Cross-age study on secondary school students' views of stars

Philipp Bitzenbauer,* Sarah Navarrete, and Fabian Hennig
Friedrich-Alexander-Universität Erlangen-Nürnberg, Professur für Didaktik der Physik, Staudtstr. 7, 91058 Erlangen, Germany

Malte S. Ubben
Technische Universität Braunschweig, Institut für Fachdidaktik der Naturwissenschaften, Bienroder Weg 82, 38106 Braunschweig, Germany

Joaquin M. Veith
Stiftungsuniversität Hildesheim, Institut für Mathematik und Angewandte Informatik, Samelsonplatz 1, 31141 Hildesheim, Germany

(Received 23 June 2023; accepted 16 November 2023; published 7 December 2023)

Research in astronomy education has uncovered that many learners possess limited and fragmented understanding of stars. The corresponding misconceptions manifest in various areas such as star formation, size, the relationship between stars and planets, and their position in space and have been shown to persist across different age groups and educational settings, highlighting the need for further investigation. This paper presents the findings of an empirical study that examines secondary students' views of stars and their evolution throughout their secondary school education. We designed and evaluated an instrument for assessing students' views of stars in five domains (stars and the solar system, formation and evolution of stars, general properties and motion of stars, (sub-)stellar objects, as well as color and brightness). The instrument creation process involved several steps, including literature-based item development, an expert survey with faculty members, and a quantitative pilot study with a sample of $N = 390$ secondary school and college students. This process led to a final version of the instrument that exhibits good psychometric properties. We used this new instrument in a cross-age study to investigate the alignment of secondary students' ideas about stars with scientific views across different stages of secondary education. The sample of this main study comprised a total of $N = 366$ learners, including 148 lower (aged 13–14 years), 151 middle (aged 15–16 years), and 67 upper (aged 17–18 years) secondary school students. Our study findings reveal a progressive development of students' perspectives on star-related topics throughout their school education: Using analyses of variance and conducting pairwise *post hoc* comparisons, we observed a statistically significant increase in the proportion of responses aligning with scientific views across all aspects of stars examined in this study, as students progressed from lower secondary to upper secondary levels. We further report on widely held views of stars among our study participants that do not align with the scientific views and discuss the implications of our findings for both educational research and practice.

DOI: 10.1103/PhysRevPhysEducRes.19.020165

I. INTRODUCTION

Astrophysics research findings are increasingly making their way into the spotlight of news broadcasts and newspaper articles due to recent advancements in the field (e.g., see Ref. [1]), as indicated by the multiple Nobel Prizes awarded for astrophysical research in recent years (e.g., 2002 [2], 2006 [3], 2011 [4], 2015 [5], 2017 [6], 2019 [7], and 2020 [8]). The presence of astrophysical topics in the media even permeates the educational sphere, where space and astronomy topics are found to be of significant interest to both boys and girls [9], and this is not surprising as argued in Ref. [10]: From an early age, students are consciously confronted with astronomical questions through the inevitable act of observing the sky [11]. Furthermore, their curiosity about the origins of humanity leads them to seek scientific explanations, highlighting the enduring allure and importance of astronomy and astrophysics in the eyes of students and nonspecialists alike.

*philipp.bitzenbauer@fau.de

Consequently, astronomy concepts are on the rise in K-12 physics education [12,13]. Thereby, the topic of stars and their evolution holds significant potential for physics education from various perspectives: Stars are objects whose structure and life can be described through the interaction of different subdisciplines of physics (for an overview, see Ref. [14]). Through the application of fundamental physics principles, a comprehensive understanding can be gained on the formation and evolution of stars as well as their transition into compact objects such as white dwarfs, neutron stars, or black holes, which occurs when nuclear processes convert lighter elements (e.g., hydrogen and helium) into heavier elements up do iron (including carbon and oxygen) cease. The example of stars offers a compelling demonstration of the interconnectedness between observation (such as spectral analysis) and theoretical description (for an overview, see Ref. [15]). Hence, stars are not only a fundamental aspect of astrophysics but may also serve as a central theme in astronomy education: They provide an excellent opportunity for students to grasp core aspects of stellar evolutionary processes. Additionally, the theory of star formation remains incomplete, with unresolved questions, offering a platform for learning about the nature of science. For example, the mechanisms leading to the collapse of interstellar clouds, and hence to star formation, are still not fully understood [16].

In this paper, we report on a cross-age study investigating secondary students' views of stars and their evolution throughout secondary education. Building on previous research on student learning of astronomy concepts in general, and stars in particular (see Sec. II), we formulate our research questions in Sec. III and describe the design and evaluation of the research tool used in this study in Sec. IV. We present the findings of our study in Sec. VI and discuss these results against the backdrop of prior research in Sec. VII. Finally, we provide recommendations for both future research and practice in astronomy education at the secondary school level (see Sec. IX).

II. RESEARCH BACKGROUND

A. The big ideas of astronomy education research

1. Topics covered and target groups addressed

Lelliott and Rollnick [17] conducted a comprehensive review of peer-reviewed astronomy education studies published between 1974 and 2008. Their review highlighted that the majority of these studies focused on teaching and learning of five big ideas, "all involving the Earth in relation to its satellite and the Sun" [17] (p. 1777): (i) Earth [18–25], (ii) gravity [18,20,21,24–26], (iii) the day-and-night cycle [27–29], (iv) the seasons [30–34], and (v) the Earth-Sun-Moon system [22,35,36]. However, there is a scarcity of studies exploring student learning in other areas of astronomy, such as the Big Bang [12,37–39], black holes [40], stars [41,42], and aspects related to sizes and distances of astronomical objects [43]. A recently published review article by Salimpour and Fitzgerald [44] further highlights the development of "a fertile landscape for research into Cosmology Education" (p. 819). It is noteworthy that research in the field of astronomy education sketched above has focused on questions regarding the teaching and learning of astronomy concepts in various target groups: These groups range from students, including school-aged students [29,30,43,45,46], as well as college and university students [42,47–50], to pre- and in-service teachers [51–54].

2. Tools to assess students' conceptions of astronomy topics

To assess students' conceptual understanding and students' conceptions of the aforementioned astronomy topics, a number of concept inventories have been developed and evaluated, e.g., the Lunar Phases Concept Inventory [55], the Moon Phases Concept Inventory [56], the Light and Spectroscopy Concept Inventory [57,58], the Astronomy Diagnostic Test [59–61], the Test of Astronomy Standards [62], the Astronomy and Space Science Concept Inventory [63], and the Star Properties Concept Inventory [64]. However, in Ref. [65], the authors emphasize the necessity of developing further instruments that allow for the valid assessment of students' conceptions of more advanced topics such as black holes [40] or the Big Bang [12] on the one hand, and stars, especially with regard to aspects beyond their basic properties, on the other. Only one such tool has been developed so far, namely the Cosmology Concept Inventory [66]. And, given the broad range of facets exhibited by cosmological topics, it seems necessary to develop further diagnostic tools to address the gap in the existing literature concerning students' understanding and students' ideas related to the topic of stars. In the upcoming Sec. II B, we will provide an overview of astronomy education research results on this topic.

B. Students' conceptions of stars

Agan's study [67] shed light on students' ideas about stars across various educational levels, revealing a multitude of little elaborate ideas: These encompass the twinkling nature of stars or the students' idea that Polaris (the North Star) would be the brightest star. Notably, a significant number of students fail to recognize the Sun as a star, harbor the notion that stars are immortal, and mistakenly assume that all stars end in supernovae. Moreover, an erroneous view persists among students that all stars in the celestial sphere are equidistant from the Earth [68,69]. Further research has unveiled that students often describe stars as round objects without edges [68] and perceive them as motionless entities in the night sky while recognizing the Sun's apparent motion during the day [28,70]. Their comprehension of daily celestial motion and knowledge of significant celestial objects, such as Polaris and the ecliptic, also display limitations [71,72]. Additionally, misconceptions regarding the size of stars persist among learners [73].

To (a) address these persistent ideas that do not align with the scientific view and (b) enhance students' learning experiences of astronomy concepts, researchers have explored the impact of out-of-school learning, particularly through planetarium visits. Lelliott [74] and Dunlop [75] conducted investigations into the effects of planetarium visits on secondary school students' understanding of astronomy topics. Lelliott's study [74] hints at different cognitive levels of students' knowledge, with planetarium visits leading to changes in their initial ideas and a shift toward more scientifically accurate notions of celestial motion.

C. Research desiderata

To succinctly summarize, research in astronomy education has revealed that many learners possess limited and fragmented ideas about stars. Prevailing students' ideas that do not align with the scientific views encompass various aspects such as star formation, size, their relationship to planets, and their position in space. These challenges persist across diverse age groups and educational settings, underscoring the need for further investigation. Specifically, two key research desiderata emerge:

1. *Comprehensive analysis of students' views of stars*: The studies presented in Sec. II B have predominantly focused on isolated aspects of stars, leaving some areas unexplored. Consequently, there is a dearth of instruments capable of capturing students' views comprehensively, encompassing the multifaceted nature of the topic [65]. A comprehensive analysis and identification of students' views of stars, spanning the various subaspects (e.g., from star formation to spectral aspects), remains elusive. Recognizing the prevalent ideas among learners that contradict current scientific views, however, is vital, as it enables the development of targeted instructional interventions. Thus, a pressing research need exists to develop an instrument that can holistically assess learners' views on different aspects related to stars.
2. *Cross-age exploration of students' learning about stars*: Although existing cross-age studies in astronomy have examined students' learning during limited periods of their school paths [76,77], a comprehensive overview of students' learning about stars across the entirety of secondary education is lacking. Understanding the evolution of students' views of stars over time is crucial for designing instructional strategies that align with their cognitive development and evolving needs. In fact, as there is evidence that interest [9] and cognitive processes [78] are changing during adolescence (roughly school years 7–12), these years of change are of special interest for research on cognitive development and the development of ideas in physics. By gaining insights into the progression of students' views throughout this part of their secondary education, educators and curriculum developers can tailor interventions that effectively address conceptual challenges at different stages of development.

In this article, we tackle both of these research desiderata: On the one hand, we aim at developing an instrument to economically assess learners' views of different aspects regarding stars. On the other hand, we use this new instrument to explore the development of secondary school students' views of stars throughout their secondary education during adolescence, and hence, aim at gaining a comprehensive overview of students' ideas of stars that may not align with the current scientific views and their development through school years 7 to 12.

III. RESEARCH QUESTIONS

The first research objective of this paper is to develop an instrument that can economically assess learners' views on various aspects of the stars and that performs well psychometrically on a sample of secondary school students. The second research objective is to gain insights into secondary school students' views of stars by examining their continuous development throughout secondary education and identifying ideas that contradict the current scientific views. Hence, we aim at clarifying the following research questions:

RQ1: How do the proportions of students' ideas about stars aligning with the scientific views compare among lower, middle, and upper secondary school students?

RQ2: What ideas that do not align with the scientific view on various aspects of stars are prevalent among secondary school students?

IV. DESIGN AND EVALUATION OF THE INSTRUMENT

In this section, we provide an overview of the development and evaluation of an instrument suitable for the assessment of learners' views of various aspects of stars. This endeavor aligns with the research objective outlined in the previous section (see Sec. III) and, hence, contributes to the validity and reliability of the research findings put forth in this paper.

