not to judge students' (or any person's) views as more or less correct but to characterize them along known topical lines to gain a structured and comprehensive view.

A. Relevant aspects of NOS

In the context of physics education, we are concerned with more specific aspects of the nature of science: What actual pieces of information about NOS should students learn? From a pragmatic point of view, one could first distinguish information about scientific knowledge from information about the way we as a scientific community come to this knowledge. Hofer and Pintrich call this the *nature of knowledge* and the *nature of knowing* [12]. Schwartz, Lederman, and Lederman use the terms *nature of scientific knowledge* (NOSK) and *nature of scientific investigations* (NOSI) instead [13]. Osborne *et al.* use an additional category about *science* as a social phenomenon and its institutions [14].

From here, there are several ways to further differentiate and generate a catalog of relevant aspects. One way is to analyze existing curricula and educational standards like McComas and Olson did [15]. A different approach was used in the Delphi study by Osborne *et al.* [14]. A third way was used by Lederman *et al.* who mainly dealt with the usefulness of the knowledge in question [11,13,16,17].

Several of these studies resulted in some kind of catalog of statements about the nature of science. Neumann and Kremer give a juxtaposition of several of these catalogs (Table I).

One might also notice that these quite different catalogs often correspond to different test instruments or approaches. For practical reasons we will use the questionnaire from our previous study [10] as discussed below.

TABLE I. Aspect of NOS based on the juxtaposition by Neumann and Kremer [18], employing the works by McComas and Olson [15], Osborne *et al.* [14], Lederman [11] (p. 304), and Schwartz *et al.* [13] (p. 4ff).

	Reference
Scientific knowledge ...	
is tentative but stable	[11,14,15]
is empirically based	[11,13–15]
uses theories and laws	[14]
is replicable and critically tested	[14,15]
Scientific investigations ...	
Aim at knowledge gain	[13–15]
are social phenomena	[11,13–15]
employ multiple methods	[13,14]
recognize and handle anomalous data	[13]
Science ...	
is linked with technology	[14,15]
is culturally embedded	[11,14,15]
has global implications	[15]
must be reported clearly and openly	[15]

However, this instrument does not cover aspects in the third block of Table I.

B. High school students' views of NOS

High school students' views constitute the baseline for any further development at university. We, thus, start with this demographic.

Kampa *et al.* report four groups of high school students with varying degrees of sophistication in their views of nature of science [19]: In each about 40% of their sample exhibit *slightly sophisticated* or *sophisticated* views. This corresponds to moderate or better scores in their test instrument without notable differences between the measured NOS aspects. However about 10% of the sample can be described as having *evidence-based* or *dynamic* views (i.e., they value empirical evidence but show naive views concerning the dimensions source and certainty of scientific knowledge) and additional 10% as having a *multiplistic view* (i.e., they question and challenge scientific knowledge, but do not appreciate the importance of evidence).

In an international study, Neumann [20] compares U.S. and German high school students and finds slightly more sophisticated views among the U.S. sample—this corresponds to the fact that the nature of science does not play a major role in German curricula when compared to the U.S. However, several studies suggest that these views are somewhat domain specific [21]. Quantum mechanics, for instance, seems to be an area that is typically used to explicitly discuss NOS at school. Students thus show comparatively well-informed views in this domain [22].

Several studies show the benefit of adequate views of NOS [9]: Students with high NOS scores may learn physics content more efficiently and show different approaches to understanding hard-to-grasp concepts [23], are better scientific thinkers and problem solvers [24,25], and show differences when evaluating conflicting evidence [26]. For the transition from high school to undergrad studies in physics a positive selection along more expertlike views of NOS and attitudes towards learning physics may be observed [27].

Here, we might notice that, while we hesitate to call a person's views on NOS "right" or "wrong" (especially when it comes to the more professional level [10]) we still might identify views that do enhance or hinder learning as well as views that are more "adequate" when we compare them to results from the epistemology (i.e., philosophy) or history of science [20,28,29]. This is not identical to the views being "expertlike" as, e.g., professors [10] or teachers [30] might be experts with views that are not necessarily adequate when seen from an epistemological or historical point of view.

C. Acquisition during the introductory phase

During the introductory phase, i.e., the first couple of semesters at university, undergrad students get confronted

with a plethora of new ways of thinking and working in the domain of physics (and probably beyond) [1]. Through the lens of a constructivist understanding of learning, all these interactions shape the way students construct their understanding of physics as a science, its methods and knowledge. Central elements, which shape students' learning and understanding, can be identified as (a) their and their peers' prior knowledge or views, (b) their instructors' (professors' or teaching assistants') views and how they enact them, and (c) learning structures (e.g., teaching methods in lectures, instructions, exercises, lab work).

Our previous work [10] shows that German university professors by and large tend to understate the tentative nature of scientific knowledge and to overstate the rigidity of methods for scientific investigations, and thus could be considered some kind of empiricist—a not overly sophisticated view in the categories of Kampa *et al.* [19]. While they stress the importance of teaching students about the nature of science they rarely do so in their own lectures.

However, as the professors play the role of *experts* and possibly also as *role models*, one might expect a trend of students' beliefs towards their professors' beliefs—or at least their enacted beliefs or their “hidden curriculum” [6]. As these aspects are often not made explicit there are multiple ways imaginable that, from here, might lead to an “epistemic deterioration” during university courses that do not explicitly consider NOS [31].

