Epistemic belief structures within introductory astronomy

Keith Johnson* and Shannon D. Willoughby
Montana State University, Bozeman, 59717-3840 Montana, USA

(Received 20 February 2018; published 7 June 2018)

The reliability and validity of inventories should be verified in multiple ways. Although the epistemological beliefs about the physical science survey (EBAPS) has been deemed to be reliable and valid by the authors, the axes or factor structure proposed by the authors has not been independently checked. Using data from a study sample we discussed in previous publications, we performed exploratory factor analysis on 1,258 post-test EBAPS surveys. The students in the sample were from an introductory Astronomy course at a mid-sized western university. Inspection suggested the use of either a three-factor model or a five-factor model. Each of the factors is interpreted and discussed, and the factors are compared to the axes proposed by the authors of the EBAPS. We find that the five-factor model extrapolated from our data partially overlaps with the model put forth by the authors of the EBAPS, and that many of the questions did not load onto any factors.

DOI: 10.1103/PhysRevPhysEducRes.14.010135

I. INTRODUCTION

This paper is a continued exploration of student epistemological beliefs within introductory astronomy as presented in previous literature [1,2]. The goal is to identify and analyze response patterns of students in an introductory astronomy class using exploratory factor analysis on data from the epistemological beliefs assessment for physical science (EBAPS). We believe the findings will help further understanding of student epistemologies in introductory astronomy, an area where literature is lacking. This analysis also provides some insight into the EBAPS instrument itself. Fundamentally, we seek to answer the following research question: What factor structure is present for astronomy students who take the EBAPS and what do these factors represent?

The EBAPS suffers from a distinct lack of validity studies within literature, as others have also noted, despite its prominence as one of the more well-known epistemic instruments within physics education research [3–7]. It is important that instruments as influential as the EBAPS experience multiple validation studies so as to further establish the legitimacy of what the instrument claims to assess [8–10]. We do not claim that this work is a complete and thorough validation study of the EBAPS (validation of this assessment is expanded upon in Sec. II A), but rather we seek to gain insight into how students are responding to the assessment by exploring the constructs present within the instrument and how they relate to the constructs as proposed by the authors of the assessment.

The analyses done herein, which utilize five years of post-test data from an introductory astronomy course, will begin to explore the psychometric properties of the EBAPS. The factor structure revealed by the exploratory factor analysis done will be further refined using written student responses of items belonging to these factors. Last, the implications of these findings will be discussed as will future work regarding continued validation work and the EBAPS.

*keith.johnson12@montana.edu

II. METHODS

A. EBAPS

The epistemological beliefs assessment for physical science (EBAPS) is a 30-item forced-choice instrument designed to assess students' epistemologies within the physical sciences. The EBAPS was developed by a research group (Andrew Elby, John Frederiksen, Christina Schwarz, and Barbara White) from University of California, Berkeley. The instrument is designed to focus attention on personal epistemologies as opposed to expectations [11]. The authors claim the EBAPS assesses five nonorthogonal dimensions of epistemological beliefs: Structure of scientific knowledge, nature of knowing and learning, real-life applicability, evolving knowledge, and source of ability to learn. Precise definitions of these axes have been provided in Table IX in the Appendix. Additional information regarding the EBAPS may be found on the host website [12]. Each item on the EBAPS is scored on a scale from 0 to 4, where 0 represents an unsophisticated view and 4 represents a more sophisticated or more expertlike view. Questions are mapped onto one of the five axes defined by the authors of the instrument. Axis scores represent an average of all item scores belonging to an axis; which items correspond to a particular axis may also be seen in Table IX.

The authors of the EBAPS explicitly state that they "[…] don't want to assume that each subscale corresponds to a stable, consistent belief (or set of beliefs)" and provide compelling reasoning as to why they take this stance [11]. In essence, the authors of the EBAPS follow a cognitive theory which consists of "fine-grained cognitive resources" which are activated upon context, similar to DiSessa's p prims [13]. What this means is that items that belong to the same subscale, category, or axis may be answered differently depending on particular contextual cues, but in a general sense should still be placed within a certain subscale. It may be worth wondering why bother promoting axes at all within the instrument. However, experts may still provide a general category for which these items should belong; this categorization along axes is an approach many epistemic instruments within physics education research follow. Whether the constructs (beliefs) that are found in this study represent a contextual activation of local cognitive resources or they represent fundamental beliefs with varying levels of sophistication, they may be used to help better understand and address student epistemologies within introductory astronomy.

B. Context and participants

The EBAPS was given as a pretest and post-test in a large enrollment introductory astronomy course at a medium-sized, midwestern, land-grant institution. The course had active learning elements and typically consisted of a 45 minute lecture with a 30 minute group activity. Lecture itself made frequent use of iClickers, writing activities, and student-student interactions while group activities made use of the tutorials for introductory astronomy to supplement lecture [14]. The data in this study were taken from fall of 2012 through spring of 2017 across two identical sections of the course, 2334 students in all. Although precise information of classroom demographics over this time could not be acquired, in fall of 2017 the university as a whole had an ethnic distribution which was 83% Caucasian, 4% Hispanic, 2% American Indian, and 1% Asian. The remaining 10% were African American, Pacific Islander, or a combination of multiple ethnicities. Roughly 47% of students were female and 53% were male. The instructor notes that each section consisted of nearly 200 students a semester and was predominantly Caucasian with an approximately equal distribution of males and females. The same instructor (one of the authors) taught the course across both sections and all five years of the study.

C. Exploratory factor analysis

Factor analysis is a statistical method that transforms a set of variables into factors whose items contain some type of underlying similarity. In general, there are three reasons for conducting an exploratory factor analysis (EFA). The first is to discover the underlying structure responsible for guiding student responses. This is what the current paper will focus on, as the underlying epistemic belief structures influencing student responses to the EBAPS are sought. The second is to reduce a set of data into simplified components. For instance, if a set of variables are being utilized to measure a trait such as intelligence, there could arise a scenario in which several of these variables have strong linear relationships with one another. Here the second method, typically known as principal components analysis, will reduce these variables of high multicollinearity into a single predictor. Last, EFA may be used in the creation of an instrument, assisting the researcher in determining whether the variables put forth are measuring the latent variable they intend to measure. It is not unusual for an EFA to be followed up with a confirmatory factor analysis (CFA). Briefly, a CFA is a type of factor analysis that measures the consistency of any factors that exist within a model proposed by the researchers. A CFA may be used on the same data as an EFA to help refine items and/or item relationships. Alternatively, a CFA may be used on a separate set of data as an EFA to determine how well the proposed model fits with alternate data.