A. Design of the instrument

1. Determination of target group and question format

Research question 1 addresses the evolution of learners' views of stars throughout their secondary education. Hence, the primary target group are secondary school students. To allow for the economic (i.e., time saving) identification of students' ideas about stars on a large scale (see research question 2), the use of rating scale items seems to be a sensible approach, which has been used in previous research aimed at economic assessment of students' views in science education research [79,80]. In addition, unlike dichotomous items, the use of a rating scale makes it easier to move away from a simple "right/wrong" classification of students' responses—which is particularly important when assessing students' views, as only two rating options (e.g., yes/no, right/wrong, agree/disagree) may force respondents to adopt a view that they would not have naturally formulated in some cases. Consequently, in our instrument, we decided to include statements about different aspects of stars alongside a four-point rating scale (1 corresponds to "I do not agree," 2 to "I rather do not agree," 3 to "I rather agree," and 4 to "I agree"). We decided to include a response option for abstaining to ensure that participants were not compelled to choose either agreement or disagreement when uncertain. Additionally, this reduces the likelihood of participants guessing their responses (for similar arguments, see Ref. [81]).

It is crucial to emphasize that the ratings provided by the participants for the statements in the instrument were not evaluated dichotomously. Instead, we employed a categorization approach to classify the students' ratings into the following distinct categories: (a) "in line with the scientific view," (b) "not aligning with the scientific view," and (c) "abstained from voting." This categorization decision is well justified given the dynamic nature of research on stars and their formation, which continues to grapple with unresolved questions in the field, as thoroughly discussed in the introduction of this article (cf. Sec. I). However, we refrain from a more fine-grained categorization that takes into account the degree of (dis)agreement among students since the meaning assigned to the rating options is highly subjective and may therefore vary from person to person. For example, if a questionnaire item contained a statement that deviated from the currently accepted scientific view of stars and a student selected the response option "I rather agree," we categorized the student's rating as "not in line with the scientific view".

Finally, we highlight that we deliberately ask for students' views or students' ideas instead of students' conceptions in the research questions underlying this study. The term conceptions would refer to deeper underlying notions or functional patterns that people give to phenomena and representations [82]. However, we assume that in this study, we also collect new *ad hoc* triggered mental connections of the learners, which are provoked by the items of the research instrument developed in this study. These are then not already internalized conceptions, but we refer to those as students' views or ideas that potentially are more superficial in nature. The distinction of the terms students' views or ideas from the term students' conceptions is in line with diSessa's knowledge-in-pieces perspective on learning (e.g., see Refs. [83,84]) in which students' views or ideas are regarded the "product of occasional mismatches between p-prims and contexts" [85] (p. 10) and aims to emphasize this mismatch.

2. Description of the content domain and item development

With our instrument, it should be possible to capture students' views of precisely those aspects of stars that are relevant for astronomy education at the secondary level. Therefore, the aspects to be included were initially identified based on (German) secondary school curricula using physics school textbooks from four different German federal states (Lower Saxony, Saxony, North Rhine-Westphalia, and Bavaria). Additionally, we used typical undergraduate textbooks (e.g., [10,86,87]) as well as scientific articles on astronomy education research that cover student learning (e.g., see Refs. [39,67,75,88,89]) for the description of the relevant content domains. This procedure led to the identification of five thematic domains covering the relevant aspects of stars:

- Domain 1: Stars and solar system. This domain covers the Earth-Sun-Moon relationship as well as their classification as celestial objects.
- Domain 2: Formation and evolution of stars. This domain covers topics such as stars' origin, age, life, and death.
- Domain 3: General properties and motion of stars. This domain includes key aspects regarding the properties and motions of stars, such as size or apparent motion.
- Domain 4: (Sub-)stellar objects. This domain comprises the main aspects of binary stars, brown dwarfs, white dwarfs, and pulsars.
- Domain 5: Color and brightness. This domain covers questions on stars' specific colors and their brightness.

These domains cover aspects of the key ideas relevant to the description of stars and their functioning, as identified in the course of developing the teaching-learning sequences on stars in their evolution published by Galano *et al.* [90] (e.g., stars' parameters, as touched upon in our domain 3) or Colantonio *et al.* [91] (e.g., functioning and evolution, as touched upon in our domain 2), while extending them with respect to (sub)stellar objects (domain 4).

In the initial iteration of the instrument development, we meticulously devised a comprehensive set of 82 rating scale items, encompassing the five thematic domains (i.e., subscales) mentioned above, to be subjected to thorough evaluation (see Sec. IV B). These 82 items were either developed newly or adopted from earlier instruments and empirical findings gained from prior research: for example, distractors used in Bailey's instrument, see Ref. [92], (p. 110), which were chosen by many students, were used as item statements in our instrument, thus building on students' views.

B. Evaluation of the instrument

1. Expert survey

To ensure the content validity of the instrument, we conducted an expert survey involving a panel of three esteemed faculty members with extensive expertise in astronomy research and teaching. The expert survey encompassed two key aspects: content evaluation and linguistic refinement.

Regarding content evaluation, the experts provided valuable insights on the scientific accuracy of the items and their alignment with the five content domains covered in Sec. IV A 2. Their comments helped identify any necessary adjustments or reallocations. Simultaneously, the experts scrutinized the language employed in the items, offering recommendations for rephrasing where deemed essential.

An exemplary change resulting from the expert survey was the adaptation of the items to represent positive statements. Therefore, for example, the item "Not all stars have a planetary system" was changed to "All stars have a planetary system" (item 2–3, for the item formulations, see Tables IX–XIII). Also, the wording was changed in several places to make the items more accessible for the students or to be more precise: Item 3–4, for example, had the original wording of "There is no gravitational force acting on stars," which was then changed to "Stars underlie gravitational pull." Similarly, item 3–5, which was initially worded "Stars have no gravitational force." was changed to "Stars exert attraction on things" after the expert input. The experts also suggested further items to be added regarding the different content domains: For example, item 2–16 "A supernova immediately destroys a large part of the galaxy" was provided as an additional suggestion for the content domain 2 "Formation and Evolution of Stars."

Based on the invaluable feedback obtained from the expert survey, we refined the item set, resulting in a total of 65 revised items. These revised items were then subjected to further piloting.

2. Psychometric characterization

Finally, we conducted a psychometric evaluation of the remaining items. Therefore, the items were administered to a sample of $N = 390$ secondary school and college students. The psychometric characterization of the items based on classical test theory was carried out to select the final item set for the five subscales. This involved examining the items' difficulties, with an accepted range of 0.2–0.8 according to Ref. [93], as well as their discriminatory powers, with accepted values ≥ 0.2 according to Ref. [94]. Additionally, Cronbach's alpha was calculated as an estimator of the internal consistency of all five subscales [95].

During the assessment, a total of 10 items were identified and subsequently excluded due to their inadequate psychometric characteristics. Consequently, the final instrument comprises 55 items, distributed among the five subscales. A comprehensive overview of the item distribution per subscale, along with the internal consistencies of each subscale, the average item difficulties and discriminatory powers, is given in Table I. Furthermore, to enhance clarity, we have included a sample item for each subscale. The final version of the instrument, with items arranged by subscales and alongside potential references, can be found in the Appendix of this paper.

Taken together, the rigorous process of development and evaluation culminated in a final version of the instrument that exhibits robust psychometric properties and enables a reliable assessment of secondary school students' views of various aspects of stars. Subsequently, this refined instrument was employed in our main study, which focuses on addressing the research questions outlined in Sec. III. In the following section, we provide a detailed account of the

TABLE I. Overview of the five subscales D1 to D5 (including the number N of items comprised and the Cronbach's alpha values) of the final version of the instrument used in this study to approach a clarification of the research questions alongside the average item difficulties and the average discriminatory powers of the corresponding items. We refrain from reporting the item difficulties and discriminatory powers of the single items due to the large number of items.

Subscale	N	α	∅ item difficulty	∅ discriminatory power	Example item
D1: Stars and solar system	9	0.69	0.65	0.36	The Sun is the largest star in the Universe.
D2: Formation and evolution of stars	16	0.83	0.56	0.44	Stars do not form and die, they only undergo changes over time.
D3: General properties and motion of stars	15	0.78	0.66	0.39	All stars have the same mass.
D4: (Sub-)stellar objects	8	0.79	0.35	0.50	All stars end as white dwarfs.
D5: Color and brightness	7	0.73	0.70	0.46	All stars are white.

methodology employed in our cross-age study, followed by the presentation of our findings in Sec. VI.

V. METHODS

A. Study design, sample, and instrument

A cross-age study design was chosen to approach a clarification of our research questions as has been done in previous studies concerned with similar research objectives (cf. [96,97]). The sample comprised $N = 366$ German secondary school students (from four federal states, namely Lower Saxony, North Rhine-Westphalia, Saxony, and Bavaria), divided into three different cohorts, enabling a deeper investigation of the temporal progression of students' views of stars throughout secondary education (for limitations of this approach, see Sec. VIII): We included participants from various grades that roughly cover the phases of adolescence, namely $N_1 = 148$ (81 female) students from grades 7 to 8 at the start of adolescence (aged 13–14 years), $N_2 = 151$ (70 female) students from grades 9 to 10 during adolescence (aged 15–16 years), and $N_3 = 67$ (31 female) students from grades 11 to 12 during late adolescence (aged 17–18 years). These age ranges and school years have proven useful for inclusion in earlier educational research focused on the learners' cognitive development during adolescence [98–100]. In the following, we will refer to these subsamples as cohort 1 (lower secondary school and early adolescence), cohort 2 (middle secondary school and middle of adolescence), and cohort 3 (upper secondary school and late adolescence), respectively. The participants did not receive any instruction as part of this study prior to test administration beyond their regular physics lessons and participation was completely voluntary as well as uncompensated. It it is noteworthy that current physics curricula in Germany deal with astronomy topics in a rather superficial manner, in particular, many of the topics assessed by our instrument remain a fringe topic throughout the entirety of secondary education.

The data collection, took place at the end of the school year 2021/2022 when the instrument presented in Sec. IV was administered as a paper-pencil test in the German language (for all items of the five subscales in the English language, see the Appendix). The students were allotted a complete double lesson, lasting a total of 90 min, to engage with the instrument, effectively eliminating any time constraints. The participants' processing times ranged approximately between 35 and 75 min. The Cronbach's alpha values for the five subscales calculated from the main study data (cf. Table II) were stable compared to the ones obtained in the pilot study (cf. Sec. IV B 2).

B. Data analysis

For each study cohort 1–3, we report the descriptive statistics regarding the percentage of responses in line with the scientific views on the items of each of the five subscales of the instrument, including mean, standard deviation (SD), median, as well as minimum (Min) and maximum (Max) values. Analyses of variance (ANOVAs) were conducted to check for differences between the three cohorts, with corresponding Tukey-Kramer *post hoc* tests to check for significant differences between the groups (see research question 1). As a measure of effect size for the overall comparisons, we used partial eta squared (η_p^2) where the commonly used categorization of small ($\eta_p^2 < 0.06$), medium ($0.06 \leq \eta_p^2 < 0.14$), and large ($0.14 \leq \eta_p^2$) effects was applied [101]. As a measure of effect size regarding the pairwise comparisons, we used Cohen's d alongside the established ranges of small ($d < 0.5$), medium ($0.5 \leq d < 0.8$), and large ($0.8 \leq d$) effect sizes [101]. To ensure the assumption of homogeneity underlying ANOVA, we employed Levene's test [102]. Additionally, the normal distribution of the data was assessed using the Shapiro-Wilk test [103].

TABLE II. Internal consistencies of the subscales of the instrument based on the data of the main study sample. These compare well with the ones obtained in the pilot study (cf. Sec. IV B 2). For the descriptions of the subscales, see Sec. IV A 2 and Table I, respectively.

	Cronbach's alpha		
Subscale	Cohort 1	Cohort 2	Cohort 3
1	0.67	0.70	0.74
2	0.86	0.87	0.79
3	0.74	0.79	0.79
4	0.86	0.80	0.84
5	0.71	0.73	0.71

To analyze the students' views in terms of scientific accuracy (see research question 2), we employed the categorization of responses described in Sec. IV, namely (a) "in line with the scientific view," (b) "may not align with the scientific view," and (c) "abstained form voting." The proportion of agreements with statements that may not align with the current scientific views was analyzed (see Tables IX–XIII) and an interpretation of the corresponding students' views—also in light of prior research—is given in the Sec. VII.