III. RESEARCH QUESTIONS

In this paper we will work on two research questions concerning students' views of nature of science and their development during the first 5 semesters at German universities:

- QI Do developmental patterns emerge in the students' views during their first years at university?
 QII In what way are these views similar or different from the views of German professors [10]?

IV. METHOD

A. Available test instruments

Especially in the U.S., where NOS plays a major role in science education, there is a wide variety of test instruments for all educational levels. Some of the more established ones come from projects referenced in Table I: Views of Nature of Science (VNOS [16]), Views about Science Survey (VASS [5,32]), Maryland Physics Expectations Survey (MPEX [33]), or Colorado Learning Attitudes about Science Survey (CLASS [34]). Some of those also diverge into neighboring areas like students' attitudes towards learning science.

When using a test instrument in another language, it is important to not change the items' meanings during the translations process. We thus need to either revalidate a

translated instrument or concentrate on instrument with an established and validated German translation. We find a German translation of VASS [35] and a German translations of E-CLASS [36] (a variant of CLASS focussing experimental physics, [37]) available. Some further instruments are developed in German by Kremer [21], by Neumann [20] (which is available in both German and English), as well as some scales by Riese, which are part of a German competency test for prospective physics teachers [38].

In our earlier publication [10] we used closed-form items from those instruments to construct a test instrument that is usable at the professional or university level. We chose the form of an online closed-form test to make it easily analyzed. However, this ruled out adapting items from the German VASS [35] as it uses open-form questions which need some amount of interpretative work to be analyzed.

B. Test instrument

The used test instrument queries the test person's beliefs concerning several aspects of NOS. It is largely based on established scales (Table II [39]). The students were to rate 44 given statements on a 4-point Likert scale (1 = “totally incorrect” to 4 = “totally correct”).

All test items are formulated independent of any specific physics content to not skew the results this way.

C. Statistical tests

In the following sections we use some statistical tests to draw our conclusions: For differences between two groups, the Wilcoxon-Mann-Whitney test is calculated (U-test) [41]. This nonparametric test may be interpreted alike the common t test but does not rely on the normality of data, which is not to be expected for small sample sizes. Similarly, for differences between more groups we use the Kruskal-Wallis test as a nonparametric replacement for the more common ANOVA. Both calculations are done with the R package *stats* [42]. For effect sizes we calculate Cohen's d with the R package *lsr* [43].

As we calculate p values to check for statistical significance of differences in a lot of occasions, we need to control the familywise error rate with a Bonferroni correction [44]. In each of the presented statistical tests we regard the results as significant on a level α if $p < \alpha/m$ with $m = 16$ being the number of statistical tests. We use ** for $\alpha = 0.01$ ($p < 0.000625$), * for $\alpha = 0.05$ ($p < 0.003125$) and + for $\alpha = 0.1$ ($p < 0.00625$).

D. Sample

We conducted our study with 100 undergraduate students at the university of Paderborn, Germany, in September and October 2019. The students were recruited in 3 cohorts in their bachelor of science (B.Sc.) or bachelor

TABLE II. Overview of the test instrument. Example items marked (–) are inverted. Scales from Kremer’s tests [40] are marked *K* and from Riese [38] *R*. All test items can be found in the Appendix.

Id	Scale	Items	Example item	Source
NOSK-CRT	Certainty of knowledge	6	Even physical knowledge is not clearly provable and can change over time.	R
NOSK-DEV	Development of knowledge	8	New discoveries can change what scientists think is right.	K
NOSK-SMP	Simplicity of knowledge	5	The more complicated a scientific theory is, the higher its reputation is among scientists. (–)	K
NOSK-JST	Justification of knowledge	9	Good theories rely on the results of many different experiments.	K
NOSI-PRP	Purpose of the sciences	5	The goal of scientific theory is to give order to part of the human experience.	K
NOSI-TAL	Theories and laws	3	Theories are not yet proven, laws are fact. (–)	R
NOSI-MET	Scientific methods	3	In order to gain new physical insights one has to proceed according to the following method: generation of hypothesis—development of appropriate experiments—observation and evaluation—derivation of laws. (–)	R
NOSI-CRE	Creativity and imagination	5	Creative thinking is incompatible with logic-based science. (–)	K

of education (B.Ed.) program. Cohort A just started their studies, cohort B just began their second year at university, and cohort C are at the end of their second year. Anonymity was ensured. The students received neither compensation nor course credit. However, the students got time in their class to fill out the survey. The response rate was 83.3% (100 complete responses out of 120 responses in total).

A sample overview is given in Table III. About one-fourth to one-fifth of the students were female, five did not specify their gender. About three-fourths are enrolled in a bachelor of science program (B.Sc.) as physics majors, one-fourth in a bachelor of education program (B.Ed.) as

prospective teachers. In general, the participation counts are in concordance with the typical university dropout rates in physics at German universities [45].

For comparison, we draw on the sample of 50 German university professors which were tested in a similar fashion in May and June 2018 [10].

