

# Astronomy And Science Student Attitudes (ASSA): A Short Review And Validation Of A New Instrument

Sophie Bartlett, Cardiff University, United Kingdom  
Michael T. Fitzgerald, Edith Cowan University, Australia  
David H. McKinnon, Edith Cowan University, Australia  
Lena Danaia, Charles Sturt University, Australia  
Jasmina Lazendic-Galloway, Monash University, Australia

## ABSTRACT

*A new survey for measuring students' astronomy and science attitudes that has been validated for use with high school students (with a future intent to expand to middle school and university) is presented. We initially present a short review of instruments in the literature that attempt to measure attitudes in astronomy together with the difficulties encountered in measuring these by researchers in the subsequent analyses of results. To illustrate this, we present an example from an Astro101-level university course to display the problems with the current, most commonly used, astronomy attitude instrument. We then present the initial design and the Factor Analysis of a new instrument designed to address the deficiencies of this existing instrument from a sample of students in a high school-level astronomy education project. The factors identified by this instrument include: Interest in Astronomy, Interest in Science Outside of School, Practical Work in Science, Teacher's Actions in science, Perceptions of Ability in Science, Future Aspirations in Science, Benefits of Science, and Personal Relevance of School Science, all of which possess high internal response consistency and construct validity.*

**Keywords:** Attitudes Towards Science; Astronomy Education; Student Interest in Science

## INTRODUCTION

The history of science education reform contains numerous calls for a reassessment of the current approaches due to flagging student interest in science careers and in science in general (e.g., American Association for the Advancement of Science (AAAS) 1990; Millar & Osborne, 1998; Goodrum, Druhan, & Abbs, 2012). Some have gone as far as to call for a 're-enchantment' of science education (Tytler 2007). All rely on arguments that there is some \*better\* approach, in terms of impact on student attitudes and perceptions of science (Danaia, McKinnon, Parker, Fitzgerald, & Stenning, 2012). For a more detailed description of the nature of attitudes as one component of affect, see Slater, Slater, Heyer and Bailey (2015).

Students do not think of one monolithic 'science'. Rather, they think of "physics", "chemistry" and "biology". That is to say, they think of each component of science separately (e.g., Havard, 1996) and "astronomy" is no different. It is very easy to imagine someone very interested in astronomy fainting from boredom in a physics class, even though physics and astronomy tend to co-exist together in many tertiary institutions. Hence, it is very important not only to measure interest in a generic 'science' but also a quite specific interest in the scientific field itself, in this case, astronomy.

Astronomy Education Research (Bailey & Lombardi, 2015) as a content-context specific subset of science education is also concerned with these "better" approaches. A few quotes below from project descriptions of robotic telescope education projects (Gomez & Fitzgerald, 2017) in the literature clearly display this intent:

*“... there is a worrying lack of interest in science and science-related subjects among school children. .... Clearly, astronomy can be used to stimulate an interest in science as a whole.” – (Newsam, Carter, & Roche, 2003) (Liverpool Telescope/National Schools Observatory);*

*“... the Faulkes Telescope Project ... would make a contribution towards improving school students’ attitudes to scientific subjects and towards inspiring interest in them.” – (Beare, 2007) (Faulkes Telescope Project);*

*“We have used astronomy projects ..... in order to encourage their pursuit of science and technology careers.” (Kadooka, Bedient, Hu, Hemphill, & Meech, 2008) (HISTAR);*

*“The EU-HOU project aims at re-awakening the interest for science through astronomy and new technologies ...” (Ferlet, 2008) (EU-HOU);*

*“..... this project offers a straightforward means to enthuse high school students and the general public about radio astronomy ...” (Hollow et al. 2008) (Pulse@Parkes);*

and,

*“[ a goal is to ] ... increase the probability of students choosing science, other than as a potential personal interest, as a topic for higher level study or as a potential future career path or, at the very least, help them discover they are actually interested in science,” (Fitzgerald, McKinnon, & Danaia, 2015).*

These are obviously noble causes or potentially powerful mission statements. Unfortunately, while these motives are claimed, there is little evidence in the literature that these significant shifts in student attitudes and perceptions towards science are being made on any scale (Fitzgerald, Hollow, Rebull, Danaia, & McKinnon, 2014). There are two aspects for this lack of evidence. The first and simpler aspect is that it is rare that anyone is attempting to measure attitude change with any rigor. The second aspect, which may be impacting the first, is there has not been a survey instrument available with which to measure these changes with rigor or relevance.