VI. RESULTS

In the following, we report the results of our study, separated by domain. For each of the five investigated domains, we will first report descriptive statistics and will, second, give the ANOVA results to evaluate the development of secondary school students' views of stars. To provide a concise overview, the ANOVA results are gathered in Table III. A more in-depth view of the students' responses, and hence, the ideas on stars apparent among the secondary school students is provided in Tables XIV–XVIII. In Sec. VII, we will elaborate on ideas among secondary

TABLE III. Results of ANOVAs comparing the percentage of responses aligned with the scientific view on all items across the three cohorts (1 corresponding to lower secondary students, 2 to middle secondary students, and 3 to upper secondary students) and across all subscales (domains D1 to D5). The p values reported in the last three columns belong to a Tukey-Kramer *post hoc* test. Cohen's d coefficients as measures of effect size for the pairwise comparisons are provided in the corresponding Figs. 1–5.

								Post hoc test		
Domain		Sum of squares	df	Sum of squares of residual error	F	p	η_p^2	1–2	1–3	2–3
D1	Between groups	3.30	2	1.65	39.7	<0.001	0.18	<0.01	<0.01	<0.05
	Within group	15.10	363	0.04						
D2	Between groups	2.48	2	1.24	28.0	<0.001	0.13	<0.01	<0.01	<0.01
	Within group	16.06	363	0.04						
D3	Between groups	1.82	2	0.91	24.0	<0.01	0.12	<0.01	<0.01	<0.01
	Within group	13.78	363	0.04						
D4	Between groups	0.65	2	0.32	4.46	<0.05	0.02	0.94	<0.05	<0.05
	Within group	26.26	363	0.07						
D5	Between groups	1.89	2	0.94	13.2	<0.001	0.07	<0.05	<0.001	<0.01
	Within groups	25.99	363	0.07						

school learners. We refer to certain items of the instrument with the abbreviation x-y, which stands for item y of subscale x (again, see Tables XIV–XVIII).

A. Domain 1: Stars and solar system

Table IV shows the descriptive statistics for all cohorts regarding items of domain 1. While cohort 1 students averaged 48.4% scientifically accurate answers, the proportion of responses aligning with scientific views for cohort 2 is 64.2% and further increases to 72.6% for cohort 3. A similar observation can be made for the median.

The trend observed in Table IV is statistically substantiated by the ANOVA results (cf. Table III). The difference between the three cohorts is statistically significant [$F(2,363) = 39.7, p < 0.001; \eta_p^2 = 0.18$]. Comparing the three cohorts directly yields a statistically significant difference between cohort 2 and cohort 3 ($p < 0.05$) with an effect size of $d = 0.42$. Cohort 1, on the other hand, differs highly significantly from both cohort 2 ($p < 0.01$) and cohort 3 ($p < 0.01$) with medium to high effect sizes of $d = 0.52$ and $d = 1.07$, respectively. These results are summarized in the form of boxplots in Fig. 1—for the presentation of our boxplots, the whiskers indicate $1.5 \times$ IQR throughout this article, where IQR is the interquartile range. A more in-depth view of the students' ideas is provided in Table XIV. In Sec. VII, we will elaborate on the ideas apparent among secondary school learners in terms of content and against the backdrop of prior research.

TABLE IV. Descriptive statistics for the percentage of responses on all items of domain 1 that are in line with the scientific views, separated by cohort.

Cohort	Mean	SD	Median	Min	Max
1	48.4	22.9	55.6	0.0	88.9
2	64.2	18.6	66.7	11.1	100
3	72.6	18.3	77.8	22.2	100

B. Domain 2: Formation and evolution of stars

Table V shows the descriptive statistics for all cohorts regarding items of domain 2. While mean and median are, on average, slightly lower compared to domain 1, we again observe an increase of all metrics with the exception of standard deviation.

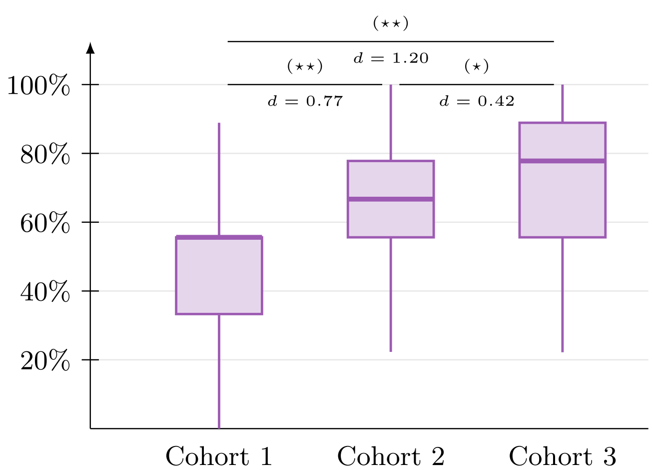

FIG. 1. Boxplot for the percentage of responses on all items of domain 1 that are in line with the scientific views. Asterisks indicate the statistical significance of Tukey-Kramer *post hoc* pairwise comparisons ($^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$), whereas Cohen's d is reported as a measure of effect size.

TABLE V. Descriptive statistics for the percentage of responses on all items of domain 2 that are in line with the scientific views, separated by cohort.

Cohort	Mean	SD	Median	Min	Max
1	42.9	21.4	43.8	0.0	75.0
2	53.8	20.8	50.0	6.3	100
3	65.4	20.9	68.8	6.3	100

TABLE VI. Descriptive statistics for the percentage of responses on all items of domain 3 that are in line with the scientific views, separated by cohort.

Cohort	Mean	SD	Median	Min	Max
1	55.7	22.2	60.0	0.0	93.3
2	64.2	17.5	66.7	13.3	100
3	75.3	17.2	80.0	20.0	100

The trend observed for domain 1 also holds for domain 2 (cf. Table III): The difference between the cohorts is very highly significant [$F(2,363) = 28.0$, $p < 0.001$; $\eta_p^2 = 0.13$]. Likewise, all between-group comparisons show highly statistical significance ($p < 0.01$) with effect sizes ranging from $d = 0.52$ between cohorts 1 and 2, $d = 0.55$ between cohorts 2 and 3 as well as $d = 1.07$ between cohorts 1 and 3. Hence, not only do the groups differ highly significantly, but medium to high effect sizes can be associated with each of the pairwise comparisons, indicating a steady increase of scientifically accurate views across secondary education. These results are summarized in the form of boxplots in Fig. 2. A more in-depth view of the students' ideas is provided in Table XV. In Sec. VII, we will elaborate on the ideas apparent among secondary school learners in terms of content and against the backdrop of prior research.

C. Domain 3: General properties and motion of stars

The descriptive statistics for all cohorts regarding items of domain 3 are provided in Table VI. The values overall compare very well to the ones from domains 1 and 2 (cf. Tables IV and V).

The ANOVA results for domain 3 show almost identical values with those of domain 2 (cf. Table III). The cohorts differ highly statistically significantly [$F(2,363) = 24.0$, $p < 0.01$; $\eta_p^2 = 0.12$] and the same is true for all pairwise cohort comparisons ($p < 0.01$ each). The corresponding effect sizes are all high with $d = 0.44$ between cohorts 1 and 2, $d = 0.57$ between cohorts 2 and 3 as well as $d = 1.01$ between cohorts 1 and 3. Hence, as before, a continuous improvement in scientifically accurate views held by students can be observed from lower to middle and, finally, higher secondary education. A summary of these results in terms of boxplots is provided in Fig. 3. A more in-depth view of the students' ideas is provided in Table XVI. In Sec. VII, we will elaborate on the ideas apparent among secondary school learners in terms of content and against the backdrop of prior research.

D. Domain 4: (Sub-)stellar objects

Table VII provides an overview of descriptive statistics for all items of domain 4. This domain holds the overall lowest descriptive statistics. While cohort 1 students averaged 32.9% scientifically accurate answers, the proportion of responses aligning with scientific views for

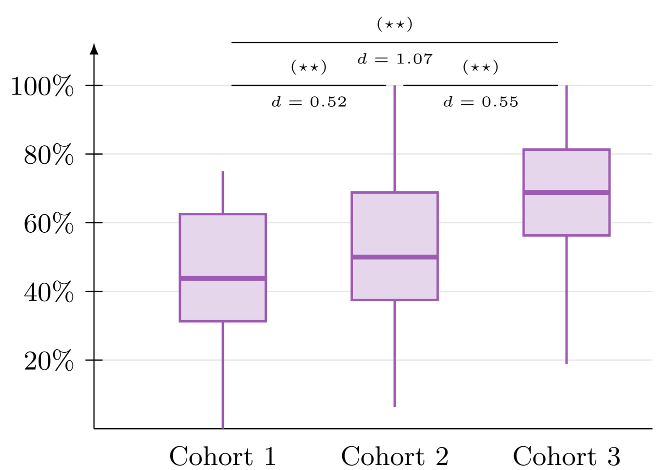

FIG. 2. Boxplot for the percentage of responses on all items of domain 2 that are in line with the scientific views. Asterisks indicate the statistical significance of Tukey-Kramer *post hoc* pairwise comparisons ($^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$), whereas Cohen's d is reported as a measure of effect size.

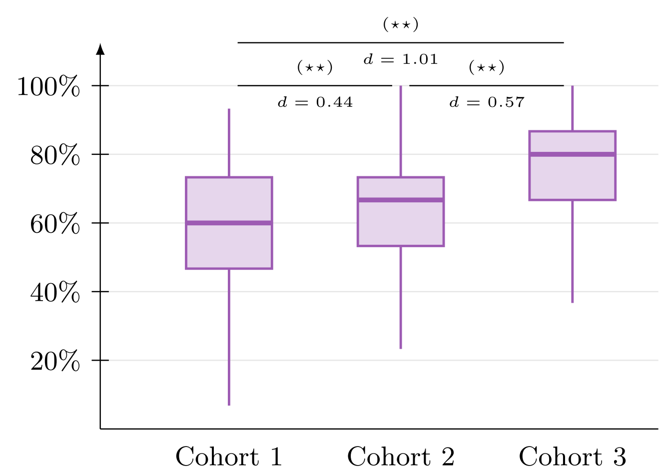

FIG. 3. Boxplot for the percentage of responses on all items of domain 3 that are in line with the scientific views. Asterisks indicate the statistical significance of Tukey-Kramer *post hoc* pairwise comparisons ($^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$), whereas Cohen's d is reported as a measure of effect size.

TABLE VII. Descriptive statistics for the percentage of responses on all items of domain 4 that are in line with the scientific views, separated by cohort.

Cohort	Mean	SD	Median	Min	Max
1	32.9	27.9	37.5	0.0	87.5
2	31.8	26.4	25.0	0.0	100
3	43.1	25.7	37.5	0.0	100

TABLE VIII. Descriptive statistics for the percentage of responses on all items of domain 5 that are in line with the scientific views, separated by cohort.

Cohort	Mean	SD	Median	Min	Max
1	58.9	29.5	71.4	0.0	100
2	67.3	24.9	71.4	0.0	100
3	78.9	24.3	85.7	0.0	100

cohort 2 is 31.8% and 43.1% for cohort 3. The median reflects this performance and, for the first and only time, there are participants with no responses that are in line with the scientific view. Thus, on average, secondary school students' views differed the most from the current scientific view regarding (sub-)stellar objects. We will elaborate on this observation in Sec. VII.