As a (weak) check for a positive selection, the test persons were asked for their perceived background in NOS. We used a single item where they could rate their own experience in comparison to their respective peers on a scale from 1 = “much less than average” to 5 = “much more than average”. While the professors rate their background as slightly higher

TABLE III. Sample overview. For each cohort, the full sample, and for the professors’ sample from Ref. [10], we give the person count N , the percentage of persons giving “female” or “male” as their gender (participants could leave the question blank), the study program (numbers missing to 100% are due to missing answers), and the perceived background in NOS (1 = “much less than average” to 5 = “much more than average” compared to their respective peers).

Cohort		N	Gender		Study program		Background	
			Female	Male	B.Sc.	B.Ed.	Mean	SD
A	(1st sem.)	52	23.1%	76.0%	72.1%	26.9%	3.12	0.83
B	(3rd sem.)	27	25.9%	73.1%	74.1%	25.9%	2.89	1.05
C	(4th or 5th sem.)	21	19.0%	78.9%	90.5%	4.8%	3.05	0.86
All Students		100	23.0%	72.0%	77.0%	22.0%	3.04	0.90
Professors		50	3.50	0.94

than average ($M = 3.50$, $SD = 0.94$; average relative experience would be 3.0), the students could not be identified as a positive selection in this regard ($M = 3.04$, $SD = 0.89$). The difference between professors and students is significant ($W = 1605$, $p = 0.0018$) with a median effect size of $d = 0.51$. To check if the three cohorts of students differ in this respect, we conducted an Kruskal-Wallis test with the cohort as predictor. This did not indicate significant differences between the cohorts ($\chi^2 = 1.145$, $df = 2$, $p = 0.56$).

In a Kruskal-Wallis test, the three cohorts do not differ in the amount of female participants ($\chi^2 = 0.206$, $df = 2$, $p = 0.9$) nor in their study program ($\chi^2 = 4.268$, $df = 2$, $p = 0.11$). However, it should be noted that in cohort C only 1 of 21 students (4.8%) is enrolled as a student teacher compared to about one-fourth in the other cohorts.

E. Validity concerns

To establish validity, one has to present an argument why a test score might be interpreted as a measure of the trait it is intended to represent [46].

As our test instrument is based on already validated scales from other instruments, we refer the interested reader to the qualitative analysis done in the original publications. One concern, however, might be the different target population samples. As the source instruments are constructed for high school students we do not expect a bigger loss of validity when using it with students in their early semesters at university as they might show quite a similar understanding. To (at least partly) support this claim, we will analyze the test's internal consistency as an indicator of students' understanding of test items. If students' understanding of single items differs significantly, one might expect a considerably lower item-scale correlation.

However, as university students enrolled in physics programs may be seen as a positive selection when compared to high school students, we might expect a ceiling effect as our sample may, as a group, tend to more adequate views [27].

V. RESULTS

A. Scale analysis

We first check for scale reliability to ensure general usability with this sample. This is of special importance for the following two reasons: (i) Although the used scales were developed and tested with (high school) students as presented in the original publications, students from our sample might interpret some items differently than intended. This would lead to the affected scale having lower internal consistency and, thus, bad reliability measures. (ii) One might argue that the students' beliefs are not more or less stable cognitive traits but merely decisions made in the moment. This again would lead to (maybe only slightly) lower reliabilities.

As a reliability measure we use the standardized Cronbach's α [47]. Results are given in Table IV. We regard scales with $\alpha < 0.65$ as not reliable enough to be used. We put the cutoff at this comparatively low value (instead of the widely used 0.7) as several scales exhibit an α just under 0.7, which would reduce our analysis more than necessary.

The somewhat low reliabilities just below the usual cutoff of about $\alpha < 0.7$ might be attributed to the early stage in the students' science learning. At this stage a lot of their views might not have formed fully, resulting in some inconsistent answering what might correspond to reason (ii) above. This might be identified as a variant of the typical pretest effect [48].

With this cutoff, we need to drop the scale NOSI-TAL (theories and laws) with α as low as 0.28. Four other scales each had one item with low item-total correlation. After dropping those items, they reached a reasonable α . The specific items dropped are indicated in the test instrument in the Appendix. Based on this analysis, the scales are reliably usable with our student sample.

For all remaining scales we calculate scores as item means with items rated from 1 = "totally incorrect" to 4 = "totally correct". Some items are inverted as indicated in the Appendix, to let higher values represent more adequate views.

B. Students' views

We are now prepared to have a look at the students' views on the remaining 4 NOSK and 3 NOSI scales. Figure 1 gives a first overview of the data.

The general tendency to observe is that all scale means are above the theoretical midpoint of 2.5. This should indicate more adequate views, in general, and could correspond to the positive selection compared with the originally intended audiences as discussed above. However, there are further nuances: The scales NOSK-CRT (certainty of knowledge) and NOSI-MET (scientific

TABLE IV. Scale overview for the whole student sample. For each scale the item count and standardized Cronbach's α is given. Item count is given indicating dropped items as well. For the retained scales arithmetic mean and standard deviation (SD) are given as well.