It is this second aspect that provided the motivation for construction of a new, well validated, instrument focused specifically on astronomy rather than just science in general. The goal in this paper is to provide an astronomy and science focused attitude survey for use with projects at middle-school, high-school and college-level students. A particular motivation for the construction of this instrument is to have a standard survey with which to appraise the effectiveness of various approaches and implementations in shifting students’ attitudes.

The purpose of this paper is to present a new survey that has been validated for use with high school students that measures six science attitude factors and one interest in astronomy factor. We initially present a short history of attempts to measure attitudes in astronomy and their difficulties. To illustrate this, we present an example from an Astro101-level university course to display the problems with the current, most commonly used, astronomy attitude instrument. We then present the initial design and then the Factor Analysis of a new instrument designed to address the deficiencies of the previous instrument from a sample of students in a high school-level astronomy education project. Reliability and Construct Validity analyses for the scales are also presented. Future directions and uses of this instrument are then discussed.

### **SHORT HISTORY OF ATTITUDES IN ASTRONOMY**

While content knowledge has typically been the dominant topic of astronomy education research, research into affect, identity and attitude (e.g., Wittman, 2009) are not too far behind (Slater et al. 2016). Of the 111 articles and theses tagged as “Affective” in the iSTAR abstract database (istardb.org) at the time of writing, 56 were coded as quantitative, 35 as qualitative and 20 as mixed methods. Inspecting the 76 quantitative and mixed methods articles and theses, we found that 42 were focused on our target groups (middle-school, high-school and college students) while also directly inquiring about student attitudes. Of these, 33 tackled the measurement of attitudes with a

traditional Likert scale questionnaire. Out of these questionnaires only 18 (or around half) used an instrument that possessed any semblance of, or attempt at, validation.

This means that approximately half of the quantitative research undertaken into student attitudes in astronomy thus far has been undertaken with non-validated instruments. Commonly, this was in the form of a survey created from scratch by the author or by taking an existing survey and altering it so dramatically that it would have required re-validation, which was not reported in the research.

The majority of papers *with* validated scales used Zeilik's Attitude towards Astronomy Instrument (1997, 1999) which is discussed at length below, while others used instruments such as CLASS (Adams, Perkins, Dubson, Finkelstein, & Wieman, 2005) or EBAPS (White, Elby, Frederiksen, & Schwarz, 1999). There has also been some work in Turkey with an instrument called the 'Astronomy Attitude Scale' (Bektasli, 2016; Turk & Kalkan, 2015; Bilici, Armagan, Cakir, & Yuruk, 2012).

The "Survey of Attitudes towards Astronomy" (ATA) from Zeilik, Schau, Mattern, Hall, Teague, and Bisard (1997) and Zeilik, Schau, and Mattern (1999) was based upon the work of Schau, Stevens, Dauphinee, and Del Vecchio (1995) in the domain of tertiary statistics education. The instrument posits four factors, but it is not clear that these factors were independently factor analysed or just reinterpreted or rephrased from Schau et al. (1995)'s Attitudes Towards Statistics survey. The four factors are:

- 1) Affect (A)-the student's positive and negative feeling concerning astronomy;
- 2) Cognitive Competence (CC)- attitudes about intellectual knowledge and skills applied to astronomy;
- 3) Value (V)- attitudes about the usefulness, relevance, and worth of astronomy in personal and professional life; and,
- 4) Difficulty (D)- attitudes about the difficulty of astronomy as a subject.

Zeilik et al. (1999) saw little change in these four factors over one semester of teaching using the ATA and commented that shifting these attitudes in such a short time frame was a difficult or an impractical task especially when evidence at the time showed that undergraduate physics studies tended to lead to *decreases* in favourable attitudes of students in their classes (Redish, Steinberg, & Saul, 1998). This resistance to change in attitude scores over the course of a semester-long implementation is also mirrored in all of the relatively infrequent later uses of the ATA instrument (Kareva, Miller, Foster & James, 2014; Bektasli, 2013; Ucar & Demircioglu, 2011; Miller & Redman, 2010; Richwine, 2007; Platco, 2005; Slater & Jones, 2004; Zeilik & Morris, 2003; Zeilik, Bisard & Lee, 2002). In all of these studies there were no significant changes detected from the pre- to post-instruction occasion. There is no known study that the authors have found where a change, significant or otherwise, was reported using this instrument in a repeated-measures design.