With the scores being relatively low in each cohort, the variance in performance did not differ as much compared to the other domains. While overall group comparison still results in a statistically significant difference [$F(2, 363) = 4.46$, $p < 0.05; \eta_p^2 = 0.02$], no statistically significant difference was found between cohorts 1 and 2 ($p = 0.94$). The differences between cohorts 1 and 3 as well as cohorts 2 and 3 were found to be statistically significant by Tukey-Kramer *post hoc* tests ($p < 0.05$ each), but the corresponding effect sizes were low with $d = 0.38$ and $d = 0.42$, respectively. Thus, for the domain of (sub-)stellar objects, we record the overall least distinguishable progress throughout secondary education, with only the later grades (11 and 12) meaningfully separating themselves from the rest. The boxplots reflecting this observation are presented in Fig. 4. A more in-depth view of the students' ideas is provided in Table XVII. In Sec. VII, we will elaborate on the ideas apparent among secondary school learners in terms of content and against the backdrop of prior research.

E. Domain 5: Color and brightness

Table VIII shows the descriptive statistics for all cohorts regarding items of domain 5. In contrast to domain 4, domain 5 records the overall highest mean percentages of responses in accordance with the scientific views with 58.9% for lower, 67.3% for middle, and 78.9% for higher secondary school students. The minimum and maximum in each cohort are 0.0% and 100.0%, respectively.

The ANOVA results for domain 5 demonstrate a much-pronounced improvement along the trajectory of secondary education (cf. Table III). The overall difference between the three cohorts is highly statistically significant [$F(2, 363) = 13.2$, $p < 0.001; \eta_p^2 = 0.07$]. A Tukey-Kramer *post hoc* test indicates that the advancement from grades 7–8 to grades 9–10 is statistically significant ($p < 0.05$) with an effect size of $d = 0.31$, whereas the difference between grades 9–10 and 11–12 is highly statistically significant ($p < 0.01$) with an effect size of $d = 0.48$. Finally, the difference between cohorts 1 and 3 is very highly

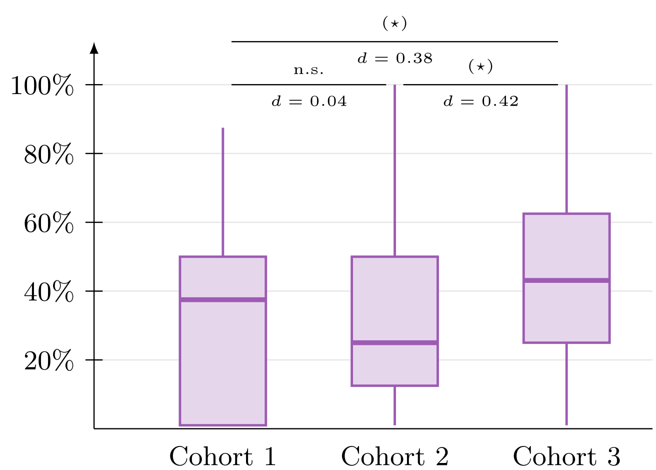

FIG. 4. Boxplot for the percentage of responses on all items of domain 4 that are in line with the scientific views. Asterisks indicate the statistical significance of Tukey-Kramer *post hoc* pairwise comparisons ($^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$), whereas Cohen's d is reported as a measure of effect size.

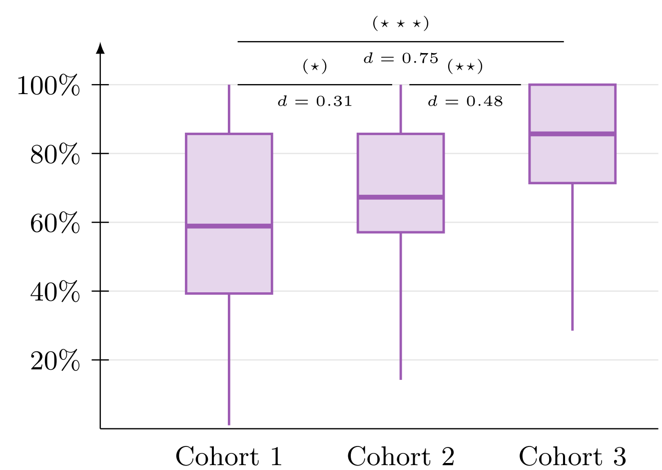

FIG. 5. Boxplot for the percentage of responses on all items of domain 5 that are in line with the scientific views. Asterisks indicate the statistical significance of Tukey-Kramer *post hoc* pairwise comparisons ($^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$), whereas Cohen's d is reported as a measure of effect size.

statistically significant ($p < 0.001$) with a medium effect size of $d = 0.75$. The boxplots reflecting this continuous increase in accurate scientific views are illustrated in Fig. 5. A more in-depth view of the students' ideas is provided in Table XVIII. In Sec. VII, we will elaborate on the ideas apparent among secondary school learners in terms of content and against the backdrop of prior research.

VII. DISCUSSION

In this section, we contextualize our findings against the backdrop of prior astronomy education research. In particular, we shed light on the extent to which our findings are consistent with prior research, identify areas where differences have emerged, and highlight new contributions to the understanding of secondary school students' views of stars. Moreover, we discuss the implications of our findings for both astronomy education research and practice.

In general, our cross-age analysis reveals a progressive development of students' perspectives on star-related topics throughout their secondary school education. The statistically significant increase in the percentage of responses aligned with current scientific views from lower to upper secondary school applies to all aspects of stars examined in this study, as demonstrated through ANOVAs and *post hoc* pairwise comparisons (see Sec. VI). Furthermore, our study expands upon existing literature, which primarily focuses on students' views of the nature, apparent motion, and properties of stars [20,64,76].

A. Discussion of findings regarding domain 1

The percentage of responses aligning with scientific views in domain 1 (stars and solar systems) shows an increase from 48.4% among lower, and 64.2% among middle to 72.6% among upper level secondary school students (see Table IV). An item with a very big gain of alignment toward scientific views was, for example, item 1–9, where progressively more students disagreed with the statement that the Sun is the largest star in the Universe (75.0% of the lower, 43.7% of the middle, and 34.3% of the upper secondary students). On the single-item level, we also found that a majority of the study participants (75.0% of the lower, 70.9% of the middle and 62.7% of the upper secondary students) agreed that there are hundreds of stars in our solar system (item 1–5, see Table XIV). This idea does not align with the scientific view and has already been reported in Ref. [104]. This student view can possibly be explained by the confusion of the terms solar system and stellar system as indicated by Rajpaul *et al.* [43]. A further item that stands out due to a large share of views that may not align with scientific views is item 1–4, indicating that between 32.4% (lower), 46.1% (middle), and 41.8% (upper) secondary school students believe that "the planets and the Sun were formed at the time of the Big Bang." Another perspective among our study participants is that metals have existed since the Big Bang (item 1–6, see Table XIV): 46.6% of the lower, 40.4% of the middle, and 29.9% of the upper secondary students were of that opinion, while in addition, around a third of the students at all grade levels were undecided (32.4%, 25.8% and 28.4%, respectively). This finding supports the study by Slater *et al.* [70], in which around a third of the students have been reported to hold the view that heavy atoms have existed since the Big Bang and that the Big Bang directly created everything.

In contrast to previous research by Dunlop [75], Philips [105], or Comins [104], in our study, we did not find evidence of students misconceiving that the Sun is not a star: Only approximately 10% of lower and middle secondary school students, and merely 3% of upper secondary school students, subscribed to this view (see item 1–8 Table XIV). It is noteworthy, that it remains unclear which research methods Philips [105] and Comins [104] employed to uncover this confusion to be among learners. Additionally, Comins [104] stated that many students would believe that the Sun is bigger than other stars. In our study, we identified this view to be among the lower secondary school students (60.1%, see Table XIV), while only 37.7% of the middle and 29.9% of the upper secondary students agreed with item 1–9 ("The Sun is the largest star in the Universe."). In general, a view more in line with a scientific view is progressively more facilitated during the years, though some ideas connected to the Big Bang appeared to be less malleable. Still, the general change toward a more scientific view in domain 1 is a positive result and the current instructional environments facilitate this development.

B. Discussion of findings regarding domain 2

Similar to domain 1, the percentage of responses aligning with scientific views in domain 2 (formation and evolution of stars) shows an increase from 42.9% among lower secondary school students to 53.8% among middle secondary school students, and 65.4% among upper secondary school students (see Table V). An item with a very big gain of alignment toward scientific views was, for example, item 2–10, where progressively more students agreed that stars can change in color (16.9% of the lower, 35.8% of the middle, and 79.1% of the upper secondary students). But on the single-item level, we also found that—in line with earlier research conducted by Agan [67]—more than half of our participating students agreed with item 2–13 stating that stars would fade and disappear over time (63.5% of the lower, 58.3% of the middle, and 53.7% of the upper secondary students). Finally, item 2–16 reveals that almost half of the students of each cohort responded that "a supernova immediately destroys a large part of the galaxy" indicating a skewed perception of astronomic scales regarding the relation between the galactic scale and the scale of a supernova. This is congruent to previous research by

Salimpour *et al.* [106] who found that students often struggle with reasoning when it comes to spatial relations in the context of cosmology, due to hard-to-grasp numbers that "offend everyday logics" [106] (p. 122). In summary, the developmental aspects of stars and the relevant scales appear to develop toward a scientific view with the exception of some views about a star's fate. Again, the general change toward a more scientific view in this domain is a positive result and the current instructional environments appear to facilitate this development.

C. Discussion of the findings regarding domain 3

A study similar to the one presented in this paper, focusing on star-related aspects, was conducted by Plummer [76]. In her cross-age study, Plummer aimed to assess students' views of celestial motion and identify any misconceptions held at different grade levels (1st, 3rd, and 8th-grade students). One part of her study specifically focused on students' views of the apparent motion of stars. Among the sample students, two main perspectives emerged: those who could provide a general description of stars moving slowly across the night sky (40% of 8th graders, 50% of 3rd graders, and 35% of 1st graders) and those who believed that stars never move (40%, 40%, and 25%, respectively). The remaining students held various perspectives, including the idea that stars only move at the end of the night. Our study complements Plummer's findings by providing insights into the views of older students during adolescence regarding the apparent motion of stars. Among our cohort 1 students (grades 7 and 8), only 18.2% believed that stars are stationary and fixed in the sky (item 3–11, see Table XVI). This percentage was even lower among cohort 3 students (grades 11 and 12) at 14.9%. In addition, our findings align with the research conducted by Agan [67], revealing a comparable proportion of students (approximately 14%) who share the view that stars are stationary. This stands in contrast to studies published in Refs. [28,46,70], which reported a higher prevalence of this view among approximately 40% of students. On a further note, our findings are supported by Plummer's study, according to which more than "one-half of the students in the first, third, and eighth grades do not think that we see different stars in the sky during the night (65%, 60%, and 65%), respectively" [76] (p. 1598). In our study, 60.8% of lower secondary school students, 75.5% of middle secondary school students, and 82.1% of upper secondary school students agreed with item 3–12, which states that we see different stars over the course of a night (see Table XVI). In summary, our findings suggest a continued evolution of students' views on the apparent motion of stars throughout secondary education.

Another cross-age study on basic astronomy concepts was conducted by Trumper, which included both junior (grades 7–9) [107] and senior (grades 10–12) [108] high school students. In Trumper's study, only 36% of junior high school students were aware that stars are the farthest objects from the Earth [107] (p. 1117), while this percentage increased to 49% among senior high school students [108] (p. 103). In our study, the percentages of students agreeing with the scientific views on this matter were slightly higher in the sample cohorts. Among lower secondary school students, 43.9% agreed that stars are farther away from the Earth than the Sun (item 3–9, see Table XVI). This percentage increased to 51.7% among middle secondary school students and 53.7% among upper secondary school students. Therefore, it appears that students' views regarding distances develop throughout their educational careers. This observation is further supported by the decreasing agreement with item 3–10, which states that the distance between stars is about the same as the distance between planets. While half of the lower secondary school students disagreed with this item, the disagreement percentage increased to 62.9% among middle secondary school students and 74.6% among upper secondary school students (see Table XVI). Our findings align with earlier research on learners' views of astronomical object distances from the Earth's surface (e.g., see Refs. [12,43,46]).