The authors of the EBAPS have attributed values (between 0 and 4) to each question within the instrument. EFA traditionally utilizes the Pearson correlation matrix, which is based on the assumption that the data exhibit an equal interval scale with linear relationships between items. However, many items on the instrument have a Likert-style response and as such cannot necessarily be assumed to exhibit these traits [15]. Violations of this assumption (i.e., Likert-style data) could potentially result in Pearson correlations that underestimate the strength of correlations between items thereby leading to improper factors and/or factor loadings [16,17]. Alternative options could include the use of polychoric correlations with EFA. Polychoric correlations are specifically designed to estimate correlations for ordinal data from a bivariate normally distributed population. In the end, it was decided to implement traditional EFA but with the principal axis factoring extraction methodology, which makes no assumptions of the underlying distribution of the data [18]. This same approach to EFA has been utilized on other epistemic instruments within the sciences as well [19].

Ideally, if proper factors underlying student epistemic beliefs in astronomy can be identified, then future work may focus on addressing these traits within the classroom, either for improvement or refinement. The EBAPS itself may benefit from this work as well, given that more data regarding response behavior to the instrument will be acquired.

D. Data analysis

The following section is composed of two parts, first determining the number of factors to retain, and then assessing the stability of the factors. All analyses within were done utilizing IBM SPSS software. This provided a well-vetted template from which to conduct work and allowed for easy partitioning and thorough exploration of the data.

1. Number of factors to be retained

Parallel analysis and scree plots were generated to assist in determining the number of possible factors involved. Also utilized was an EFA with the Kaiser criterion method, which states that factors having eigenvalues greater than or equal to 1 are to be retained [20]. We proceed by discussing the results of all three methods. The methods below make use of the post-test data of students from section one, which had 1258 students across five years. Section two will later be utilized in future work with confirmatory factor analysis.

A scree plot takes the factors from a correlation matrix of assessment items and plots them against their corresponding eigenvalues. In a scree plot it is typical that the eigenvalues associated with most factors tend to lie along a single line. To determine which factors to retain, look for when the slope of this line makes a distinct change and keep the factors whose eigenvalues rise above that line. The scree plot generated from the 30 EBAPS items, seen in Fig. 1, indicates three prominent factors with upwards of six possible factors (seven if the inflection point is included) [21].

Parallel analysis involves random generation of pseudoclasses (based off of raw data input) from which a correlation matrix is created and its factor eigenvalues calculated. These eigenvalues are thus values arising from random noise within the raw data. As such, only the eigenvalues from the factor analysis that are larger than these noise-based eigenvalues are not likely to be artificial constructs created from deviations in data. The parallel

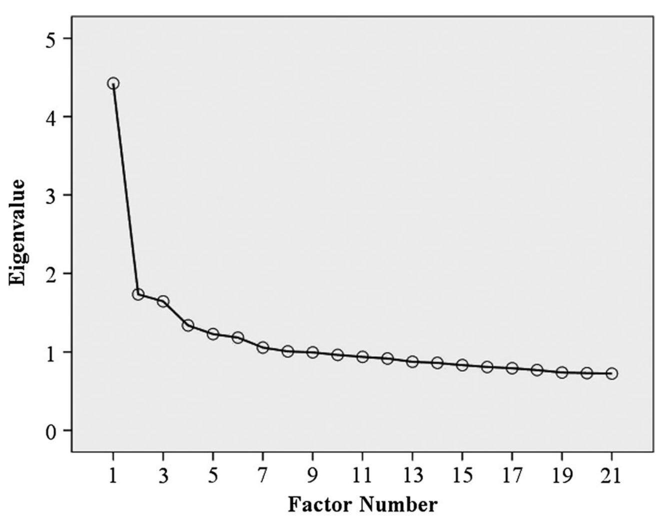

FIG. 1. Scree plot for the 30 EBAPS items (post-test). Factors after 21 were truncated.

analysis was done using an add-on to SPSS that can be obtained online [22,23]. Eigenvalues generated in the analysis were done with permutations of the raw 1258-student data set using Castellan's algorithm [24]. The results of the parallel analysis were similar to that of the scree plot, revealing three to five prominent factors, but up to seven possible factors overall.

In the final attempt to determine the number of axes to retain, an exploratory factor analysis using the Kaiser criterion was performed with all students from section one and using all thirty items of the EBAPS. In doing so, this analysis did not restrict the number of factors that could be present. This method also did not rely purely on the number of eigenvalues greater than unity to determine factors, but instead looked at how many eigenvalues were over unity after other filtering criteria (detailed later) were applied. The EFA implemented the direct oblimin oblique rotation method. An oblique rotation was utilized as we expect some general correlation of epistemological constructs [18]. Other often-used oblique rotations include the quartermin and promax methods, although there is reason to believe that all these oblique rotations lead to a similar outcome [18].

Principal axis factoring was utilized instead of principal components analysis (PCA) as we sought to find underlying relationships between items and avoid distributional assumptions, as opposed to a PCA, which is utilized to reduce data into simplified minimal components [25–28]. In conducting the EFA the determinant of the correlation matrix for this data was nonzero (determinant $= 0.026$), suggesting that linear combinations of data (i.e., factors) are possible. Furthermore, Bartlett's test of sphericity indicated that this determinant is indeed significantly different from zero ($p < 0.001$). The Kaiser-Meyer-Olkin (KMO) test revealed that the degree of common variance among the items was acceptable as Meritorious (KMO $= 0.846$) [29]. Last, the diagonal elements of the anti-image correlation matrix were all above 0.50, a sign that these items may be suitably explained by underlying factors [30]. Some issues did arise when parsing the correlation matrix, however. Ideally an item should have at least one correlation exceeding 0.30, however, several items within the data did not reach this cutoff [25].

Upon analyzing the correlation matrix, we chose to remove items which did not have a correlation coefficient greater than 0.2. This value was chosen instead of 0.3 because most items either exhibited at least one correlation greater than 0.2 or were notably far beneath this measure. The items 13, 18, 19, 20, 23, 24, 26, and 29 were removed from the factor analysis during this process, and the analysis was redone. Ideally all items in a survey can be kept, however, it is common for some items to simply not group strongly enough to indicate the presence of a prominent latent variable. Item 2 was removed in this second iteration when it was found that it had a loading less than 0.32 [25]. Again the analysis was redone and at this point the communalities of the items were inspected. Items whose communality was less than 0.20 were removed one at a time, with the lowest value being removed first, redoing the factor analysis after each item was removed [31]. This led to the removal of items 3, 7, 15, 17, and 27. The pattern matrix for the final iteration of this exploratory factor analysis is displayed in Table I; also provided within Table II are the corresponding factor correlations for the analysis. As can be seen, there are four primary factors left and none of the fifteen items exhibited a crossloading greater than 0.40 [32]. All but one factor retained at least three items which is a recommended criteria for factor stability [33]. This final factoring also upheld all basic assumptions, namely, a KMO $= 0.845$, and Bartlett's test of sphericity having $p < 0.001$.