This is likely not just a property of this instrument. Slater, Slater and Shaner (2008) used a modified version of the "Attitudes toward Science Inventory" (Gogolin & Swartz, 1992; Weinburgh & Steele, 2000) to test attitude change over a semester-long pre-service elementary teacher course. Their study showed no change in attitudes despite seeing significant gains in content knowledge.

Fitzgerald, McKinnon, Danaia, and Deehan (2016) used a modified and extended version of the Secondary School Science Questionnaire (Goodrum, Hackling, & Rennie, 2001) partially to probe affective changes in 15-16 year-old high-school students' participation in an inquiry-based astronomy course. Again, while good gains were made in content knowledge and statistically significant shifts in views of their science classroom were also seen, items related to their affect towards science or their opinions of science and school science were little changed. Similarly, large shifts in views of their science classroom but minimal shifts in their affect and opinion were seen using the same tool in junior high school (Danaia, McKinnon, & Fitzgerald, 2017).

It is either something about the typical design of these astronomy subjects that does not seem to shift attitude or there is something about the current questionnaires that do not capture any changes in student attitudes towards astronomy or science. Maybe different techniques and approaches truly need to be used to shift attitude.

Alternatively, attitude itself may not change on such restricted (<1 semester) timescales. The only exception that we have found is a paper using a custom designed five-scale instrument (Slater, Safko, & Carpenter, 1999) which did see shifts in Astronomy attitudes that were sustained over the longer term by in-service teachers undertaking an “Astronomy for Teachers” course.

### **AN ILLUSTRATIVE EXAMPLE OF THE ZEILIK ATA INSTRUMENT**

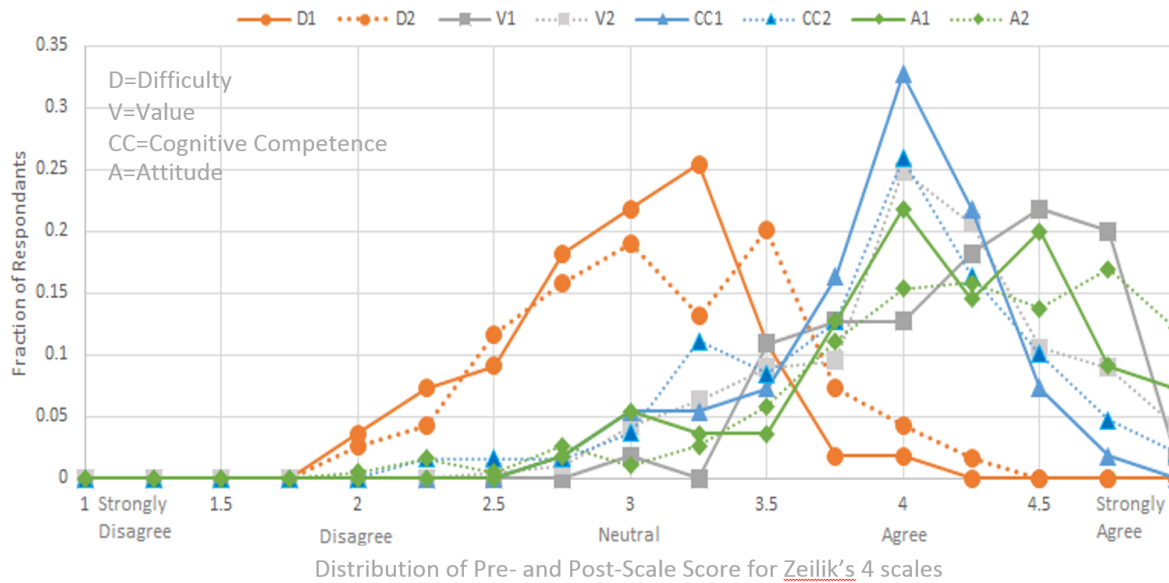
As an example, we use some recent data from Monash University, Australia, obtained in an Astro101-level subject (Lazendic-Galloway, Fitzgerald, & McKinnon, 2016). A total of 55 students completed the pre-test and 189 completed the post-test using the Zeilik ATA survey. Due to the large number of items compared with the small number of responses, Factor Analysis of these data was not undertaken. We thus assume the factor structure from the prior publications (Schau et al. 1995, Zeilik et al. 1997). We did compute reliability analyses for items in the reported scales. Cronbach’s Alphas for the scales were Affect (0.824), Cognitive Confidence (0.830), Value (0.74) and Difficulty (0.66). Tukey’s test for additivity for each scale was adequate suggesting that the values for the Likert scale responses could be added together to form scales for each factor.