There are further items that stand out in this domain: For example, Comins [104] reported the view held by students that stars do not rotate, which is supported by our study since 43.9% of lower secondary school students (33.1% middle and 26.9% upper) agreed with item 3–15 stating "Stars don't rotate." Furthermore, 43.9% of lower, 29.1% of middle, and 37.3% of upper secondary school students held the misguided view that stars do not underlie gravitational pull, viewing stars as stationary celestial objects that do not interact through gravity with their surroundings. All in all, the learning environments facilitated a positive development when comparing cohort 1 to cohort 3 in domain 3.

D. Discussion of findings regarding domains 4 and 5

While the percentage of responses aligning with scientific views in domain 4 remains below 50% for students across all grade levels (see Sec. VII), we found a majority of students demonstrating views in line with current scientific understandings of the topic color and brightness of stars in domain 5 (see Sec. VIII). No statistically significant difference can be observed in the percentage of responses aligning with scientific views for items in domain 4 between lower (32.9%) and middle secondary school students (31.8%), see Table VII. However, from a deeper look into the response patterns on all items provided in Table XVII, it becomes clear that a substantial part of the students did not hold inaccurate views but rather abstained from voting. In terms of content, items 4–3, 4–5, as well as 4–7 all address brown and white dwarfs and have in common that on average, one third of all participants express views that may not align with the scientific view across these items. This finding can likely be attributed to the omission of (sub-)stellar objects from the German astronomy curricula at both lower and middle school levels. Insights into current curriculum developments regarding astronomy in German secondary schools can be found in Ref. [12].

The topic of color and brightness of stars shows a similar pattern: Surprisingly, a small but statistically significant difference exists in the percentage of responses aligning with scientific views for items in domain 5 between lower (58.9%) and middle secondary school students (67.3%). This discrepancy can possibly be attributed to either implicit learning, which occurs unconsciously as students engage with different topics, such as atomic physics (for more details on implicit learning, see Reber [109]), or the influence of informal learning environments [110], such as planetariums [75]. However, the data collected in this research do not allow for further unpacking of the impact of implicit learning or informal learning environments on students' progress with star-related concepts in secondary education. Hence, in future research, it would be valuable to explore (a) which topics in the secondary school curriculum facilitate implicit learning of astronomy and (b) the sources from which students gain insights into astronomy topics in informal settings. This seems particularly important as Wang [111] emphasized "that implicit learning is both a source of and an influence on intuitive scientific knowledge that is important for conceptual change" (p. 110). In terms of content, two items of this domain regarding the brightness of stars stand out:

- Item 5–6 ("The brightness of stars is constant", see Table XVIII), where a shift toward a perspective aligned with scientific views was observed from lower to upper secondary school levels: 39.2% of the lower secondary students disagreed with this statement but 86.6% of the upper secondary students agreed with this statement.
- Item 5–7 which states that "it is said that stars twinkle because they change their brightness": 45.9% of lower, 39.1% of middle, and 25.4% of higher secondary school students agreed with this statement.

In congruence with the response pattern on item 3–7, the brightness of stars constitutes a recurring theme that poses difficulties to students.

E. Implications for educational research

The results of this study imply that, in general, students' ideas about stars show a progressive alignment with the current scientific views. This quite positive development, however, lacks a clear explanation at this point, meaning that further extensive research is advised to clear up the various causes for the developments presented in this study: While teaching in the classroom appears to play a role, it is important to consider the potential influence of informal learning environments such as out-of-school visits [112,113] or explanatory videos [114]. These supplementary resources may contribute to the positive evolution of students' understanding. Additionally, no data about the traits of the students have been gathered. As previous research has shown, astronomical topics tend to rank highly in studies on interesting topics for students [9]. A thorough investigation into how and what affective traits facilitate the increase in alignment with scientific topics during the secondary school years could give new and more detailed insights into the role of interests in learning processes. Furthermore, the instrument utilized in this study has the potential to be employed in future research to identify the specific factors and properties of learning materials and methodologies that facilitate these favorable learning outcomes. By using this instrument, researchers can gain insights into the effective strategies and approaches that support students' development of ideas in the domain of stars and related astronomical concepts, complementing previous instruments devised for this purpose (e.g., see Refs. [55,92])—e.g., in terms of content regarding aspects of (sub-)stellar objects or in terms of methodology regarding the use of rating scale items. Though there is a progression toward a scientific view from cohort 1 to cohort 3 for all items (items 4–4 and 4–8), several items align less with scientific views during year 9–10 (cohort 2). These are in particular: 1–4, 2–9, 2–11, 2–15, 2–16, 3–5, 3–7, 3–11, 3–13, 4–1, 4–2, 4–3, 4–7, 4–8, 5–4 and 5–5. It is not clear whether this is due to students of this age being in the middle of adolescence from this fact alone, but this result aligns with the drop in interest during these years [9]. A closer examination of this result and its possible causes is recommended to be the focus of future research. Further investigation is also warranted to delve deeper into the exact mechanisms driving this generally positive learning trajectory. Additionally, exploring the comparative impact of various educational interventions and materials could shed light on the most effective approaches for promoting accurate scientific understanding of stars among students.

F. Implications for educational practice

From the findings presented in this study, we surmise that current educational practices available between the lower and upper secondary school level likely facilitate a basic development of ideas about stars that in general align progressively more with scientific views during these years. However, some persistent ideas have also been isolated that have also been found in corresponding literature (e.g., see Refs. [67,76,115]). Mainly, static ideas have been found especially in the lower high school classes and confusions of central ideas like stellar system and solar system have also been documented. Ideas of change and dynamics (stars changing brightness or stars changing color) might be better facilitated by incorporating learning materials that depict these dynamic properties (such as videos or comics, cf. [114,116]) or even enable learners to interact with the material (such as simulations, cf. [117]). Some topics that showed a generally small improvement toward scientific views were some aspects of the Big Bang and the final stages of stars. We therefore recommend that there should be careful emphasis on these topics in educational settings.

VIII. LIMITATIONS

While our study on students' views of stars throughout secondary education provides valuable insights, it is important to acknowledge several limitations. First, it is crucial to note that our research adopted a cross-age study design rather than a conventional longitudinal approach: Instead of longitudinally tracking the views of a specific group of students, we assessed students from different grade levels at a single time point. Our approach offers valuable cross-sectional data but (a) limits our ability to capture individual students' developmental trajectories and the specific changes in their views of stars over time and (b) may be subject to cohort bias [118].

Second, it is crucial to acknowledge that our study's focus on secondary school students from Germany poses limitations on the generalizability of our findings to other educational systems. Additionally, the diversity in secondary school astronomy curricula among the different German federal states further influences the outcomes of this study. To address this limitation, we intentionally included students from various federal states across Germany (north: Lower Saxony; middle: Saxony, North Rhine-Westphalia; south: Bavaria) in all age groups of our sample. Specifically, the respective federal states exhibit varying degrees of emphasis on astronomy within their respective secondary school curricula. Moving forward, to extend the applicability of the findings presented in this study, it is imperative to conduct future investigations into secondary school students' (mis-) conceptions of star-related topics beyond the confines of German samples. Such studies are vital to (a) validate and substantiate the conclusions drawn here, and (b) provide evidence-based recommendations for astronomy curricula, both in general and specifically tailored to Germany's educational context. Notably, despite the aforementioned restrictions, our sample includes learner cohorts representing three distinct age groups across the entire secondary school period: cohort 1 comprised students aged 13–14 years, cohort 2 with students aged 15–16 years, and cohort 3 with students aged 17–18 years. This diversity enables direct comparisons of our results with those obtained from learners of different age groups and stages of development, various educational settings, and across different countries in future research.

Third, the subsamples of lower and middle secondary school students consisted of approximately 150 participants each, whereas the number of participants in the upper secondary school group was only about half that size. This discrepancy arises from the fact that in Germany, students make a decision after completing grade 10 regarding whether to pursue a physics course in upper secondary school. Consequently, the total cohort of possible study participants becomes significantly smaller at this grade level. Furthermore, teachers are often less inclined to participate in research studies during the 12th grade due to the impending final exams. These factors imposed limitations on our sampling approach, resulting in an asymmetric distribution of participants among the three subsamples.

Another limitation stems from the question format used in our instrument, which consisted of closed-ended rating-scale items. While this format facilitated efficient data collection on students' views of stars on a large scale, it is important to recognize that the predetermined statements may have influenced participants' responses, potentially leading to the generation of *ad hoc* conceptions. To mitigate this limitation, future research should incorporate qualitative data collection methods, such as mind mapping or concept mapping [119–122], to gather in-depth insights that validate and expand upon our findings. This way, not only views held by students are compiled but also the reasons underlying those views can be explored more thoroughly. Furthermore, it is crucial to acknowledge that our study primarily focused on assessing students' ideas and views of stars, rather than their conceptual understanding of relevant aspects related to stars. While our study provides valuable perspectives, it is essential to complement these findings with investigations into students' conceptual understanding. This, perhaps, can be achieved by additionally utilizing the aforementioned Cosmology Concept Inventory [66] and comparing the assessed conceptions of students with their views on the topic under investigation.

Finally, at the qualitative level, an important limitation must be considered, which stems from the fact that the items of the survey represent a compromise of the following criteria: (i) intelligible wording, (ii) coverage of overarching concepts of astronomy and stars, rather than addressing all the details of a topic, and (iii) subject-specific correctness.

Item 2–13 of the survey (see Sec. II in the Appendix), which deals with stars fading, may serve as a prime example of this tradeoff: While there are objects that could be considered "stars" that do indeed fade rather quickly, such as variable stars or even space debris that appears like a star when observed from the Earth, in this item, we address a more surface-level idea—that stars are generally not fading during timescales of human observations. Hence, this item is intended to evaluate whether students have understood that nearly all observable stars are not, in fact, fading and that the stars in the sky are more or less permanent in brightness. In contrast, assessing students' more detailed ideas of special objects like variable stars is outside the scope of this item (see bullet (ii) in the above list). Again, developing additional items that cover further topics and aspects related to stellar and substellar objects and their evolution could help alleviate the constraints imposed by the aforementioned compromise.

IX. CONCLUSION

From the data collected and analyzed, our findings about students' ideas about stars compared to scientific views have shown that progressing from lower to upper secondary school, ideas start to align more with the current scientific views. This development in itself is positive, though the exact factors for it still need to be further determined. From the data gathered, some of the ideas are already being developed rather effectively, such as ideas about Sun's size when compared to other stars (see item 1–9 in Table XIV and 3–2 in Table XVI) or that stars can change both color and brightness (see item 2–10 in Table XV and 5–6 in Table XVIII) while others seem more robust, such as the idea that there are hundreds of stars in the solar system (item 1–5, see Table XIV) or that white dwarfs are suns (item 4–5, see Table XVII). In summary, our analysis indicates a positive trend of students' ideas aligning more closely with scientific views regarding stars as they progress through secondary school, although the specific contributing factors remain to be determined.

Anonymized data from the study is available on request from the authors.

APPENDIX: THE INSTRUMENT USED IN THIS STUDY

In this Appendix, we provide all items of the instrument used in this study, sorted by the five subscales included.

1. Subscale on domain 1: Stars and solar system

TABLE IX. Overview of the rating scale items related to the subscale on domain 1. The table includes the item numbers (No.), the statements themselves, indications of whether the statements do or do not align with the scientific view, and any references or sources of inspiration. The statements have been translated to the best of the authors' abilities into English.