Scale	Item count	Reliability (std. α)	Mean	SD
NOSK-CRT	5	0.68	2.81	0.60
NOSK-DEV	8	0.87	3.63	0.44
NOSK-SMP	5	0.69	3.15	0.54
NOSK-JST	8	0.78	3.64	0.36
NOSI-PRP	4	0.65	3.38	0.43
NOSI-TAL	3	(0.28)	scale dropped	
NOSI-MET	2	0.68	2.82	0.73
NOSI-CRE	5	0.79	3.27	0.56

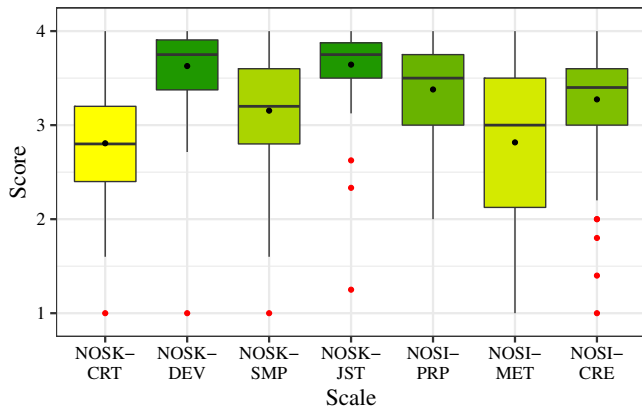


FIG. 1. Adequateness of students' beliefs. Higher scores indicate more appropriate beliefs in the respective areas. Horizontal bars indicate scale medians, black dots indicate arithmetic means, red dots represent outliers. Raw numbers can be found in Table IV.

methods) exhibit relatively low means of only about 2.8. In contrast, the scales NOSK-DEV (development of knowledge) and NOSK-JST (justification of knowledge) show rather high means of above 3.6.

The diagram also indicates some outliers at the lower end of most scales. All together, 19 students score below 2 on at least one scale. Of those, only two test persons have a low score on 2 scales and only one person on 5 scales at once. 9 low scores are found on the scale NOSI-MET.

The overall impression we might get from these findings could be that our students are well aware of the fact that scientific knowledge develops and that experiments play a role in this process, but they often lack a proper understanding of the interplay of theory and experiment and the degree of certainty to be expected from scientific findings.

We will now divide the student sample into the three cohorts: A (1st semester), B (3rd semester), and C (4th and 5th semester) as in Table III. Although these are distinct persons, we might interpret this longitudinally and infer developmental patterns. Here, some caution is needed as the three cohorts' composition is not completely equal: Cohort A is of much bigger size (because of dropout), cohort C consists of much fewer student teachers than the earlier ones. As this design is not able to differentiate between-person effects from within-person effects, we have to check all effects for group differences.

The plot in Fig. 2 gives the same data as Fig. 1 by split by cohorts. For most scales, the three cohorts differ only slightly. If we conduct a Kruskal-Wallis test to test for the cohorts' effect on each scale we only get significant differences for one scale.

There is a marginally significant effect on the scale NOSK-CRT (certainty of knowledge; $\chi^2 = 11.161$, $df = 2$, $p = 0.005$). Here we observe the students' score to be lower in later semesters. However, in a pairwise U-test, only the difference between cohorts A ($M = 2.97$,

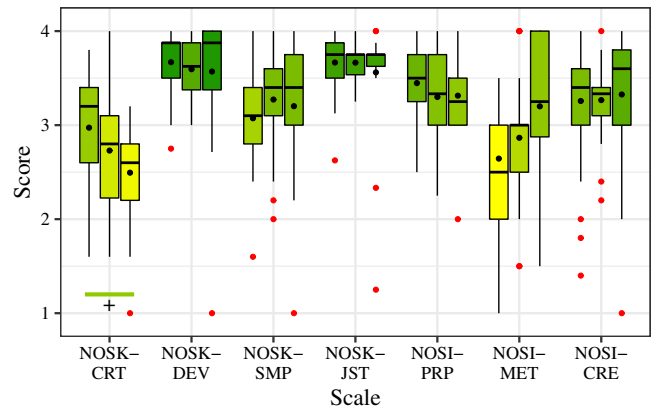


FIG. 2. Adequateness of students' beliefs by cohorts (A, B, C from left to right). Higher scores indicate more appropriate beliefs in the respective areas. Horizontal bars indicate scale medians, black dots indicate arithmetic means, red dots represent outliers.

$SD = 0.56$) and C ($M = 2.49$, $SD = 0.57$) is statistically significant ($W = 809$, $p = 0.0013$) with a big effect (Cohen's $d = 0.86$). This effect might be explained by cohort composition as the later cohorts consist of fewer student teachers. We check this by calculating the effect of the study program on this scale. We do not find a significant effect ($W = 933$, $p = 0.46$) which supports the notion of a cohort effect.

A trend into the opposite direction in the scale NOSI-MET (scientific methods) is not significant after the Bonferroni correction ($\chi^2 = 7.958$, $df = 2$, $p = 0.019$). Here, the students' scores appear higher in later cohorts.

When combined, these findings are quite interesting: Students score rather highly (indicating adequate beliefs) on most scales. However, while the students seem to get a better grasp of physics' methodical details and of the interplay of theories and laws over time, they also regard physical knowledge as increasingly certain and less tentative.

C. Comparison to professors' views

In our previous study we gathered the data of 50 physics professors at German universities [10]. Although the tested professors are not the persons responsible for the university education of any student from our sample, this is a representative sample from the whole of Germany.