The difference in response rate between the two occasions led to two statistical tests being employed. The first treated the pre- and post-populations as independent groups to exploit the larger statistical power from the larger N available. The second was a matched-pairs analysis, with the 40 identifiable students who completed the same survey on both occasions. For the comparison of the pre- and post-responses as two independent sets requiring four separate univariate tests, we use the average inter-item correlation of 0.418 to calculate a modified Bonferroni-correction yielding a p-value of < 0.023 below which significance was indicated. As the distribution of responses for each scale were not normally distributed (6 out of the 8 distributions failed the Shapiro-Wilk normality test), a Mann-Whitney U/Wilcoxon Rank Sum test was computed for each scale.

Three of the scales showed no statistical significance between the pre- and post-tests, (D 0.173, CC 0.756 and A 0.501). However, the ‘Value’ Scale did have a significant change (Sig 0.015 < p = 0.023). While it is significant, the effect is small (pre mean = 4.1 SD = 0.47, post mean = 3.9 SD = 0.51). If this is a real effect, it may be due to the fact that students who valued astronomy less were less likely to make sure they handed in their survey on the pre-occasion compared to the post-occasion and, hence is an artifact of the data collection strategy rather than a truly statistically significant difference. Thus, the result should be treated with caution.

In the analysis of the pre- and post-responses as a matched pre/post data set with the much smaller population, we used a Wilcoxon Signed-Rank Test and found no significant differences (A 0.035, D 0.328, CC 0.980, V 0.812, all above the criterion of p = 0.023) from the pre- to the post-occasion. Figure 1 shows the pre- and post-results of the attitude survey for the four scales derived from the instrument. Of note in this graphical representation is the large negative skew in the distributions of the scales. For three of the four scales, there would likely be a ceiling effect making it even more difficult to detect any changes.

**Figure 1.** Distribution of student scores on the four scales using Zeilik’s (et al. 1997, 1999) Attitude towards Astronomy Scale. Solid lines are pre-test, dashed lines are post-test.



It is likely that as Monash University is one of the top science and engineering universities in the country with the highest Australian Tertiary Admission Rank (ATAR) requirements to a Bachelor of Science program, the responding students already highly value science. The students’ appreciation of astronomy is already as high as could be reasonably expected to be achieved while their perception of the level of Difficulty seems to be at an acceptable level. The results from the attitude test are ‘optimal’. By this we mean that if you were asked to draw a hypothetical profile for a cohort of students who best represented good, but also reasonable, attitudes towards astronomy and astronomy education, these measured profiles would be a close match. Nonetheless, detecting significant change is rendered more difficult.

### DESIGN OF NEW INSTRUMENT

Although existing Attitudes towards Astronomy questionnaires exist, we are yet to find one that has been thoroughly tested and validated, is not limited by floor or ceiling effects and is therefore capable of measuring change. A new instrument was trialed with students in the Our Solar Siblings program (Fitzgerald, et al. 2018, in an attempt to provide a tool that counters these problems and reliably and validly measures the effectiveness of relevant intervention projects in science education with a high degree of accuracy.

This new instrument is not brand-new per se. Rather, it is an adaptation of previous instruments but focused specifically on astronomy rather than general science. The 34 items from the Attitudes Towards Science instrument constructed and validated by Kind, Jones and Barmby (2007) were combined with a further 20 items deemed potentially useful from previous studies by Fitzgerald et al. (2016) and Danaia et al. (2017) that used the Secondary School Science Questionnaire (Goodrum et al. 2001) with an additional 11 newly created items intended to probe students’ interest in astronomy more directly. These 11 new items are presented below.