Given the findings for the Kaiser criterion factor analysis, scree plots, and parallel analysis, it has been decided that both a three-factor model and a five-factor model are to be presented. We neglect formally discussing this potential four-factor Kaiser criteria-based model because the five-factor model does a satisfactory job of retaining and yet adding on to the information provided by the four-factor model. A six-factor model was also considered, however, some compromises to commonality criteria as well as crossloading issues with item 12 led to the dismissal of the model. Beyond that, models of seven and eight factors were explored but it was found that loading and cross-loading issues became too frequent to retain viable independent factors. It is possible that future work will reconsider a four-factor and six-factor model with confirmatory factor analysis. Future CFA work will also consider items that were removed from their original factors due to the strict filtering criteria applied here.

TABLE I. Factor loadings from EFA utilizing the Kaiser criterion. Loadings less than 0.3 have been suppressed. EFA involved all section 1 students. These factors represent factors whose eigenvalues are over unity after other filtering criteria were applied to the original 30 items.

Items	Factor 1	Factor 2	Factor 3	Factor 4
14	0.583			
08	0.556			
06	0.467			
10	0.442			
04	0.430			
22		0.719		
21		0.603		
25		0.415		
30			0.593	
28			0.417	
16				0.558
05				0.535
11				0.513
12				0.437
09				0.353

TABLE II. Factor correlation matrix results for EFA utilizing the Kaiser criterion. EFA involved all section 1 students. These factors represent factors whose eigenvalues are over unity after other filtering criteria were applied to the original 30 items.

Factors	1	2	3	4
1	1.000	...	...	...
2	0.310	1.000	...	...
3	0.451	0.347	1.000	...
4	0.619	0.206	0.313	1.000

2. Factor stability

Having decided the number of factors to retain, an EFA comparison was conducted for both a three-factor model and a five-factor model. To test item stability, students from the above section one data were split into two roughly equal groups for both the three-factor model and five-factor model; the groups were then compared to one another [34]. The split itself was based alphabetically by last name, as there is no reason to believe that a strong correlation exists between last name and epistemic beliefs within this demographic. The number of students from each group is not consistent, as items between groups and models are not similar and SPSS does not consider students with incomplete data for an item. Roughly, the first group (group A) contained around 580 unique students and the second group (group B) contained around 630 unique students. The approach in factor simplification within these analyses utilized a similar methodology as was done for the Kaiser criteria factor analysis above except the number of factors were simply restricted in SPSS to be 3 and 5, accordingly. It is worth noting that some items removed via this process may again be considered during CFA model refinement. For now, we are utilizing this process to provide the most mathematically sensible model possible with EFA.

Three-factor model.—The initial results for the three-factor models of both groups A and B are shown in Tables III and IV, respectively. A direct comparison between these models reveals similar factors with mostly consistent item loadings. As we seek a dependable theoretical model, dissimilar items were removed. This resulted in items 2 and 4 being removed from factor 1 in group A while items 28 and 30 were removed from factor 3 in group B. Similarly, item 3 was also removed from factor 3 in group A. An EFA was redone for each group utilizing only the remaining common items, the results of which are shown in Table X (group A) and Table XI (group B) within the Appendix. We thus propose a theoretical three-factor model shown in Table V.

Five-factor model.—Table VI (group A) and Table VII (group B) reveal the results for each group within the

TABLE III. Initial three-factor loadings for group A. Loadings less than 0.25 have been suppressed. EFA involved approximately half of the section 1 students. These loadings represent results after a filtering of the original 30 items.

Items	Factor 1	Factor 2	Factor 3
08	0.780		
14	0.564		
02	−0.466		
06	0.465		
10	0.423		
04	0.390		
22		0.737	
21		0.617	
25		0.405	
11			0.556
05			0.539
03			0.477
16	0.262		0.411
12			0.405
09			0.356

TABLE IV. Initial three-factor loadings for group B. Loadings less than 0.25 have been suppressed. EFA involved approximately half of the section 1 students. These loadings represent results after a filtering of the original 30 items.

Items	Factor 1	Factor 2	Factor 3
16	0.610		
09	0.505		
12	0.487		
05	0.464		
11	0.433		
21		0.642	
22		0.603	
25		0.465	
08	0.258		0.465
30			0.452
28			0.450
14			0.407
10			0.398
06			0.377

TABLE V. Theoretical three-factor model.

Factor 1	Factor 2	Factor 3
Question 06	Question 21	Question 05
Question 08	Question 22	Question 09
Question 10	Question 25	Question 11
Question 14		Question 12
		Question 16

TABLE VI. Initial five-factor loadings for group A. Loadings less than 0.25 have been suppressed. EFA involved approximately half of the section 1 students. These loadings represent results after a filtering of the original 30 items.

Items	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
08	0.721				
02	−0.554				
06	0.471				
14	0.430				
10	0.324				
22		0.743			
21		0.651			
25		0.359			
11			0.615		
05	0.250		0.414		
03			0.386		
01				0.669	
12			0.354	0.354	
09					−0.661
04					−0.389
16			0.272		−0.284

five-factor fitting. As can be seen, there are modest discrepancies between both results. However, prominent factors should be similar between the three-factor theoretical model and these five-factor theoretical models. As a result, each group model was reworked. Group A had item 1, 2, and 3 removed as these items were not present in the theoretical three-factor model, nor are items 1 and 2 present in the five-factor model of group B. Item 16 was added to group A as were items 28 and 30 due to their presence in both group B and in their previous association with the three-factor model of group A (they were initially removed due to weak correlations, $r < 0.25$). For similar reasons, item 10 was added to group B while item 7 was removed. Although these were not minor changes, results indicate relative agreement between the refined 5-factor models of group A (Table XII) and group B (Table XIII) as seen in the Appendix. Utilizing these findings, a five-factor theoretical model is hence put forth in Table VIII. Despite inconsistencies between the loading of items 4 and 9, as well as the concerning presence of factors with fewer than three items, we will discuss shortly how student responses have provided insight as to why items 4 and 9 may represent some form of underlying belief.