As with the professors more scales were dropped out of the study, we are left with only 5 scales to directly compare between the two samples. The results are shown in Fig. 3. Here, we observe quite a few similarities between the professors' and students' views. A U-test only exhibits significant difference in one of the five scales:

On the scale NOSI-CRE (creativity and imagination), the professors score very high ($M = 3.5$, $SD = 0.46$). The students score significantly lower ($M = 3.27$, $SD = 0.56$; $W = 1629$, $p < 0.0005$) with a medium sized effect

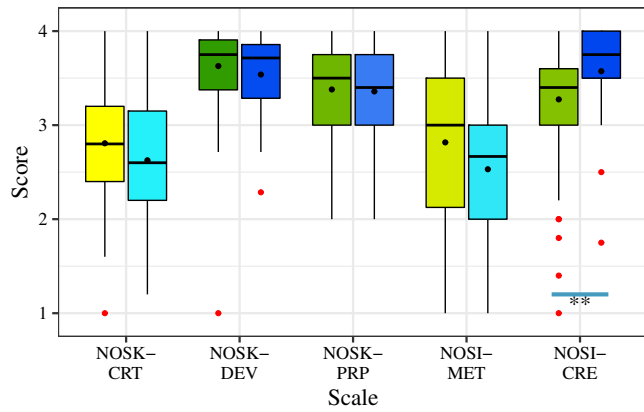


FIG. 3. Adequateness of students' (left) and professors' (right) beliefs. Higher scores indicate more appropriate beliefs in the respective areas. Horizontal bars indicate scale medians, black dots indicate arithmetic means, red dots represent outliers. Significant differences between groups in U-test are marked.

($d = 0.56$). Here again, we do not observe the theoretically expected trend towards the professors' views. A plausible explanation might be that classes in experimental physics at university largely confront the students with kind-of ready-made experimental questions. Conversely, the professors might draw their views from their research, where the creative act of finding interesting questions or interpreting data in new ways might be more important.

The slight difference in the scale NOSI-MET (scientific methods) between professors ($M = 2.53$, $SD = 0.77$) and students ($M = 2.82$, $SD = 0.73$) is not significant after Bonferroni correction ($W = 2903.5$, $p = 0.036$).

VI. IMPLICATIONS FOR LEARNING NOS AT UNIVERSITY

While learning about the nature of science has a long-standing tradition in U.S. curricula, it is a relatively new topic in German physics education research and is rarely represented in German curricula. One might thus argue that NOS is only learned implicitly but only rarely made an explicit topic in class—both at school and at university.

It, thus, can be viewed positively that our students exhibit relative adequate views in most aspects of NOS. In all areas we see mean scores above the scale mean. This is even the case in the area of scientific methods where the professors' views in our earlier study tended to be lower than expected [10].

However, one worrying result is the tendency to view scientific knowledge as certain and largely stable. This effect can be observed with both students and professors and with the former even worsening over time. Here, we might suggest several approaches to better that situation.

One possible consideration might be to more explicitly deal with uncertainty in experimental data [49]. Research suggests that most teachers lack the competence to

adequately handle uncertain evidence—probably they were never exposed to them in their studies. This corresponds to our finding of perceived high certainty of knowledge.

Another approach could strategically use the history of science to teach about NOS as well [50]. As the development of physical knowledge might be suitable to demonstrate its tentativeness and the discussions of what is established fact, this could lead to more adequate views in the long term. Interesting, albeit laborious, ideas for actual lab work come from experimental science history where historical experiments are reconstructed and discussed in regards to their development, explanatory power and impact on the scientific community [51,52].

In any case, our study is not the first one to find less adequate belief patterns among students that might hinder their learning [9,23–26]. Other studies suggest that the negative effects of latter views can be circumvented with specifically targeted instruction [53]. It would be worthwhile to generate similar concepts for university courses that could specifically target NOS.

Looking ahead, the next step in this research project should be to uncover the specific processes that lead to the students' acquisition of views on NOS as well as their views' effects on their concrete work in physics research. Additionally, we then might shed some light on the importance of views in the professional stage of a career in physics [54].

Another topic to research might be the perceived importance of topics like NOS in the view of our students. The professors in our earlier study perceived most topics as rather important—with the exception of the certainty and development of knowledge and the role of creativity and nonscientists in physics research. As of now we are not able to assess possible similarities or assimilation or learning effects with a similar instrument.

VII. SUMMARY AND DISCUSSION

For this study we did a complete survey with 100 physics students in the first 2 years at a German university for their beliefs on several aspects of the nature of science. The used online instrument was adapted from an earlier study where it was used with 50 physics professors at German universities. It, thus, can be regarded as established for the use in an university context. The overall reliability turned out just acceptable. This might be a common pretest effect where students did not form coherent beliefs on some of the topics yet [48]. However, the tests validity may be questioned as the validation studies carried out for the adapted test scales involved no interview or think-aloud studies that might ensure the test persons understand the items as intended.

The students show rather adequate views on most areas tested. Notable exceptions are the certainty of scientific knowledge where they tend to a more certain view than the academic consensus would suggest. They also regard scientific methods as relatively rigid. However, these views

resemble the professors', who tend to even less adequate views in some of these areas. In comparison with the professors, the students also regard scientific research as less creative. If we characterized the professors as adhering "to some kind of naive empiricism" (p. 9) in our earlier study [10] the same can be said about the students. If we use Kampa's categories [19], these students exhibit evidence-based or dynamic views (similar to about 10% of the tested high school students). This suggests the positive selection effect found in literature [27].