1. I have a strong interest in learning about the Universe
2. I like to use scientific tools to study the world around me.
3. I would rather do a science experiment than read about science.
4. I find astronomy interesting
5. I feel drawn to the night sky

6. I am interested in the stories behind the constellations.
7. I would like to own my own telescope
8. I would like to undertake astronomy as a career.
9. I would like to do astronomy as a hobby.
10. I would prefer to use a large research telescope through the internet than a smaller telescope in my backyard
11. I would prefer to look at the night sky directly with my eye rather than take a much higher quality picture with a camera.

The adopted items from Kind, Jones and Barmby and the original items generated a 65-item questionnaire with eight constructs considered to guide students' attitudes towards science and astronomy. These eight constructs were General Interest in Science, Ability in Science, Practical Work in Science, Interest in Science Outside of School, Interest in Astronomy, Future Aspirations in Science, Teacher's Actions and Personal Relevance of Science

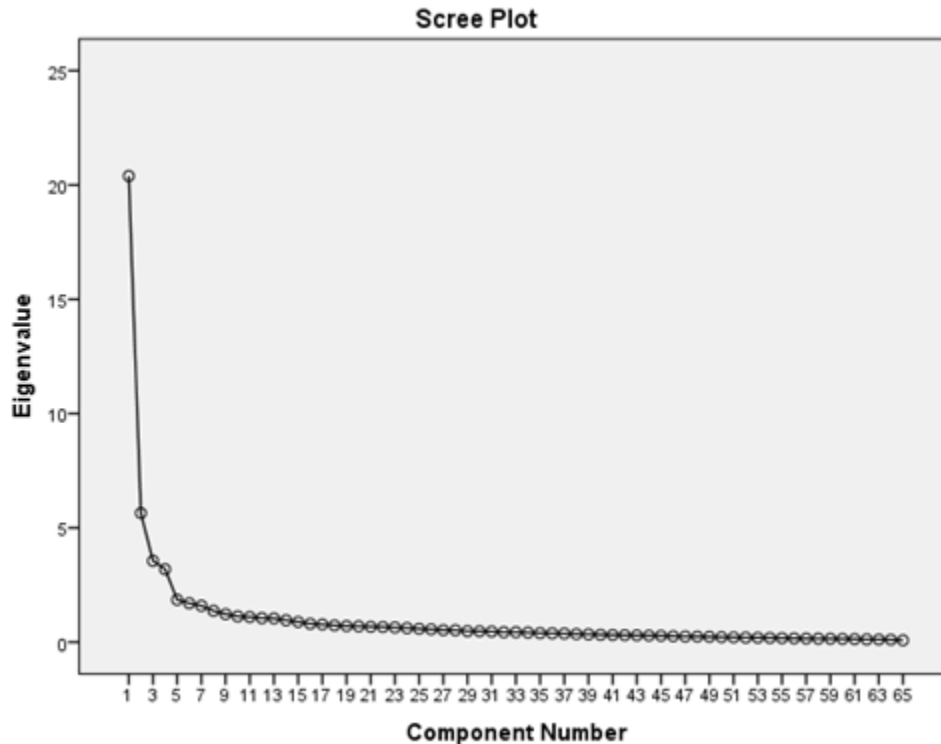
Students responded to all statements using a five-point Likert Scale: Strongly Agree, Agree, Undecided, Disagree, Strongly Disagree. All respondents were high school students participating in the Australian-based Our Solar Siblings project (Fitzgerald et al. 2018). A total of 429 students supplied matched pre- and post-occasion data from eight high schools ranging across Grades 9 to 11 during 2016 academic year.

### **FACTOR ANALYSIS OF THE NEW INSTRUMENT**

Exploratory Factor Analysis is a commonly used statistical procedure to reduce the number of individual items in a questionnaire to a much fewer number of "latent factors" that influence respondents' reactions to the items. It is based on the hypothesis that these latent factors influence the responses and that the influence can be detected through the correlations amongst the items. The net result of this search for these latent factors is that many fewer calculations need be entertained. That is to say, the search for parsimony reduces the researcher's final computational effort. For example, if a questionnaire designed to be used in both a pre- and a post-test for a class contained the same 50 items, then 50 univariate calculations would have to be computed to determine if any changes occurred. Moreover, given that there are 50 univariate calculations to be computed on data that were acquired at the same times, one cannot use the commonly accepted significance level of 0.05. The same questionnaire may only contain four or five latent factors that influence individuals' responses to the entire 50 items. Hence there would only need to be four or five calculations (Analysis of Variance with repeated measures on the occasion of testing), and sometimes only one (a Multivariate Analysis of Variance with repeated measures on the occasion of testing if certain statistical conditions are met), that the researcher needs to compute in order to detect change or changes.