It is important to consider all models presented within this section as future CFA work may reveal that one particular model is more robust when applied to any alternative student data. Ideally, the five-factor model would have adequate fit statistics in CFA and be comparable to the three and four-factor models. Should that be the scenario, the five-factor model will be favored as it details a larger swath of the collective students' epistemic belief structure.

TABLE VII. Initial five-factor loadings for group B. Loadings less than 0.25 have been suppressed. EFA involved approximately half of the section 1 students. These loadings represent results after a filtering of the original 30 items.

Items	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
16	0.594				
05	0.523				
09	0.471		0.271		
11	0.464				
12	0.445				
03	0.279				
22		0.735			
21		0.573			
25		0.461			
04			0.542		
30				0.590	
28				0.367	
06					0.539
07					0.415
08					0.404
14					0.376
10					0.336

III. RESULTS

To assist in the interpretation of the proposed factors presented in Tables V and VIII, written student responses were utilized. In the latter years of the study students had access to an online homework site called Sapling Learning [35]. As such, EBAPS questions were able to be presented to the students in such a manner that they could type an open-ended response to the question asked. Their response was voluntary and no credit was awarded for participation. From these student responses, information pertaining to all factor analysis questions in both models was able to be obtained (with the exception of question 4). The responses, which typically numbered between 40 and 90 responses per question, were analyzed and partitioned for various patterns. It is advised that the reader utilize the EBAPS alongside the discussions, as there are frequent interpretations of the questions themselves. The discussion will predominantly involve an analysis of the five-factor theoretical model, as it has many similarities with the three-factor model while containing more information. That is, both models have three nearly identical factors which test the same latent variable, but the five-factor model also includes an additional two factors.

TABLE VIII. Theoretical five-factor model.

Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Question 06	Question 21	Question 05	Question 28	Question 04
Question 08	Question 22	Question 11	Question 30	Question 09
Question 10	Question 25	Question 12		
Question 14		Question 16		

A. Theoretical five-factor model

1. Factor 1: Structure of science

The first factor in the theoretical five-factor model includes EBAPS questions 6, 8, 10, and 14. The general theme seen in student responses here may be classified as *network of science* or *structure of science*. The majority of written responses to question 14 involve the idea that science permeates everyday life and is thus a necessary field in which politicians should be familiar. Question 8 written responses discuss the belief that theories act as a webbing which connects and accounts for information. These theories are thus vital for the accumulation of knowledge within science. Written responses for question 6 tend to have students focus on the scientific community and the idea that science is a reliable, rigorous process or entity. Last, question 10 written responses convey a type of trust-in-science mindset, exemplified by many students making the conclusion that if a scientific theory or principle does not make sense, then the student must be missing something. Again, basing the idea that the theory must be capable of being understood because science created it and science is a rigorous and reliable process.

Note that item 10 is likely grouping to this factor because of the latter portion of the question, which states "[⋯] not everything in science is supposed to make sense." Question 10 was designed to address the "structure of scientific knowledge" (see Table IX) but some responses did indicate a metacognitive ("nature of knowing and learning") aspect which focused on the "[⋯] you have to accept it and move on [⋯]" question segment.

Overall, this factor seems to be testing student views of the philosophical, psychological, and occasionally sociological aspects of the nature of science (NOS). Aspects of NOS such as "science is an attempt to explain phenomena," "scientists work collaboratively," "science aims to be consistent," and "science relies on skepticism" were often seen within the written student responses [36]. In conclusion, it may be best to categorize this factor as structure of science, given that student responses largely display a belief of science as a network of people, processes, or information involved in understanding everyday life. This is distinct from structure of scientific knowledge but not exclusively so. Written responses often involved discussions of the nature of science, which may then manifest in their views regarding the Structure of scientific knowledge. Epistemologically, this factor could be probing student views with respect to the philosophical and psychological aspects of the nature of science.

2. Factor 2: Innate ability vs hard work

Questions 21, 22, and 25 make up factor two within the five-factor model. These questions are quite likely probing the extent to which students attribute hard work or natural ability to achievement within science, and life in general.

Although no explicit responses were requested for questions 21 and 22, students were asked to define natural ability and hard work. Written response analysis revealed students categorized hard work as dedication to learning the material, perseverance, and/or working to fully understand the material. In summary, one may say students see hard work as dedication and perseverance toward understanding the material. Natural ability, however, was less agreeable in its categorization. The first, and most obvious, response was that natural ability is a trait governing scientific proficiency and that you are born with it. Other students did not clarify if natural ability was truly nature or nurture, but did associate it with an ability to learn the material while having put in considerably less work than their peers. Several of these responses specifically referred to being able to learn the material quickly. Last, a non-negligible number of students referred to natural ability as being something that a student has acquired and brought into the classroom; these students associated natural ability with either prior knowledge, work ethic, and/or interest in the subject. Written responses to question 25 reinforced these findings and further displayed the dichotomy of hard work versus natural ability.

Although this factor seems best suited to be labeled *innate ability vs hard work*, it does seem to still partially align with "source of ability to learn" which the EBAPS associates with questions 22 and 25. Interestingly, there were also traces of Schommer's quick learning present within these questions (via the student written responses) as well, specifically in how students viewed natural ability [37,38]. That is, typical student responses associated natural ability with being able to either learn the material quickly or on their first attempt [37,38]. Epistemologically, this factor appears to pit the influence of hard work against that of natural ability when assessing the role that they play in student views of success.

3. Factor 3: Source of ability to learn

Factor three in the five-factor theoretical model includes EBAPS items 5, 11, 12, and 16. According to the EBAPS (Table IX), the items in this factor are either nature of knowing and learning (11 and 12) or source of ability to learn (5 and 16). Upon analyzing student responses, the one clear trend for all these items is the belief that science can be learned.

Written response data for question 5 focuses on the belief that there are a variety of ways one can learn material, naming strategies such as memorization, incremental studying, and group studying. The same overarching belief, that science can be learned, was present for question 11 in the form of content familiarity. In particular, familiarity with test material as developed via time devoted to studying. Epistemologically, question 12 responses focused on two segments. Based on written response data students likely read item 12 as either "When learning science, it is possible for people to understand the material" or "When learning science, it is beneficial to relate the material to your own ideas." The latter here would test nature of knowing and learning, whereas the former is more closely related to source of ability to learn. This is further made obvious as question 1 occasionally grouped with question 12 in the many approaches or iterations to factor analysis that were utilized (as seen outside of what has been presented within this paper).