No.	Item	In line with scientific view	Not in line with scientific view	Reference/ Inspired by
1–1	The Earth orbits the Sun and the Moon.		✗	
1–2	The Sun and the Moon orbit the Earth.		✗	
1–3	The Earth and the Moon orbit the Sun.	✗		
1–4	The planets and the Sun were formed at the time of the Big Bang.		✗	[39]
1–5	There are hundreds of stars in our solar system.		✗	[67,88]
1–6	Metals have existed in the Universe since the Big Bang.		✗	
1–7	The Moon is a star.		✗	[89]
1–8	The Sun is a star.	✗		[75]
1–9	The Sun is the largest star in the Universe.		✗	[67]

2. Subscale on domain 2: Formation and evolution of stars

TABLE X. Overview of the rating scale items related to the subscale on domain 2. The table includes the item numbers (No.), the statements themselves, indications of whether the statements do or do not align with the scientific view, and any references or sources of inspiration. The statements have been translated to the best of the authors' abilities to English.

No.	Item	In line with scientific view	Not in line with scientific view	Reference/ Inspired by
2–1	Stars already existed before the Big Bang.		✗	[39,88]
2–2	Stars don't form and die, they just change over time.		✗	
2–3	All stars have a planetary system.		✗	
2–4	Stars have spikes.		✗	
2–5	The surface of stars is almost completely covered by volcanoes.		✗	[67]
2–6	All stars are the same age.		✗	
2–7	We see the stars exactly as they are in this moment.		✗	
2–8	Stars undergo changes.	✗		
2–9	Stars change their size.	✗		

(Table continued)

TABLE X. *(Continued)*

No.	Item	In line with scientific view	Not in line with scientific view	Reference/ Inspired by
2–10	Stars change their color.	✗		
2–11	Massive stars can evolve into neutron stars.	✗		
2–12	Massive stars can evolve into black holes.	✗		
2–13	Stars fade and disappear over time.		✗	[67]
2–14	Supernova eruptions occur only in our galaxy.		✗	
2–15	All stars end up in a supernova.		✗	
2–16	A supernova immediately destroys a large part of the galaxy.		✗	

3. Subscale on domain 3: General properties and motion of stars

TABLE XI. Overview of the rating scale items related to the subscale on domain 3. The table includes the item numbers (No.), the statements themselves, indications of whether the statements do or do not align with the scientific view, and any references or sources of inspiration. The statements have been translated to the best of the authors' abilities to English.

No.	Item	In line with scientific view	Not in line with scientific view	Reference/ Inspired by
3–1	All stars are the same size.		✗	
3–2	There are stars that are larger than the Sun.	✗		[67]
3–3	All stars have the same mass.		✗	
3–4	Stars underlie gravitational pull.	✗		
3–5	Stars exert attraction on things.	✗		
3–6	All stars are about the same distance from the Earth as the Moon.		✗	
3–7	The brightest stars are closest to the Earth.		✗	[67]
3–8	All stars are the same distance from the Earth.		✗	
3–9	Stars are farther away from the Earth than the Sun.	✗		
3–10	The distance between stars is about the same as the distance between planets.		✗	
3–11	All stars are stationary—they are fixed in the sky and do not move.		✗	[36,76,123]
3–12	During the course of a night we see different stars.	✗		[76]
3–13	Stars seem to rise and set.	✗		[67,104]
3–14	The observable motion of the stars is a result of the Earth's rotation around its own axis.	✗		[17,73]
3–15	Stars don't rotate.		✗	[67,104]

4. Subscale on domain 4: (Sub-)stellar objects

TABLE XII. Overview of the rating scale items related to the subscale on domain 4. The table includes the item numbers (No.), the statements themselves, indications of whether the statements do or do not align with the scientific view, and any references or sources of inspiration. The statements have been translated to the best of the authors' abilities to English.

No.	Item	In line with scientific view	Not in line with scientific view	Reference/ Inspired by
4–1	Two orbiting stars (binary star system) would quickly collide.		✗	
4–2	There are long-standing binary star systems.	✗		
4–3	Brown dwarfs describe the final stage of a star.		✗	
4–4	White dwarfs are planets.		✗	
4–5	White dwarfs are suns.	✗		
4–6	White dwarfs are stars.	✗		
4–7	All stars end up as white dwarfs.		✗	
4–8	Pulsars are stars that alternately emit light of different intensities.		✗	

5. Subscale on domain 5: Color and brightness

TABLE XIII. Overview of the rating scale items related to the subscale on domain 5. The table includes the item numbers (No.), the statements themselves, indications of whether the statements do or do not align with the scientific view, and any references or sources of inspiration.

No.	Item	In line with scientific view	Not in line with scientific view	Reference/ Inspired by
5–1	All stars have the same color.		✗	
5–2	All stars are white.		✗	
5–3	All stars are yellow.		✗	[67,104]
5–4	There are blue stars.	✗		
5–5	Stars emit many colors of light.	✗		
5–6	The brightness of stars is constant.		✗	
5–7	Stars twinkle because they change their brightness.		✗	[67,104]

6. Distribution of students' responses on all items

TABLE XIV. Share of responses to all items of domain 1 that are in line with scientific views (+), not aligning with scientific views (−) or abstained form voting (○), for all three cohorts.

Item	Cohort	−	○	+	Item	Cohort	−	○	+
1–1	1	41.9	10.1	48.0	1–6	1	46.6	32.4	20.9
	2	13.9	2.0	84.1		2	40.4	25.8	33.8
	3	9.0	0.0	91.0		3	29.9	28.4	41.8
1–2	1	36.5	6.8	67.6	1–7	1	19.6	10.1	70.3
	2	12.6	0.7	86.6		2	11.9	6.0	82.1
	3	11.9	0.0	88.1		3	10.4	1.5	88.1
1–3	1	25.7	6.8	67.6	1–8	1	10.8	6.8	82.4
	2	21.2	3.3	75.5		2	11.3	2.0	86.6
	3	9.0	0.0	91.0		3	3.0	0.0	97.0
1–4	1	32.4	17.6	50.0	1–9	1	60.1	14.9	25.0
	2	46.1	10.6	43.0		2	37.7	6.0	56.3
	3	41.8	4.5	53.7		3	29.9	4.5	65.7
1–5	1	75.0	10.1	14.9					
	2	70.9	0.0	29.1					
	3	62.7	0.0	37.3					

TABLE XV. Share of responses to all items of domain 2 that are in line with scientific views (+), not aligning with scientific views (−) or abstained form voting (○), for all three cohorts.

Item	Cohort	−	○	+	Item	Cohort	−	○	+
2–1	1	52.0	18.2	29.7	2–9	1	14.2	20.9	64.9
	2	37.7	18.5	43.7		2	24.5	13.2	62.3
	3	23.9	11.9	64.2		3	11.9	9.0	79.1
2–2	1	71.6	10.1	18.2	2–10	1	50.7	32.4	16.9
	2	41.7	9.3	49.0		2	49.0	15.2	35.8
	3	25.4	3.0	71.6		3	38.8	9.0	79.1

(Table continued)

TABLE XV. *(Continued)*

Item	Cohort	−	∘	+	Item	Cohort	−	∘	+
2–3	1	50.0	18.2	31.8	2–11	1	7.4	50.7	41.9
	2	24.5	15.2	60.3		2	12.6	56.3	31.1
	3	11.9	9.0	79.1		3	10.4	44.8	44.8
2–4	1	7.4	24.3	68.2	2–12	1	37.2	35.1	27.7
	2	11.3	9.9	78.8		2	22.5	27.8	49.7
	3	4.5	9.0	86.6		3	16.4	22.4	61.2
2–5	1	33.1	27.0	39.9	2–13	1	63.5	10.8	25.7
	2	29.1	24.5	46.4		2	58.3	11.3	30.5
	3	13.4	20.9	65.7		3	53.7	10.4	35.8
2–6	1	18.2	11.5	70.3	2–14	1	27.7	29.7	42.6
	2	9.9	6.0	84.1		2	11.9	25.2	62.9
	3	9.0	1.5	89.6		3	13.4	14.9	71.6
2–7	1	14.9	20.9	64.2	2–15	1	38.5	25.0	36.5
	2	19.2	5.3	75.5		2	38.4	30.5	31.1
	3	16.4	6.0	77.6		3	34.3	17.9	47.8
2–8	1	3.4	17.6	79.1	2–16	1	50.0	21.6	28.4
	2	5.3	2.6	92.1		2	35.1	36.4	28.5
	3	6.0	4.5	89.6		3	47.8	22.4	29.9

TABLE XVI. Share of responses to all items of domain 3 that are in line with scientific views (+), not aligning with scientific views (−) or abstained form voting (∘), for all three cohorts.

Item	Cohort	−	∘	+	Item	Cohort	−	∘	+
3–1	1	0.0	14.2	85.8	3–9	1	35.1	20.9	43.9
	2	2.0	2.0	96.0		2	33.1	15.2	51.7
	3	1.5	0.0	98.5		3	35.8	10.4	53.7
3–2	1	48.6	14.2	37.2	3–10	1	17.6	32.4	50.0
	2	28.5	7.3	64.2		2	13.9	23.2	62.9
	3	10.4	6.0	83.6		3	7.5	17.9	74.6
3–3	1	6.8	25.0	68.2	3–11	1	18.2	10.1	71.6
	2	6.6	4.6	88.7		2	26.5	11.3	62.3
	3	4.5	0.0	95.5		3	14.9	9.0	76.1
3–4	1	43.9	18.9	37.2	3–12	1	25.0	14.2	60.8
	2	29.1	22.5	48.3		2	15.9	8.6	75.5
	3	37.3	13.4	49.3		3	11.9	6.0	82.1
3–5	1	25.0	14.9	60.1	3–13	1	35.8	20.9	43.2
	2	25.8	25.8	48.3		2	43.0	20.5	36.4
	3	19.4	11.9	68.7		3	34.3	11.9	53.7
3–6	1	14.2	21.6	64.2	3–14	1	25.7	28.4	43.2
	2	9.3	7.3	83.4		2	18.5	21.9	59.6
	3	1.5	3.0	95.5		3	11.9	14.9	73.1
3–7	1	39.9	10.8	49.3	3–15	1	43.9	18.9	37.2
	2	42.4	8.6	49.0		2	33.1	22.5	44.4
	3	28.4	4.5	67.2		3	26.9	11.9	61.2
3–8	1	0.0	22.3	77.7					
	2	6.0	2.0	92.1					
	3	1.5	1.5	97.0					

TABLE XVII. Share of responses to all items of domain 4 that are in line with scientific views (+), not aligning with scientific views (−) or abstained form voting (∘), for all three cohorts.

Item	Cohort	−	∘	+	Item	Cohort	−	∘	+
4–1	1	29.7	31.8	38.5	4–5	1	38.5	48.0	13.5
	2	30.5	37.1	32.5		2	28.5	45.0	26.5
	3	25.4	29.9	44.8		3	47.8	26.9	25.4
4–2	1	21.6	21.6	56.8	4–6	1	17.6	43.9	38.5
	2	13.9	43.0	43.0		2	16.6	44.4	39.1
	3	13.4	29.9	56.7		3	9.0	20.9	70.1
4–3	1	39.2	36.5	24.3	4–7	1	20.9	47.3	31.8
	2	30.5	48.3	21.2		2	20.5	48.3	31.1
	3	37.3	34.3	28.4		3	26.9	32.8	40.3
4–4	1	10.8	51.4	37.8	4–8	1	36.5	41.9	21.6
	2	5.3	45.7	49.0		2	36.5	55.0	11.9
	3	10.4	26.9	25.4		3	41.8	44.8	13.4

TABLE XVIII. Share of responses to all items of domain 5 that are in line with scientific views (+), not aligning with scientific views (−) or abstained form voting (∘), for all three cohorts.