The structure of our student sample gave us the opportunity to differentiate three cohorts: A just started their studies, B at the beginning of their second year at university, and C at the end of their 2nd year at university. If we interpret these longitudinally, we only observe little development: More advanced students regard scientific knowledge as slightly more certain and demonstrable by experiment. This finding only partly corresponds to what Johnson and Willoughby describe as epistemic deterioration [31].

From this point on several possible next steps could be imagined: With a bigger sample from a wider range of backgrounds, we could try to find groups with distinct belief patterns or differences in their development over time. This would enable us to evaluate the effects different kinds of university courses have on the students' views.

It would also be worthwhile to conduct longer interviews with at least some of the students. Here, determining factors for their beliefs could be researched. With techniques like story telling, one could paint a picture of the genesis of their respective views and give explanations for the found patterns.

APPENDIX: QUESTIONNAIRE

Please note that the presented test items are translated solely for English language publication. In the study they were used in their German original. For the sources of the scales see Table II.

Items that were inverted for analysis are marked (–), dropped items (due to low discriminatory power) are marked (X).

1. NOSK-CRT: Certainty of knowledge

- Physics, like humanities, cannot provide absolute true knowledge.
- In the sciences, valid evidence is relevant, therefore, there is no discussion about what is considered to be secure knowledge. (–)
- Physical knowledge can be clearly demonstrated (e.g., by appropriate experiments) (–)
- (X) Scientific findings are not purely objective, but also are influenced by the bias of the researchers.
- Even physical knowledge is not clearly provable and can change over time.

- Physics distinguishes itself from other sciences (e.g., legal studies) in that one can clearly decide without long discussions whether a theory is right or wrong. (–)

2. NOSK-DEV: Development of knowledge

- Scientific theories are changed or replaced when new evidence is available.
- Sometimes concepts change in the sciences.
- Sometimes scientists change their mind about what's true in their field.
- New discoveries can change what scientists think is right.
- There are many questions in the sciences that even scientists cannot answer.
- Some concepts in the sciences are different today than what scientists used to think.
- The concepts in science books sometimes change.
- Scientific theories change and evolve over time.

3. NOSK-SMP: Simplicity of knowledge

- Scientific theories are often more complicated than they should be. (–)
- Scientific theories and laws are more complicatedly formulated than simply. (–)
- The more complicated a scientific theory is, the higher its reputation is among scientists. (–)
- Scientists strive to establish as many theories and laws as possible. (–)
- If two theories equally explain a natural phenomenon, the more complicated theory is the better one. (–)

4. NOSK-JST: Justification of knowledge

- Good theories rely on the results of many different experiments.
- When scientists conduct experiments, they determine in advance some aspects of the exploration.
- (X) It is important to have a concrete idea before starting an experiment.
- For scientists, experiments with unexpected results are worthless. (–)
- It is important to do experiments more than once to ensure results.
- The ideas for science experiments come from being curious and thinking about how something works.
- In the sciences, new concepts can emerge from one's own questions and experiments.
- There can be several ways in science to verify concepts.
- An experiment is a good way to find out if something is true.

5. NOSI-PRP: Purpose of the sciences

- The goal of scientific theory is to give order to part of the human experience.

- (X) Scientists conduct experiments to make new discoveries.
- The goal of scientific theories is to explain natural processes.
- Scientists study natural phenomena and explain why they occur.
- Scientist conduct experiments to explain how certain events come about.

6. NOSI-TAL: Theories and laws (scale dropped)

- Physical theories are true representations of reality. (–)
- A theory is the preliminary stage of a law. (–)
- Theories are not yet proven, laws are fact. (–)

7. NOSI-MET: Scientific method

- Without the results and data from appropriate experiments, no new physical theories can be established. (–)

- (X) In order to gain new physical insights one has to proceed according to the following method: generation of hypothesis—development of appropriate experiments—observation and evaluation—derivation of laws. (–)
- New theories are always developed from the results of experiments. (–)

8. NOSI-CRE: Creativity and imagination

- Science theories and laws have nothing to do with creativity. (–)
- Scientific knowledge is also a result of human creativity.
- Creative thinking is incompatible with logic-based science. (–)
- The scientific knowledge shows the creativity of scientists.
- The creative thinking of scientists is too untrustworthy to achieve scientific advances. (–)