Question 1 clearly tests one prominent aspect of epistemic beliefs (as seen with written response data), the belief that making personal connections to the material can be beneficial (nature of knowing and learning). This same belief is one of those two previously mentioned beliefs tested by question 12. Question 1 grouping with 12 but not loading well onto factor three is an indication that the items on factor three are indeed exploring some aspect of source of ability to learn, which is discussed shortly. Last, question 16 data indicates this same repeating belief that science can be learned, only this instance has students focusing on the need for hard work as a means by which to learn. Keep in mind that, in their written responses, these students favor hard work as being dedication and perseverance toward understanding the material, understanding likely developed through the more detailed insight given above with the written question response patterns for 5 and 11.

This factor best represents a student's belief regarding whether or not science can be learned and aligns up quite well with source of ability to learn. There is also commonality with "source of knowledge" present in epistemic literature, in that the knower is a constructor of meaning [39–42]. Epistemologically, this factor seems to measure students beliefs for the efficacy of hard work and good study habits.

4. Factor 4: Nature of knowing and learning

The fourth factor in the five-factor theoretical model has only two items, 28 and 30. Question 28 appears to do an acceptable job of probing student views regarding the tentativeness of scientific knowledge and its limitations. From a larger perspective, however, many students' responses were focused around the view that second-hand data opens up room for interpretation. This question (28) asks students to discuss why multiple theories for the fate of the dinosaurs exist, a context which students primarily consider in their answering of the question. Numerous written responses reveal a student idea that the large time gap involved obscures data and/or that since humans were not around at the time, we simply cannot know for certain. On question 30, students largely agree with the argument that "[$\cdots$] it is possible to get the right answer without really understanding what it means." Their justification favors both a sophisticated view where they may be missing out on conceptual understanding and a more unsophisticated view in that they are missing out on understanding the algorithm necessary to solve the problem. It is also of note that a non-negligible number of students agreed with Jessica, stating that it may be best to simply move on in the interest of time (as opposed to correctness, which the question intended to focus on).

The most common theme between these two items is that a number of conclusions may be achieved despite having uncertainty in the process driving that conclusion. Based on written response data, students believe there to be valuable knowledge in evaluating the process by which conclusions have been drawn. Questions 28 and 30 thus appear to be gauging the value a student places on evaluating current knowledge in the context of science, particularly when students are aware of possible ambiguity in the process by which that knowledge was acquired. In question 28 students evaluate the knowledge put forth by the scientific community whereas in question 30 students evaluate their own knowledge. With respect to the EBAPS axes, these questions appear to be centered on the nature of knowing and learning (see Table IX). There are also similarities to a justification of knowledge component within the epistemic theories of both King and Kitchener and Kuhn—where knowledge requires no justification and there is an acceptance of facts as compared to critical evaluation of the knowledge and judgments to which one is exposed [42,43].

As an aside, it was seen that question 27 occasionally had a strong loading with this factor as well (but struggled with issues of low commonality). Beyond just strong loadings, there is a case to be made for item 27 being considered on this axis. A portion of the argument in item 27 stands out, specifically in the pseudostudent Julia who states, "I still think science applies to almost all real-world experiences. If we can't figure out how, it's because the stuff is very complicated, or because we don't know enough science yet." No written response data exist for question 27, but Julia's statement represents one with which factor four is familiar: evaluation of knowledge. In analyzing items 28 and 30 students are reflecting on how the current state of knowledge came to be, assessing both personal (such as question 30) and external (such as question 28) conclusions. Question 27 upholds the theme of acknowledging uncertainties and their role in the current state of knowledge and knowing.

The relationship of items 27, 28, and 30 to the work of King and Kitchener in their reflective judgment model is intriguing enough to be expanded upon in slightly greater detail. The reflective judgement (RJ) model outlines key stages in the development of critical thinking skills in the context of "ill-structured" problems, or problems which lack an absolute truth and cannot be approached algorithmically [42]. Each stage within the model attempts to link epistemic views to decisions (judgments) made for ill-structured problems. Keen are King and Kitchener on "epistemic cognition," which pertains to how individuals think and reason in relation to their beliefs of "[⋯] the limits of knowing, the certainty of knowing, and the criteria for knowing" when encountering these types of problems [44]. In written responses to the data, students frequently displayed aspects of all three primary domains in the RJ model: prereflective thinking, quasireflective thinking, and reflective thinking [42]. The most frequently occurring responses align with the quasireflective domain where students often made statements of "missing knowledge" and "uncertainty" within the knowledge claims, leading them to have individualistic choices of evidence in supporting their conclusions. It is possible that the items pertaining to factor four are measuring the sophistication of a student's epistemic cognition. This is satisfactory, as metacognition alone could not account for how students were viewing and working with knowledge in their answering of items 28 and 30 for the written response data.

5. Factor 5: Quick learning

Questions 4 and 9 make up the fifth and final factor of the five-factor theoretical model. Recall how in testing stability of the factors that there would arise inconsistencies between group A and group B for items 4 and 9. Namely, group A had 4 and 9 load alone on a factor together whereas group B saw item 9 load onto factor three while item 4 represented its own factor of one item. From simply reading question 4 it is possible that it is assessing quick learning (that science is learned quickly or not at all), like that of Schommer [37,38]. Although no interview data of item 4 are available, we do have data regarding student definitions of natural ability, a word which question 4 (or more specifically, quick learning) has some ties with. To students, one of the primary traits of natural ability is for a student to grasp material in a much shorter time than others. The term "natural ability" appears along with "learn" in question 9 and as such, it is plausible that some students may also view question 9 as stating "Someone who does not learn the material quickly can still learn the material well [⋯]." Hence, it would be possible that question 9 is probing quick learning as well as the belief to which factor three belongs (the degree to which students believe they can learn science by utilizing hard work and good study habits).

Recall that for group A, item 16 was seen to group with 4 (and 9) as well for the five-factor model. In responses to item 16 (which belongs more so to factor 3, source of ability to learn) most students were observed to state that given enough time, anyone can learn science. However, a noticeable portion of these responses revealed the belief that learning occurs at different rates for different people. This belief is not necessarily exemplifying a belief in quick learning and brings into question what items 4 and 9 may be testing. Epistemologically, factor five could be assessing the extent to which the ability to learn the material quickly affects whether material can be learned at all. However, this factor could also merely be assessing the extent to which they believe that learning may occur at different rates for different people, but inevitably can occur. There is currently an effort being made to acquire written response data on item 4 prior to the CFA analysis to clarify this ambiguity. Although this factor appears to be our weakest factor, and could be dropped within CFA, it is retained as it appears to be measuring a belief which is prominent and distinct within students.