Item	Cohort	−	∘	+	Item	Cohort	−	∘	+
5–1	1	20.3	10.8	68.9	5–5	1	15.5	14.4	69.6
	2	13.9	7.9	78.1		2	18.5	12.6	68.9
	3	4.5	7.5	88.1		3	11.9	7.5	80.6
5–2	1	14.2	10.8	75.0	5–6	1	38.5	22.3	39.2
	2	13.9	8.6	77.5		2	19.2	6.6	74.2
	3	9.0	7.5	83.6		3	9.0	4.5	86.6
5–3	1	10.8	18.2	70.9	5–7	1	45.9	29.1	25.0
	2	13.2	9.9	76.8		2	39.1	19.2	41.7
	3	4.5	7.5	88.1		3	25.4	19.4	55.2
5–4	1	14.9	21.6	63.5					
	2	30.5	15.9	53.6					
	3	19.4	10.4	70.1					

[1] H. T. Tran, P. Russo, and V. de Bakker, Effectiveness of a child-friendly astronomy news platform for science learning–an exploratory study, Eur. J. Sci. Math. Educ. **6**, 113 (2018).

[2] Raymond Davis Jr., Masatoshi Koshiba, and Riccardo Giacconi, Nobel Prize in Physics in 2002, Nobel Prize Outreach AB 2023.

[3] John C. Mather and George F. Smoot, Nobel Prize in Physics in 2006, Nobel Prize Outreach AB 2023.

[4] Saul Perlmutter, Brian P. Schmidt, and Adam G. Riess, Nobel Prize in Physics in 2011, Nobel Prize Outreach AB 2023.

[5] Takaaki Kajita and Arthur B. McDonald, Nobel Prize in Physics in 2015, Nobel Prize Outreach AB 2023.

[6] Rainer Weiss, Barry C. Barish, and Kip S. Thorne, Nobel Prize in Physics in 2017, Nobel Prize Outreach AB 2023.

[7] James Peebles, Michel Mayor, and Didier Queloz, Nobel Prize in Physics 2019, Nobel Prize Outreach AB 2023.

[8] Roger Penrose, Reinhard Genzel, and Andrea Ghez, Nobel Prize in Physics in 2020, Nobel Prize Outreach AB 2023.

[9] N. Holstermann and S. Bögeholz, Interesse von Jungen und Mädchen an naturwissenschaftlichen Themen am Ende der Sekundarstufe I [Gender-specific interests of adolescent learners in science topics], Z. Didakt. Naturwiss. **13**, 71 (2007).

[10] K.-H. Spatschek, *Astrophysik* (Springer, New York, 2003).

[11] J.-P. Meyn, Observation of planetary motion using a digital camera, Phys. Educ. **43**, 525 (2008).

[12] F. Hennig, M. Lipps, M. S. Ubben, and P. Bitzenbauer, From the big bang to life beyond earth: German preservice

physics teachers' conceptions of astronomy and the nature of science, Educ. Sci. **13**, 475 (2023).

[13] S. Salimpour, S. Bartlett, M. Fitzgerald *et al.*, The gateway science: A review of astronomy in the OECD school curricula, including China and South Africa, Res. Sci. Educ. **51**, 975 (2021).

[14] V. Bromm, Formation of the first stars, Rep. Prog. Phys. **76**, 112901 (2013).

[15] I. Hubeny, Stellar atmospheres theory: An introduction, in *Stellar Atmospheres: Theory and Observations* (Springer, New York, 2007), pp. 1–68.

[16] D. B. Fisher, A. D. Bolatto, R. Herrera-Camus, B. T. Draine, J. Donaldson, F. Walter, K. M. Sandstrom, A. K. Leroy, J. Cannon, and K. Gordon, The rarity of dust in metal-poor galaxies, Nature (London) **505**, 186 (2014).

[17] A. Lelliott and M. Rollnick, Big ideas: A review of astronomy education research 1974–2008, Int. J. Sci. Educ. **32**, 1771 (2010).

[18] J. Nussbaum, Children's conceptions of the earth as a cosmic body: A cross age study, Sci. Educ. **63**, 83 (1979).

[19] M. Kose, M. A. Kurtulus, and K. Bilen, How do gifted elementary students perceive earth's shape and gravity?, Pedagog. Res. **7**, em0126 (2022).

[20] L. Agan and C. Sneider, Learning about the earth's shape and gravity: A guide for teachers and curriculum developers, Astron. Educ. Rev. **2**, 90 (2004).

[21] C. I. Sneider and M. M. Ohadi, Unraveling students' misconceptions about the Earth's shape and gravity, Sci. Educ. **82**, 265 (1998).

[22] B. L. Jones, P. P. Lynch, and C. Reesink, Children's conceptions of the earth, sun and moon, Int. J. Sci. Educ. **9**, 43 (1987).

[23] S. Vosniadou and W. F. Brewer, Mental models of the earth: A study of conceptual change in childhood, Cogn. Psychol. **24**, 535 (1992).

[24] J. Nussbaum and J. Novak, An assessment of children's concepts of the earth utilizing structured interviews, Sci. Educ. **60**, 685 (1976).

[25] C. Sneider and S. Pulos, Children's cosmographies: Understanding the Earth's shape and gravity, Sci. Educ. **67**, 205 (1983).

[26] M. Ampartzaki and M. Kalogiannakis, Astronomy in early childhood education: A concept-based approach, Early Childhood Educ. J. **44**, 169 (2016).

[27] A. Chiras, Day/night cycle: Mental models of primary school children, Sci. Educ. Int. **19**, 65 (2008).

[28] S. Vosniadou and W. F. Brewer, Mental models of the day/night cycle, Cogn. Sci. **18**, 123 (1994).

[29] M. Fleer, A cross-cultural study of rural Australian aboriginal children's understanding of night and day, Res. Sci. Educ. **27**, 101 (1997).

[30] J. Baxter, Children's understanding of familiar astronomical events, Int. J. Sci. Educ. **11**, 502 (1989).

[31] R. K. Atwood and V. A. Atwood, Preservice elementary teachers' conceptions of the causes of seasons, J. Res. Sci. Teach. **33**, 553 (1996).

[32] L. B. Furuness and M. R. Cohen, Children's conceptions of the seasons: A comparison of three interview techniques, in *Proceedings of the 62nd Annual Meting of the National Association for Research in Science Teaching* (National Association for Research in Science Teaching, San Francisco, CA, 1989).

[33] R. Trumper, Teaching future teachers basic astronomy concepts–seasonal changes—at a time of reform in science education, J. Res. Sci. Teach. **43**, 879 (2006).

[34] C. Sneider, V. Bar, and C. Kavanagh, Learning about seasons: A guide for teachers and curriculum developers, Astron. Educ. Rev. **10**, 010103 (2011).

[35] J. Wilhelm, M. Cole, C. Cohen, and R. Lindell, How middle level science teachers visualize and translate motion, scale, and geometric space of the Earth-Moon-Sun system with their students, Phys. Rev. Phys. Educ. Res. **14**, 010150 (2018).

[36] H. Küçüközer, M. E. Korkusuz, H. A. Küçüközer, and K. Yürümezoğlu, The effect of 3D computer modeling and observation-based instruction on the conceptual change regarding basic concepts of astronomy in elementary school students, Astron. Educ. Rev. **8**, 010104 (2009).

[37] A. Christonasis, G. Stylos, T. Chatzimitakos, A. Kasouni, and K. Kotsis, Religiosity and teachers' acceptance of the big bang theory, Eurasian J. Sci. Environ. Educ **3**, 25 (2023).

[38] S. Aretz, A. Borowski, and S. Schmeling, A fairytale creation or the beginning of everything: Students' pre-instructional conceptions about the big bang theory, Persp. Sci. **10**, 46 (2016).

[39] L. E. Trouille, K. Coble, G. L. Cochran, J. M. Bailey, C. T. Camarillo, M. D. Nickerson, and L. R. Cominsky, Investigating student ideas about cosmology III: Big bang theory, expansion, age, and history of the universe, Astron. Educ. Rev. **12**, 010110 (2013).

[40] M. S. Ubben, J. Hartmann, and A. Pusch, Holes in the atmosphere of the universe: An empirical qualitative study on mental models of students regarding black holes, Astron. Educ. Res. **2**, 1 (2022).

[41] J. M. Bailey and T. F. Slater, A review of astronomy education research, Astron. Educ. Rev. **2**, 20 (2003).

[42] J. M. Bailey, E. E. Prather, B. Johnson, and T. F. Slater, College students' preinstructional ideas about stars and star formation, Astron. Educ. Rev. **8**, 010110 (2009).

[43] V. M. Rajpaul, C. Lindstrøm, M. C. Engel, M. Brendehaug, and S. Allie, Cross-sectional study of students' knowledge of sizes and distances of astronomical objects, Phys. Rev. Phys. Educ. Res. **14**, 020108 (2018).

[44] S. Salimpour and M. T. Fitzgerald, The cosmic interaction: A review of the literature on cosmology, religion, and the big questions in the context of astronomy education research, Sci. Educ. **31**, 819 (2022).

[45] F. Gali, Secondary school children's understanding of basic astronomy concepts, J. Stud. Soc. Sci. Hum. **7**, 328 (2021).

[46] P. M. Sadler, *The Initial Knowledge State of High School Astronomy Students* (Harvard University, Cambridge, MA, 1992).

[47] R. Trumper, University students' conceptions of basic astronomy concepts, Phys. Educ. **35**, 9 (2000).

[48] M. L. Berendsen, Conceptual astronomy knowledge among amateur astronomers, Astron. Educ. Rev. **4**, 1 (2005).

[49] C. S. Wallace, E. E. Prather, and D. K. Duncan, A study of general education astronomy students' understandings of cosmology. Part I. Development and validation of four conceptual cosmology surveys, Astron. Educ. Rev. **10**, 010106 (2011).

[50] R. Trumper, The need for change in elementary school teacher training—a cross-college age study of future teachers' conceptions of basic astronomy concepts, Teach. Teach. Educ. **19**, 309 (2003).

[51] H. Kalkan and K. Kiroglu, Science and nonscience students' ideas about basic astronomy concepts in preservice training for elementary school teachers, Astron. Educ. Rev. **6**, 15 (2007).

[52] J. D. Plummer and A. Tanis Ozcelik, Preservice teachers developing coherent inquiry investigations in elementary astronomy, Sci. Educ. **99**, 932 (2015).

[53] B. Bektaşli, In-service science teachers' astronomy misconceptions, Mediterr. J. Educ. Res. **15**, 1 (2014).

[54] A. S. Arslan and U. Durikan, Pre-service teachers' mental models of basic astronomy concepts, Sci. Educ. Int. **27**, 88 (2016).

[55] R. S. Lindell and J. P. Olsen, Developing the lunar phases concept inventory, in *Proceedings of the 2002 Physics Education Research Conference* (PERC Publishing, New York, 2002).

[56] P. Chastenay and M. Riopel, Development and validation of the moon phases concept inventory for middle school, Phys. Rev. Phys. Educ. Res. **16**, 020107 (2020).

[57] E. M. Bardar, E. E. Prather, K. Brecher, and T. F. Slater, The need for a light and spectroscopy concept inventory for assessing innovations in introductory astronomy survey courses, Astron. Educ. Rev. **4**, 20 (2006).

[58] E. M. Bardar, E. E. Prather, K. Brecher, and T. F. Slater, Development and validation of the light and spectroscopy concept inventory, Astron. Educ. Rev. **5**, 103 (2007).