Consequently, in this paper, Exploratory Factor Analyses (EFA) were computed using SPSS v23 with principal components extraction and Direct Oblimin Rotation, because any factors to emerge were anticipated to be correlated. Items with loadings less than 0.3 on to any given factor were eliminated. The unconstrained analysis produced 13 factors with eigenvalues greater than 1. However, one factor produced an eigenvalue of 20.389 and was responsible for 31.368 % of the variance. This large eigenvalue suggests that there is one overarching factor that may explain students' reactions. It could be described as an "overarching interest in astronomy or science". Nonetheless, investigation of the rapid change in gradient of the scree plot shown in Figure 2 indicates that there are likely to be 4-6 factors worthy of further exploration.

Figure 2. Scree plot



Inspection of the unconstrained 13-factor output revealed that many of the factors contained only 1 to 3 items. We deemed it reasonable to explore solutions for 5, 6 and 7 factors given the rapid change in gradient of the scree plot in this region. Initially, however, given that there were eight factors hypothesized to be in the questionnaire as described above, a constrained 8-factor solution was computed. The output of this produced one factor that was not interpretable. The 7-factor constraint also produced factors that were impossible to interpret. Consequently, the 8-factor and 7-factor solutions were eliminated.

The constrained 6-factor solution yielded factors that were more interpretable. Table 1 is the Pattern Matrix that has been sorted according to the magnitude of the factor loadings and which indicate the relative importance of an item within a factor. However, some items loaded on two factors, which is not uncommon when the researcher allows the factors to be correlated. In addition, two of these factors appeared to be a mixture of two related constructs. These are Factors 1 and 4. The related constructs have been highlighted in different shades of grey.

Overall however, the six-factor solution produces factors that are more distinctive and appear to contain items that could be described as separate areas of attitude. Items describing an “Interest in Astronomy” and an “Interest in Science Outside of School” were all found to load onto Factor 1. This indicates, perhaps, that there is an overarching interest in science describing these two aspects of attitude.

Consequently, we computed a separate Factor Analysis for only these “Interest” items contained in Factor 1 of the six-factor solution. This yielded two factors that indicated these two separate but related aspects of interest, viz., “Interest in Astronomy” and an “Interest in Science Outside of School”. We repeated this with all of the items that appear in Factor 4 of the six-factor solution. This computation also produced two sub-factors, one of which appears to be related to a “Perception of Personal Ability in Science” while the second seems to be related to students’ “Future Aspirations in Science”.

Given these outcomes for Factors 1 and 4, it was decided to investigate the reliabilities of these “related” sub-factors while seeking a parsimonious solution to reduce the time required for students to complete the questionnaire yet allowing the researcher to probe the various aspects of attitude. That is to say, we wished to measure the constructs reliably, eliminate extraneous and low-loading items while retaining the rich picture of the various contributions to attitude. Thus, the remaining analyses were performed on the eight factors described as: Interest in Astronomy, Interest in Science outside of School, Practical Work in Science, Teacher’s Actions, Ability in Science, Future Aspirations in Science, Benefits of Science, and Personal Relevance of School Science.