B. Theoretical three-factor model

The three prominent factors belonging to the five-factor theoretical model (factor one, factor two, and factor three) are essentially the same factors as are present in the three-factor theoretical model, and thus assess those same beliefs. The only difference is that in the three-factor model question 9 is now grouped with factor 3, which probes *source of ability to learn*. Question 9 did assess *quick learning* (or a similar construct), but also assessed the ability to learn science, i.e., source of ability to learn. This could be seen in written response data where students seemed to focus their answer on the latter portion of question 9 and thus read it as "Someone [⋯] can still learn the material well even in a hard chemistry or physics class." With question 4 no longer present in the analysis (due to restriction to three factors and process of item removal), question 9 consequently no longer experiences a crossloading associated with quick learning. Because of the involvement with another epistemic construct, it is debatable whether or not question 9 should even be considered for this factor in the three-factor theoretical model. This is something to be explored further in future work involving confirmatory factor analysis.

C. A sixth factor

Mentioned previously was the presence of a possible six-factor model, this included another factor beyond that of the five-factor model and would be composed of items 1, 12, and 26. Written response data with question 1 revealed that students saw relating science to their own experiences as beneficial. A similar response pattern was seen for question 12, in which students responded with the belief that relating science to their own ideas was beneficial. Question 12 also had written responses which exemplified the item's association with factor three, in that students believe there are a variety of ways in which one can learn the material. No written response data are available for question 26, however, reviewing the item can still allow for some conjecture regarding what it measures. In particular, the response of pseudostudent Justin in question 26 (that putting science concepts into his own words assisted his learning) highlights a common theme across all three items (1, 12, and 26): that relating aspects of science to one's self is beneficial to learning. It is possible that factor six is outlining a specific study or learning habit that may be implicit in factor three.

Taken together, it is believed the items associated with factor six measure the value students place on making personal connections with science. The stability of this factor is brought into question though, given that question 12 also belongs to factor three and that the other pseudostudent Dave (from question 26) is likely putting emphasis on a separate epistemic belief known as source of knowledge (what does the student depend on for knowledge, themselves or an authority figure?) [45].

IV. THE EBAPS IN ASTRONOMY

In this section the axes of the EBAPS (Table IX) will be discussed in regards to their relationship with the factors found during the exploratory factor analysis. We assert that this is by no means a rigorous test of the validity of the EBAPS instrument or the items as they appear on the axes of the EBAPS. This is an exploration of the prominent epistemological beliefs introductory astronomy students are exhibiting when responding to the EBAPS and how those beliefs may be related to what the EBAPS claims to measure.

The first EBAPS axis, the *structure of scientific knowledge*, mirrors very closely the structure of science factor revealed within this study. The authors of the EBAPS claim the items pertaining to axis one measure the extent to which knowledge in science is viewed as disjoint or isolated, as compared to a unified whole. In truth, the *structure of science* is in many ways measuring what structure of scientific knowledge claims to measure. The conscious choice to not perfectly align this factor with axis one of the EBAPS was due to student written response data. These data revealed that students were drawing their views of the cohesiveness of scientific knowledge from principles related to the nature of science. Evidence hints that this factor may be measuring student sophistication of NOS tenets by proxy with EBAPS questions from structure of scientific knowledge, "real-life applicability," and "evolving knowledge."

The *nature of knowing and learning* is axis two of the EBAPS and measures the value a student places on constructing their own knowledge, as opposed to passively receiving knowledge. This axis probes a variety of methods regarding how a student constructs their own knowledge, from metacognitive strategies to personal experience. It would seem the factor nature of knowing and learning partially tests these same traits. Mentioned before, nature of knowing and learning may be more aptly called *justification of knowledge*, as this factor is an indicator of how students reflect upon what is "known" and how it may, or may not, be known. The possible sixth factor put forth in this study, what we would call *personal connections*, may also fit with this axis. Personal connections appeared to be a factor whose questions explore the extent to which students value making personal connections to science. These personal connections to science involve experiences, ideas, and/or words pertaining to the self. Similarly, nature of knowing and learning claims to test, in part, students' ability to relate "new material to prior experiences, intuitions, and knowledge."

Axis three of the EBAPS is *real-life applicability*. It is believed that no one factor aligns well with this axis, which claims to measure the extent to which science is applicable outside the classroom or laboratory; how well science applies to real life. Structure of science would have the most in common with this axis, as students actively discussed (via written response) science in everyday life for questions along that factor. In particular, question 14, where many student responses viewed science as the machinery behind everyday life (thus politicians should have a working knowledge of science). Again, structure of science was not testing real-life applicability but aspects of the nature of science, yet these NOS principles would often come forth in real-life issues concerning science.

Evolving knowledge, the fourth EBAPS axis, seeks to measure the extent to which students exhibit aspects of dualistic views or relativistic views with scientific knowledge [39]. Dualistic views hold all knowledge as being either right or wrong (absolute truths exist), whereas relativistic views hold scientific knowledge as purely subjective with no absolute truths. Not often do written responses from students take a stance of science as absolute truth which did not change, nor did they frequently exhibit complete immersion in subjectivism, unable to appreciate any evidence put forth by science. Despite not observing these extremes within the data, it is likely that they have some influence, as epistemic literature has essentially confirmed their presence [39,41–43]. The apparent lack of these epistemic perspectives emerging as a factor within this study may be due to the relatively few items which explicitly address them (6, 28, 29) and/or because they are extremes which need to be analyzed independently on a collection of separate questions.

The fifth and final axis of the EBAPS is *source of ability to learn*. This axis claims to probe students' beliefs that they can become better at science, focusing on hard work over natural ability as well as the use of good study stategies. EFA revealed that this axis is seen as two distinct axes according to the students. Specifically, views about the "efficacy of hard work and good study strategies" seems to be getting tested by the factor source of ability to learn (whose name was, of course, motivated by its EBAPS counterpart). However, the first part of the EBAPS authors' description of the fifth axis "Is being good at science mostly a matter of fixed natural ability? Or, can most people become better at learning (and doing) science? [through hard work]" appears to be linked to its own factor, which was called innate ability vs hard work.

To summarize, structure of scientific knowledge, nature of knowing and learning, and source of ability to learn all appear to play a respectable role within the factors revealed in the exploratory analysis. Evolving knowledge did not appear to play a prominent role, but likely this is due to question limitation or the need for independent testing of its constructs (absolutism and relativism). Real-life applicability, as defined by the authors (Table IX), was the only axis that did not hold much significance in student response patterns to the EBAPS. This could be because the trait was playing a more subtle role than could be seen in EFA and, in some respects, the written response data, or because it is better accounted for with other fundamental epistemic constructs.