[59] B. Hufnagel, T. Slater, G. Deming, J. Adams, R. L. Adrian, C. Brick, and M. Zeilik, Pre-course results from the astronomy diagnostic test, Pub. Astron. Soc. Aust. **17**, 152 (2000).

[60] M. Zeilik, Birth of the astronomy diagnostic test: Prototest evolution, Astron. Educ. Rev. **1**, 46 (2002).

[61] B. Hufnagel, Development of the astronomy diagnostic test, Astron. Educ. Rev. **1**, 47 (2002).

[62] S. J. Slater *et al.*, The development and validation of the test of astronomy standards (toast), J. Astron. Earth Sci. Educ. **1**, 1 (2014).

[63] P. M. Sadler, H. Coyle, J. L. Miller, N. Cook-Smith, M. Dussault, and R. R. Gould, The Astronomy and Space Science Concept Inventory: Development and validation of assessment instruments aligned with the K–12 National Science Standards, Astron. Educ. Rev. **8**, 010111-1 (2010).

[64] J. M. Bailey, B. Johnson, E. E. Prather, and T. F. Slater, Development and validation of the star properties concept inventory, Int. J. Sci. Educ. **34**, 2257 (2012).

[65] P. Bitzenbauer and M. Ubben, Entwicklung eines Konzepttests zur Astronomie—Erste Ergebnisse, PhyDid B **1**, 29 (2022).

[66] S. Salimpour, R. Tytler, B. Doig, M. T. Fitzgerald, and U. Eriksson, Conceptualising the Cosmos: Development and validation of the Cosmology Concept Inventory for High School, Int. J. Sci. Math. Educ. **21**, 251 (2023).

[67] L. Agan, Stellar ideas: Exploring students' understanding of stars, Astron. Educ. Rev. **3**, 77 (2004).

[68] J. G. Sharp, Children's astronomical beliefs: A preliminary study of year 6 children in South-West England, Int. J. Sci. Educ. **18**, 685 (1996).

[69] J. G. Sharp, R. Bowker, and J. Merrick, Primary astronomy: Conceptual change and learning in three 10–11 year olds, Res. Educ. **57**, 67 (1997).

[70] S. J. Slater, S. P. Schleigh, D. J. Stork *et al.*, Analysis of individual test of astronomy standards (TOAST) item responses, J. Astron. Earth Sci. **2**, 89 (2015).

[71] J. Dove, Does the man in the moon ever sleep? An analysis of student answers about simple astronomical events: A case study, Int. J. Sci. Educ. **24**, 823 (2002).

[72] J. D. Plummer, K. D. Wasko, and C. Slagle, Children learning to explain daily celestial motion: Understanding astronomy across moving frames of reference, Int. J. Sci. Educ. **33**, 1963 (2011).

[73] S. Anantasook and C. Yuenyong, Thai student existing understanding about the solar system model and the motion of the stars, AIP Conf. Proc. **1923**, 030058 (2018).

[74] A. D. Lelliott, Learning about astronomy: A case study exploring how grade 7 and 8 students experience sites of informal learning in South Africa, Ph.D. thesis, University of the Witwatersrand, Johannesburg, 2007.

[75] J. Dunlop, How children observe the universe, Pub. Astron. Soc. Aust. **17**, 194 (2000).

[76] J. D. Plummer, A cross-age study of children's knowledge of apparent celestial motion, Int. J. Sci. Educ. **31**, 1571 (2009).

[77] T. Hannust and E. Kikas, Children's knowledge of astronomy and its change in the course of learning, Earl. Child. Res. Q. **22**, 89 (2007).

[78] D. A. Sturman and B. Moghaddam, The neurobiology of adolescence: Changes in brain architecture functional dynamics, and behavioral tendencies, Neurosci. Biobehav. Rev. **35**, 1704 (2011), passing the Knife Edge in Adolescence: Brain Pruning and Specification of Individual Lines of Development.

[79] R. Müller and H. Wiesner, Teaching quantum mechanics on an introductory level, Am. J. Phys. **70**, 200 (2002).

[80] B. Reinisch and M. Krell, Assessing pre-service teachers' views of scientists, their activities, and locations: The VoSAL instrument, Res. Sci. Educ. **53**, 139 (2023).

[81] J. M. Veith, P. Bitzenbauer, and B. Girnat, Assessing learners' conceptual understanding of introductory group theory using the CI^{2}GT: Development and analysis of a concept inventory, Educ. Sci. **12**, 376 (2022).

[82] G. Rickheit and L. Sichelschmidt, Mental models: Some questions, some answers, some suggestions, in *Mental Models in Discourse Processing and Reasoning* (Elsevier, Amsterdam, Netherlands, 1999).

[83] A. A. diSessa, Ontologies in pieces: Response to Chi and Slotta, Cognit. Instr. **10**, 272 (1993).

[84] A. A. diSessa, What do "just plain folk" know about physics?, *The Handbook of Education and Human Development: New Models of Learning, Teaching and Schooling* (Wiley, Hoboken, NJ, 1998), p. 681.

[85] T. Amin, C. L. Smith, and M. Wiser, Student conceptions and conceptual change: Three overlapping phases of research, in *Handbook of Research on Science Education* (Routledge, London, 2014), Vol. II, pp. 71–95.

[86] A. Hanslmeier, *Einführung in Astronomie und Astrophysik* (Springer, New York, 2002), Vol. 2.

[87] M. Scholz, *Die Physik der Sterne: Aufbau, Entwicklung und Eigenschaften* (Springer-Verlag, Berlin, 2018).

[88] J. P. Adams and T. F. Slater, Astronomy in the national science education standards, J. Geosci. Educ. **48**, 39 (2000).

[89] S. Vosniadou, Designing curricula for conceptual restructuring: Lessons from the study of knowledge acquisition in astronomy, Center for the Study of Reading, Champaign, IL, Technical Report No. 546, 1992.

[90] S. Galano, A. Colantonio, S. Leccia, E. Pudu, I. Marzoli, and I. Testa, A teaching-learning module on stellar structure and evolution, EPJ Web Conf. **200**, 01017 (2019).

[91] A. Colantonio, S. Galano, S. Leccia, E. Pudu, and I. Testa, Design and development of a learning progression about stellar structure and evolution, Phys. Rev. Phys. Educ. Res. **14**, 010143 (2018).

[92] J. M. Bailey, Development of a concept inventory to assess students' understanding and reasoning difficulties about the properties and formation of stars, Ph.D. thesis, The University of Arizona, 2006.

[93] T. J. Kline, *Psychological Testing: A Practical Approach to Design and Evaluation* (Sage Publications, London, 2005).

[94] N. Jorion, B. D. Gane, K. James, L. Schroeder, L. V. DiBello, and J. W. Pellegrino, An analytic framework for evaluating the validity of concept inventory claims, J. Eng. Educ. **104**, 454 (2015).

[95] K. S. Taber, The use of Cronbach's alpha when developing and reporting research instruments in science education, Res. Sci. Educ. **48**, 1273 (2018).

[96] M. Calik, B. Turan, and R. K. Coll, A cross-age study of elementary student teachers' scientific habits of mind concerning socioscientific issues, Int. J. Sci. Math. Educ. **12**, 1315 (2014).

[97] R. Trumper, Children's energy concepts: A cross-age study, Int. J. Sci. Educ. **15**, 139 (1993).

[98] D. Raufelder and S. Kulakow, The role of the learning environment in adolescents' motivational development, Motiv. Emot. **45**, 299 (2021).

[99] P. H. Taskinen, K. Schütte, and M. Prenzel, Adolescents' motivation to select an academic science-related career: The role of school factors individual interest, and science self-concept, Educ. Res. Eval. **19**, 717 (2013).

[100] J. Fox and M. Cater, An analysis of adolescents' science interest and competence in programs with and without a competitive component, J. Agric. Educ. **56**, 90 (2015).

[101] J. Cohen, *Statistical Power Analysis for the Behavioral Sciences* (Lawrence Erlbaum Associates, Hillsdale, NJ, 1988).

[102] H. Levene, *Robust Tests for Equality of Variances* (Stanford University Press, Stanford, 1960).

[103] S. S. Shapiro and M. B. Wilk, An analysis of variance test for normality (complete samples), Biometrika **52**, 591 (1965).

[104] N. F. Comins, *Heavenly Errors: Misconceptions About the Real Nature of the Universe* (Columbia University Press, New York, 2001).

[105] W. C. Philips, Earth science misconceptions, Sci. Teach. **58**, 21 (1991).

[106] S. Salimpour, R. Tytler, M. T. Fitzgerald, and U. Eriksson, Is the Universe infinite? Characterising a hierarchy of reasoning in student conceptions of cosmology concepts using open-ended surveys, J. STEM Educ. Res. **6**, 102 (2023).

[107] R. Trumper, A cross-age study of senior high school students' conceptions of basic astronomy concepts, Res. Sci. Technol. Educ. **19**, 97 (2001).

[108] R. Trumper, A cross-age study of junior high school students' conceptions of basic astronomy concepts, Int. J. Sci. Educ. **23**, 1111 (2001).

[109] A. S. Reber, Implicit learning and tacit knowledge, J. Exp. Psychol. **118**, 219 (1989).

[110] A. Lelliott, The concept of spatial scale in astronomy addressed by an informal learning environment, Afr. J. Res. Math. Sci. Technol. Educ. **14**, 20 (2010).

[111] J. Y.-M. Wang, Implicit learning in science: Activating and suppressing scientific intuitions to enhance conceptual change, Ph.D. thesis, University of Minnesota, 2018.

[112] L. Andre, T. Durksen, and M. L. Volman, Museums as avenues of learning for children: A decade of research, Learn. Environ. Res. **20**, 47 (2017).

[113] M. Behrendt and T. Franklin, A review of research on school field trips and their value in education, Int. J. Environ. Sci. Educ. **9**, 235 (2014).

[114] C. Kulgemeyer, A framework of effective science explanation videos informed by criteria for instructional explanations, Res. Sci. Educ. **50**, 2441 (2020).

[115] R. Trumper, A cross-college age study of science and nonscience students' conceptions of basic astronomy concepts in preservice training for high-school teachers, J. Sci. Educ. Technol. **10**, 189 (2001).

[116] L. Wallner and K. E. Barajas, Using comics and graphic novels in k-9 education: An integrative research review, Stud. Comics **11**, 37 (2020).

[117] N. Rutten, W. R. Van Joolingen, and J. T. Van Der Veen, The learning effects of computer simulations in science education, Comput. Educ. **58**, 136 (2012).

[118] J. Desrosiers, R. Hebert, G. Bravo, and A. Rochette, Comparison of cross-sectional and longitudinal designs in the study of aging of upper extremity performance, J. Gerontol. Ser. A **53**, B362 (1998).

[119] E. Deeb, L. M. Rebull, D. V. Black, J. Gibbs, and E. Larsen, Nitarp: Measuring the effectiveness of an authentic research experience in secondary astronomy education through concept mapping, in *American Astronomical Society Meeting Abstracts# 225* (2015), Vol. 225.

[120] E. Bizimana, D. Mutangana, and A. R. Mwesigye, Fostering students' retention in photosynthesis using concept

mapping and cooperative mastery learning instructional strategies, Eur. J. Educ. Res. **11**, 103 (2022).
[121] F. Malatjie and F. Machaba, Exploring mathematics learners' conceptual understanding of coordinates and transformation geometry through concept mapping, Eurasia J. Math. Sci. Technol. Educ. **15**, em1818 (2019).
[122] B. Mainali, Investigating pre-service teachers' beliefs towards mathematics: A case study, Eur. J. Sci. Math. Educ. **10**, 412 (2022).
[123] J. D. Plummer, Early elementary students' development of astronomy concepts in the planetarium, J. Res. Sci. Teach. **46**, 192 (2009).