**Table 1.** Pattern Matrix for the 6 Factor Solution

	Factor Loadings (Pattern Matrix)					
	1	2	3	4	5	6
I would like to do astronomy as a hobby.	.774					
I find astronomy interesting.	.734					
I feel drawn to the night sky.	.685					
I would like to undertake astronomy as a career.	.680					
I would like to own my own telescope.	.677					
I have a strong interest in learning about the universe.	.604					
I am interested in the stories behind the constellations.	.562					
I like watching science programmes on TV.	.556					
I like to visit science museums.	.548					
I would like to do more science activities outside school.	.545					
I like reading science magazines and books.	.535					
It is exciting to learn about new things happening in science.	.499					
I would like to join a science club.	.450			.352		
I like to use scientific tools to study the world around me.	.436					
I would prefer to use a large research telescope through the internet than a smaller telescope in my backyard.						
I would like more practical work in my science lessons.		.893				
I look forward to doing science practicals.		.830				
I would rather do a science experiment than read about science.		.810				
We learn science better when we do practical work.		.799				
Practical work in science is exciting.		.793				
I like science practical work because you don't know what will happen.		.777				
Practical work in science is boring. (Recoded.)		.679				
Practical work in science is good because I can work with my friends.		.667				
I like practical work in science because I can decide what to do myself.		.628				
My science teacher shows us how new work relates to what we have already done.			-.830			
My science teacher takes notice of students' ideas.			-.828			
My science teacher makes it clear what we have to do to get good marks.			-.814			
My science teacher tells me how to improve my work.			-.807			
My science teacher talks to me about how I am getting on in science.			-.782			
My science teacher uses language that is easy to understand.			-.690			
In my science class we learn about scientists and what they do.			-.620			
My science teacher gives us quizzes that we mark to see how we are going.			-.609			
My science teacher marks our work and gives it back quickly.			-.564			
My science teacher lets us choose our own topics to investigate.			-.552			
We learn interesting things in science lessons.			-.512			
Science lessons are exciting.	.304		-.486			

(Table 1 continued on next page)



(Table 1 continued)

	Factor Loadings (Pattern Matrix)					
	1	2	3	4	5	6
In my science class we plan and do our own experiments.			-.460			
In my science class we do experiments by following instructions			-.352			
I find science difficult.				.834		
Science is one of my best subjects.				.801		
I am just not good at Science.				.781		
I get good marks in Science.				.719		
I learn Science quickly.				.707		
I would like to study science at university.				.665		
I would like to have a job working with science.				.647		
I feel helpless when doing Science.				.636		
In my science class, I understand everything.				.629		
I like Science better than most other subjects at school.				.579		
I would like to study more science in the future.	.301			.555		
I would like to become a scientist.	.368			.514		
I would like to do more science at school.	.339			.491		
Science is boring.			-.338	.418		
In my science class I work out explanations in science on my own.				.403		.359
I look forward to my science lessons.			-.385	.389		
Science and technology makes our lives easier and more comfortable.					-.805	
Science and technology is important for society.					-.772	
The benefits of science are greater than the harmful effects.					-.620	
In my science class I copy notes the teacher gives me.					-.440	
The science we learn at school is useful in everyday life.			-.308			.518
The science we learn at school helps me make decisions about my health.						.476
The science we learn at school is relevant to my future.				.316		.467
The science we learn at school deals with things I am concerned about.			-.359			.411
The science we learn at school helps me understand environmental issues.			-.324			.354
In my science class I watch the teacher do an experiment.						.344
I would prefer to look at the night sky directly with my eyes rather than take a much higher quality picture with a camera.						

### Interpretations, Reliability and Construct Validity Analyses

The items loading on Factor 1 appear to cover two related aspects of “interest”. As described above, we have split them into two domains: an interest in things astronomical, and an interest in science more generally (highlighted in grey in Table 1). Our reason for doing this is that a person might not be interested in astronomy in particular but could be interested in another branch of science or indeed in things scientific more generally. These separate domains are described as Factor 1 and Factor 2 below, and which are separately boxed in lighter and darker shades of grey in Table 1. The same was also done for Factor 4, which was split into the two domains: students’ perceived ability in science and students’ intentions for science in their future. These two domains are described as Factor 5 and Factor 6 below, and which are also separately boxed in lighter and darker shades of grey in Table 1.

Initially, where more than six items loaded on a factor, the highest loading six items were used to compute Reliability Analyses to probe the potential for forming the items into a scale. Where results did not produce an adequate Cronbach’s Alpha of greater than 0.7, alternative combinations of all items were investigated with the elimination of those suggested by the analysis. Two criteria were adopted while computing the Reliability Analyses. These were to achieve: a Cronbach’s alpha of at least 0.7; and, a Tukey’s index close to 1.0 so that the raw Likert Scale responses could be added together to form a scale without further mathematical transformation. Where these criteria were not met individually or collectively, lower loading items from the Factor were investigated.