V. CONCLUSIONS

When exploring the underlying structure of student responses to the EBAPS survey in an introductory astronomy class, there were seen to be three to five notable factors explaining the relationship between item responses. A three-factor and five-factor theoretical model were discussed and their factors placed into context utilizing data from student responses to EBAPS items. Both theoretical models shared three factors, what we called structure of science, innate vs hard work, and source of ability to learn. Structure of Science appears to represent the extent to which students have sophisticated views regarding the nature of science, in particular the philosophical and psychological tenants of NOS [36]. Innate vs hard work directly confronts the views of students regarding the influence that natural ability has as compared to hard work on an individual's ability to succeed. Source of ability to learn is best described by the EBAPS itself in that the items belonging to this factor "probe students' epistemological views about the efficacy of hard work and good study strategies, as distinct from their self-confidence and other beliefs about themselves." [11]. The two additional factors within the five-factor theoretical model are nature of knowing and learning and quick learning. Nature of knowing and learning is exploring the value that a student places on justifying what they know and the acknowledgment that learning can occur through this justification. This factor also shares a notable relationship with what King and Kitchener call "epistemic cognition" [44]. Last, there is quick learning, which probes the extent to which students believe they can still learn the material despite not being able to grasp it as quickly as their peers with natural ability. We do note that it is possible that this final factor is exploring a belief that learning occurs at different rates for different people, but can occur. Future works hopes to obtain a clarification regarding this factor and what it is measuring.

In analyzing student responses to EBAPS items, themes regarding more fundamental epistemological beliefs were seen involving the nature of knowledge and the nature of knowing. Beliefs such as certainty of knowledge, the simplicity of knowledge, and the justification of knowing [45]. To speculate, response data hint that these core epistemic beliefs influence, if not generate, several of the factors in this paper which pertain to the nature of learning and the nature of intelligence.

VI. FUTURE WORK

Future work will involve the use of confirmatory factor analysis in testing the hypothesized variable relationships and the fit of the proposed models. The factors as put forth by the authors of the EBAPS will also be subjected to CFA. A more detailed question-by-question analysis of the EBAPS will be performed. The question-by-question analysis is intended to elaborate more extensively on how students in the astronomy course viewed and responded to questions on the EBAPS and how further clarity and orthogonality could be obtained on these items within this context (introductory astronomy).

APPENDIX: ADDITIONAL INFORMATION

See Tables IX–XIII.

TABLE IX. Description of EBAPS axes and the items associated with these axes. These are the definitions as provided by the EBAPS homepage [12].

1. Structure of scientific knowledge (Items: 2, 8, 10, 15, 17, 19, 20, 23, 24, 28)
   Is [science] knowledge a bunch of weakly connected pieces without much structure and consisting mainly of facts and formulas? Or is it a coherent, conceptual, highly structured, unified whole?

2. Nature of knowing and learning (Items: 1, 7, 11, 12, 13, 18, 26, 30)
   Does learning science consist mainly of absorbing information? Or, does it rely crucially on constructing one's own understanding by working through the material actively, by relating new material to prior experiences, intuitions, and knowledge, and by reflecting upon and monitoring one's understanding?

3. Real-life applicability (Items: 3, 14, 19, 27)
   Are scientific knowledge and scientific ways of thinking applicable only to restricted spheres such as the classroom or the laboratory? Or, does science apply more generally to real life?
   These items tease out students' views of the applicability of scientific knowledge as distinct from the student's own desire to apply science to real life, which depends on the students interests, goals, and other nonepistemological factors.

4. Evolving knowledge (Items: 6, 28, 29)
   This dimension probes the extent to which students navigate between the twin perils of absolutism (thinking all scientific knowledge is set in stone) and extreme relativism (making no distinctions between evidence-based reasoning and mere opinion).

5. Source of ability to learn (Items: 5, 9, 16, 22, 25)
   Is being good at science mostly a matter of fixed natural ability? Or, can most people become better at learning (and doing) science? As much as possible, these items probe students' epistemological views about the efficacy of hard work and good study strategies, as distinct from their self-confidence and other beliefs about themselves.

TABLE X. Finalized factor loadings for group A. Loadings less than 0.25 have been suppressed. EFA involved approximately half of the section 1 students. These loadings represent results after a filtering of the original 30 items and a comparison with group B.

Items	Factor 1	Factor 2	Factor 3
05	0.645		
16	0.576		
09	0.485		
11	0.431		
12	0.419		
22		0.733	
21		0.618	
25		0.403	
08			−0.644
14			−0.598
06			−0.496
10			−0.481

TABLE XI. Finalized factor loadings for group B. Loadings less than 0.25 have been suppressed. EFA involved approximately half of the section 1 students. These loadings represent results after a filtering of the original 30 items and a comparison with group A.

Items	Factor 1	Factor 2	Factor 3
16	0.596		
11	0.498		
05	0.491		
12	0.460		
09	0.438		
22		0.660	
21		0.637	
25		0.505	
08			0.572
14			0.556
10			0.474
06			0.467

TABLE XII. Finalized five-factor loadings for group A. Loadings less than 0.25 have been suppressed. EFA involved approximately half of the section 1 students. These loadings represent results after a filtering of the original 30 items and a comparison with group B.

Items	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
08	0.586				
14	0.579				
06	0.505				
10	0.436				
22		0.781			
21		0.609			
25		0.328			
30			0.520		
28			0.444		
11				0.551	
05				0.525	
16				0.413	
12				0.363	
09					−0.513
04	0.304				−0.482

TABLE XIII. Finalized five-factor loadings for group B. Loadings less than 0.25 have been suppressed. EFA involved approximately half of the section 1 students. These loadings represent results after a filtering of the original 30 items and a comparison with group A.

Items	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
16	0.582				
05	0.541				
09	0.498				
11	0.427				
12	0.402				
22		0.715			
21		0.585			
25		0.471			
04			0.654		
30				0.600	
28				0.378	
06					0.492
14					0.480
08					0.437
10					0.412

[1] S. Willoughby and K. Johnson, Epistemic beliefs of non-STEM majors regarding the nature of science: Where they are and what we can do, Am. J. Phys. **85**, 461 (2017).

[2] K. Johnson and S. Willoughby, Changing epistemological beliefs with Nnature of science implementations, Phys. Rev. Phys. Educ. Res. (to be published).

[3] W. K. Adams, K. K. Perkins, N. S. Podolefsky, M. Dubson, N. D. Finkelstein, and C. E. Wieman, New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey, Phys. Rev. ST Phys. Educ. Res. **2**, 010101 (2006).