-
- [1] H. T. Holmegaard, L. M. Madsen, and L. Ulriksen, A journey of negotiation and belonging: Understanding students' transitions to science and engineering in higher education, *Cultural Studies Sci. Educ.* **9**, 755 (2014).
- [2] N. J. Nersessian, Should physicists preach what they practice? Constructive Modeling in Doing and Learning Physics, *Sci. Educ.* **4**, 203 (1995).
- [3] E. Etkina, A. van Heuvelen, S. White-Brahmia, D. T. Brookes, M. Gentile, S. Murthy, D. Rosengrant, and A. Warren, Scientific abilities and their assessment, *Phys. Rev. ST Phys. Educ. Res.* **2**, 020103 (2006).
- [4] D. Hammer, Students' beliefs about conceptual knowledge in introductory physics, *Int. J. Sci. Educ.* **16**, 385 (1994).
- [5] I. Halloun and D. Hestenes, Interpreting VASS dimensions and profiles for physics students, *Sci. Educ.* **7**, 553 (1998).
- [6] E. F. Redish, Introducing students to the culture of physics: Explicating elements of the hidden curriculum, [arXiv: 1008.0578](https://arxiv.org/abs/1008.0578).
- [7] L. C. McDermott and E. F. Redish, Resource letter: PER-1: Physics education research, *Am. J. Phys.* **67**, 755 (1999).
- [8] *Teaching as a Reflective Practice: The German Didaktik Tradition*, edited by I. Westbury, *Studies in Curriculum Theory* (Erlbaum, Mahwah, NJ, 2000).
- [9] F.-Y. Yang and C.-C. Tsai, Personal epistemology and science learning: a review on empirical studies, in *Second International Handbook of Science Education*, edited by B. J. Fraser, K. G. Tobin, and C. J. McRobbie (Springer, Dordrecht, 2012), pp. 259–280.
- [10] D. Woitkowski and N. L. Wurmbach, Assessing German professors' views of nature of science, *Phys. Rev. Phys. Educ. Res.* **15**, 010108 (2019).
- [11] N. G. Lederman, Syntax of nature of science within inquiry and science instruction, in *Scientific Inquiry and Nature of Science: Implications for Teaching, Learning and Teacher Education*, edited by L. B. Flick and N. G. Lederman (Springer, Dordrecht, 2006), pp. 301–317.
- [12] B. K. Hofer and P. R. Pintrich, The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning, *Rev. Educ. Res.* **67**, 88 (1997).
- [13] R. S. Schwartz, N. G. Lederman, and J. Lederman, An instrument to assess views of scientific inquiry: The VOSI questionnaire, in *Proceedings of the National Association for Research in Science Teaching, Baltimore, MD* (National Association for Research in Science Teaching, Reston, VA, 2008).
- [14] J. Osborne, S. Collins, M. Ratcliffe, R. Millar, and R. Duschl, What “Ideas about Science” should be taught in school science? A Delphi study of the expert community, *J. Res. Sci. Teach.* **40**, 692 (2003).
- [15] W. F. McComas and J. K. Olson, The nature of science in international science education standard documents, in *The Nature of Science in Science Education: Rationales and Strategies*, edited by W. F. McComas (Kluwer, Dordrecht, 1998), pp. 41–52.
- [16] N. G. Lederman, F. Abd-El-Khalick, R. L. Bell, and R. S. Schwartz, Views of Nature of Science questionnaire: Toward valid and meaningful assessment of learners' conceptions of nature of science, *J. Res. Sci. Teach.* **39**, 497 (2002).
- [17] N. G. Lederman, Nature of science: Past, present, and future, in *Handbook of Research on Science Education*, edited by S. Abell and N. G. Lederman (Lawrence Erlbaum, Mahwah, NJ, 2007), pp. 831–879.
- [18] I. Neumann and K. Kremer, Nature of Science und epistemologische Überzeugungen—Ähnlichkeiten und