[4] D. Duncan and L. Arthurs, Improving student attitudes about learning science and student scientific reasoning skills, Astron. Educ. Rev. **11**, 010102 (2012).

[5] V. K. Otero and K. E. Gray, Learning to think like scientists with the PET curriculum, AIP Conf. Proc. **951**, 160 (2007).

[6] D. Muchoney, The implementation of an interactive engagement model of instruction in high school physics, Ph. D. thesis, The Robert Morris University, 2008, p. 71.

[7] G. Duffy, S. Chance, and B. Bowe, Improving engineering students design skills in a project-based learning course by addressing epistemological issues, *Proceedings of the 40th SEFI Conference, Thessaloniki, Greece* (2012).

[8] J. W. Pellegrino, Assessment of science learning: Living in interesting times, J. Res. Sci. Teach. **49**, 831 (2012).

[9] T. Kubiszyn and G. Borich, *Educational Testing and Measurement* (Wiley, New York, 2007).

[10] T. Hong, S. Purzer, and M. E. Cardella, A psychometric re-evaluation of the design, engineering and technology survey, J. Eng. Educ. **100**, 800 (2011).

[11] The idea behind ebaps, http://www2.physics.umd.edu/~elby/EBAPS/idea.htm.

[12] http://www2.physics.umd.edu/~elby/EBAPS/home.htm.

[13] D. Hammer and A. Elby, Tapping epistemological resources for learning physics, J. Learn. Sci. **12**, 53 (2003).

[14] E. E. Prather, T. F. Slater, J. P. Adams, and G. Brissenden, *Lecture-Tutorials for Introductory Astronomy* (Pearson, New York, 2012).

[15] M. E. Timmerman and U. Lorenzo-Seva, Dimensionality assessment of ordered polytomous items with parallel analysis, Psychol. Methods **16**, 209 (2011).

[16] U. Olsson, Maximum likelihood estimation of the polychoric correlation coefficient, Psychometrika **44**, 443 (1979).

[17] I. H. Bernstein and G. Teng, Factoring items and factoring scales are different: Spurious evidence for multidimensionality due to item categorization, Psychol. Bull. **105**, 467 (1989).

[18] L. R. Fabrigar, D. T. Wegener, R. C. MacCallum, and E. J. Strahan, Evaluating the use of exploratory factor analysis in psychological research, Psychol. Methods, **4**, 272 (1999).

[19] K. A. Douglas, M. S. Yale, D. E. Bennett, M. P. Haugan, and L. A. Bryan, Evaluation of Colorado learning attitudes about

science survey, Phys. Rev. ST Phys. Educ. Res. **10**, 020128 (2014).

[20] H. F. Kaiser, The application of electronic computers to factor analysis, Educ. Psychol. Meas. **20**, 141 (1960).

[21] R. B. Cattell, The scree test for the number of factors, Multivariate Behav. Res. **1**, 245 (1966).

[22] B. P. O'Connor, SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test, Instrum. Comput. **32**, 396 (2000).

[23] https://people.ok.ubc.ca/brioconn/nfactors/rawpar.sps.

[24] N. J. Castellan, Shuffling arrays: Appearances may be deceiving, Behavior Res. Meth. Instrum. Comput. **24**, 72 (1992).

[25] B. G. Tabachnick and L. S. Fidell, *Using Multivariate Statistics* (Allyn and Bacon, Boston, 2001).

[26] A. B. Costello and J. W. Osborne, Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis, Pract. Assess. Res. Eval. **10**, 173 (2005).

[27] J. DeCoster, Overview of Factor Analysis (1998), Retrieved January 2, 2018, from http://stat-help.com/factor.pdf.

[28] B. M. Byrne, *Structural Equation Modeling with Amos: Basic Concepts, Applications, and Programming* (Routledge, United Kingdom, 2016).

[29] H. F. Kaiser, An index of factorial simplicity, Psychometrika **39**, 31 (1974).

[30] B. A. Cerny and H. F. Kaiser, A study of a measure of sampling adequacy for factor-analytic correlation matrices, Multivariate Behav. Res. **12**, 43 (1977).

[31] D. Child, *The Essentials of Factor Analysis* (Continuum International Publishing Group, New York, 2006).

[32] D. Garson, Statnotes: Topics in Multivariate Analysis: Factor Analysis (2010), Retrieved January 1, 2018, from https://faculty.chass.ncsu.edu/garson/PA765/statnote.htm.

[33] E. Guadagnoli and W. F. Velicer, Relation of sample size to the stability of component patterns, Psychol. Bull. **103**, 265 (1988).

[34] J. W. Osborne and D. C. Fitzpatrick, Replication analysis in exploratory factor analysis: What it is and why it makes your analysis better, Pract. Assessment Res. Eval. **17**, 1 (2012).

[35] http://macmillanlearning.com/Catalog/elearningbrowsebymediatype/SaplingLearning.

[36] W. F. McComas and J. K. Olson, *The Nature of Science in Science Education: Rationales and Strategies*, The Nature of Science in International Science Education Standards Documents, (Kluwer Academic, Dordrecht, 1998), p. 41–52.

[37] M. Schommer, Effects of beliefs about the nature of knowledge on comprehension, J. Educ. Psychol. **82**, 498 (1990).

[38] M. Schommer, Epistemological development and academic performance among secondary students, J. Educ. Psychol. **85**, 406 (1993).

[39] W. G. Perry, *Forms of Ethical and Intellectual Development in the College Years: A Scheme* (Rinehart and Winston, New York, 1970).

[40] M. B. Magolda, *Knowing and Reasoning in College* (Jossey-Bass, San Francisco, 1992).

[41] M. R. Belenky, B. M. Clinchy, N. R. Goldberger, and J. M. Tarule, *Women's Ways of Knowing: The Development of Self Voice and Mind* (Basic Books, New York, 1986).

[42] P. M. King and K. Strohm-Kitchener, *Developing Reflective Judgment: Understanding and Promoting Intellectual Growth and Critical Thinking in Adolescents and Adults* (Jossey-Bass, San Francisco, 1994).

[43] D. Kuhn, *The Skill of Argument* (Cambridge University Press, Cambridge, England, 1991).

[44] K. S. Kitchener, Cognition, metacognition and epistemic cognition: A three-level model of cognitive processing, Hum. Dev. **26**, 222 (1983).

[45] B. K. Hofer and P. R. Pintrich, The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning, Rev. Educ. Res. **67**, 88 (1997).