- Unterschiede, *Z. Did. Naturwissenschaften* **19**, 211 (2013), <https://archiv.ipn.uni-kiel.de/zfdn/jg19.html#Art004>.
- [19] N. Kampa, I. Neumann, P. Heitmann, and K. Kremer, Epistemological beliefs in science—a person-centered approach to investigate high school students’ profiles, *Contemp. Educ. Psychol.* **46**, 81 (2016).
- [20] I. Neumann, Beyond physics content knowledge: Modeling competence regarding nature of scientific inquiry and nature of scientific knowledge: Dissertation, *Studien zum Physik- und Chemielernen*, Vol. **117** (Logos, Berlin, 2011).
- [21] D. Urhahne, K. Kremer, and J. Mayer, Conceptions of the nature of science—are they general or context specific?, *Int. J. Sci. Math. Educ.* **9**, 707 (2011).
- [22] H. K. E. Stadermann and M. J. Goedhart, Secondary school students’ views of nature of science in quantum physics, *Int. J. Sci. Educ.* **42**, 997 (2020).
- [23] H. Michel and I. Neumann, Nature of science and science content learning, *Sci. Educ.* **25**, 951 (2016).
- [24] W. A. Sandoval, Understanding students’ practical epistemologies and their influence on learning through inquiry, *Sci. Educ.* **89**, 634 (2005).
- [25] H.-S. Lin and H.-L. Chiu, Student understanding of the nature of science and their problem-solving strategies, *Int. J. Sci. Educ.* **26**, 101 (2004).
- [26] T. D. Sadler, F. W. Chambers, and D. L. Zeidler, Student conceptualizations of the nature of science in response to a socioscientific issue, *Int. J. Sci. Educ.* **26**, 387 (2004).
- [27] S. P. Bates, R. K. Galloway, C. Loptson, and K. A. Slaughter, How attitudes and beliefs about physics change from high school to faculty, *Phys. Rev. ST Phys. Educ. Res.* **7**, 020114 (2011).
- [28] P. Heering and E. Cavicchi, Teaching about nature of science through historical experiments, in *Nature of Science in Science Instruction, Science: Philosophy, History and Education*, edited by W. F. McComas (Springer, Cham, Switzerland, 2020), pp. 609–626.
- [29] A. Henke and D. Höttecke, Physics teachers’ challenges in using history and philosophy of science in teaching, *Sci. Educ.* **24**, 349 (2015).
- [30] F. Abd-El-Khalick and N. G. Lederman, Improving science teachers’ conceptions of nature of science: A critical review of the literature, *Int. J. Sci. Educ.* **22**, 665 (2000).
- [31] K. Johnson and S. Willoughby, Changing epistemological beliefs with nature of science implementations, *Phys. Rev. Phys. Educ. Res.* **14**, 010110 (2018).
- [32] I. Halloun, Views About Science and physics achievement: The VASS story, in *The Changing Role of Physics Departments in Modern Universities*, edited by E. F. Redish and J. Rigden (American Institute of Physics, College Park, MD, 1997), pp. 605–614.
- [33] E. F. Redish, J. M. Saul, and R. N. Steinberg, Student expectations in introductory physics, *Am. J. Phys.* **66**, 212 (1998).
- [34] W. K. Adams, D. N. Perkins, N. S. Podolefsky, M. Dubson, N. D. Finkelstein, and C. E. Wieman, New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey, *Phys. Rev. ST Phys. Educ. Res.* **2**, 010101 (2006).
- [35] B. Priemer, Ein diagnostischer Test zu Schüleransichten über Physik und Lernen von Physik—eine deutsche Version des Tests “Views About Science Survey”, *Z. Did. Naturwissenschaften* **9**, 160 (2003), <https://archiv.ipn.uni-kiel.de/zfdn/jg9a.htm#page160->.
- [36] The translation GE-CLASS is currently under validation by Micol Alemani at the University of Potsdam, Germany. To our knowledge, neither items nor results are published yet.
- [37] B. R. Wilcox and H. J. Lewandowski, Students’ epistemologies about experimental physics: Validating the Colorado Learning Attitudes about Science Survey for experimental physics, *Phys. Rev. Phys. Educ. Res.* **12**, 010123 (2016).
- [38] J. Riese, Professionelles Wissen und professionelle Handlungskompetenz von (angehenden) Physiklehrkräften: Dissertation, *Studien zum Physik- und Chemielernen*, Vol. **97** (Logos, Berlin, 2009).
- [39] The scale NOSK-SRC (Source of knowledge) from our previous study was not used with the students, because the professors indicated that learning about items from that scale is—in their eyes—far less important than the others.
- [40] K. H. Kremer, Die Natur der Naturwissenschaften verstehen: Untersuchungen zur Struktur und Entwicklung von Kompetenzen in der Sekundarstufe I, Ph.D. thesis, Universitätsbibliothek Kassel, Kassel (2010).
- [41] M. Hollander and D. A. Wolfe, *Nonparametric statistical methods*, Wiley Series in Probability and Mathematical Statistics: Applied probability and statistics (Wiley, New York, 1973).
- [42] R Core Team, R: A Language and Environment for Statistical Computing (2020).
- [43] D. Navarro, Learning statistics with R: A tutorial for psychology students and other beginners. (Version 0.5) (2015).
- [44] O. J. Dunn, Multiple comparisons among means, *J. Am. Stat. Assoc.* **56**, 52 (1961).
- [45] M. Neugebauer, U. Heublein, and A. Daniel, Studienabbruch in Deutschland: Ausmaß, Ursachen, Folgen, Präventionsmöglichkeiten, *Z. Erziehungswissenschaft* **22**, 1025 (2019).
- [46] M. T. Kane, Validating the interpretations and uses of test scores, *J. Educ. Measure.* **50**, 1 (2013).
- [47] L. J. Cronbach, Coefficient alpha and the internal structure of tests, *Psychometrika* **16**, 297 (1951).
- [48] D. W. Zimmermann and R. H. Williams, Reliability of gain scores under realistic assumptions about properties of pre-test and post-test scores, *Brit. J. Math. Stat. Psychol.* **51**, 343 (1998).
- [49] J. Ruhrig and D. Höttecke, Components of science teachers’ professional competence and their orientational frameworks when dealing with uncertain evidence in science teaching, *Int. J. Sci. Math. Educ.* **13**, 447 (2015).
- [50] A. Henke and D. Höttecke, Physics teachers’ challenges in using history and philosophy of science in teaching, *Sci. Educ.* **24**, 349 (2015).
- [51] C. C. Silva and P. Heering, Re-examining the early history of the Leiden jar: Stabilization and variation in transforming a phenomenon into a fact, *Hist. Sci.* **56**, 314 (2018).
- [52] P. Heering, J. Keck, and G. A. Rohlf, Laboratory notes, laboratory experiences, and conceptual analysis: Under-

- standing the making of Ohm's first law in electricity, *Berichte zur Wissenschaftsgeschichte* **43**, 7 (2020).
- [53] L. B. Wheeler, B. K. Mulvey, J. L. Maeng, M. R. Librea-Carden, and R. L. Bell, Teaching the teacher: Exploring STEM graduate students' nature of science conceptions in a teaching methods course, *Int. J. Sci. Educ.* **41**, 1905 (2019).
- [54] K. Le Roux, D. L. Taylor, B. Kloot, and S. Allie, Research on higher education: A perspective on the relations between Higher Education Studies and Discipline-Based Education Research, *Teach. Higher Educ.* **78**, 1 (2019).