A scale for each Factor was constructed by summing the Likert scores for each item and dividing this by the number of items that formed the reliable scale to render it in the same metric as the original Likert scale. The following Tables present the Items, Cronbach's alpha and the Tukey index for each Factor. Once a reliable scale with an adequate Tukey index was identified and comprising the items reported below, the Construct Validity was then investigated using One-way Analysis of Variance procedures with the scale score as the dependent variable (DV) and a separate item not employed in the construction of any of the scales as the independent variable (IV). The results from these processes are reported in the following sections.

### Factor 1: Students' Interest in Astronomy

Factor 1 is interpreted to mean a general "Students' Interest in Astronomy". The two strongest loading items are: "I would like to do astronomy as a hobby" (0.774) and "I find astronomy interesting" (0.734).

The six items with the highest loadings are shown in Table 2. Table 2 also shows the outcome of the Reliability analysis for these six items yielding a Cronbach's alpha of 0.862 demonstrating a high internal consistency. Moreover, Tukey's test of additivity was not significant and the index to which the individual items should be raised to achieve such additivity was 0.997. Thus, the individual raw item response scores can be added to produce a scale. The scale so formed is described as *Students' Interest in Astronomy*.

**Table 2.** Factor 1 – Students' Interest in Astronomy

Items	Loadings	Alpha	Tukey Index
I would like to do astronomy as a hobby	0.774	0.862	0.997
I find astronomy interesting	0.734		
I feel drawn to the night sky	0.685		
I would like to undertake astronomy as a career	0.680		
I would like to own my own telescope	0.677		
I have a strong interest in learning about the Universe	0.604		

To assess the construct validity of the scale, a One-way ANOVA with *post-hoc* multiple comparisons was computed. In order to protect for the fact that multiple univariate tests are computed on the independent variable means, a Student-Newman-Keuls (SNK) offering moderate protection for multiple separate comparisons to identify homogenous sub groups was employed.

The item "I am interested in the stories behind the constellations" was used as the IV. This item did not load on the scale. It can be hypothesized that should a student have a *high level of interest* in astronomy as measured by their scale score, they would also be likely to agree or strongly agree with the statement that they are interested in the stories behind the constellations. Likewise, should a student express a low level of interest in astronomy, they would tend to disagree or strongly disagree that they are interested in these stories.

Consistent with expectations, the results of the One-way ANOVA demonstrate that there is a highly significant difference in the mean scores on the *Interest in Astronomy* scale of the students who indicated they are interested in the stories behind the constellations, compared to those who indicated they are not interested, ( $F(4, 314) = 38.733, p < 0.001$ ). The *post-hoc* multiple-comparisons test (SNK) revealed where the significant differences lay.

**Table 3.** Post-hoc multiple comparisons output (SNK) for the Interest in Astronomy Scale

IV= I am interested in the stories behind the constellations	N	1	2	3	4
Strongly Agree	18	4.1581	3.5673		
Agree	32			3.2273	
Undecided	77			3.0990	
Disagree	114				2.2130
Strongly Disagree	78		0.403	1.000	1.000
Sig.		1.000			

Subset for alpha = 0.05

Table 3 shows that the students who Strongly Agreed with the statement of interest `in the stories behind the constellations had significantly greater mean scores in comparison to all other students who were not interested. Students who Agreed with the statement had a mean score significantly greater than those who were undecided or disagreed in any way with the statement. While the students who responded with Undecided had a higher mean score than those who Disagreed, the difference was not significant. Both of these groups, however, had a significantly higher mean score than those who Strongly Disagreed with the statement that they were interested in the stories behind the constellations. These results indicate that this particular scale possesses high construct validity. It can be concluded that this scale appears to measure Students' Interest in Astronomy.

#### Factor 2: Interest in Science Outside of School

Factor 2 is interpreted to mean a more general "Interest in engaging with Science Outside of School" rather than an interest purely in astronomy (Factor 1) even although in the 6-Factor solution these items were all identified as the same factor. The two strongest loading items are: "I like watching science programs on TV." (0.556) and "I like to visit science museums" (0.548).

This second factor appears to indicate a more general interest in science than the items focused on solely astronomy in Factor 1. Table 5 shows the six items, their factor loadings and the outcome of the reliability analysis yielding a Cronbach's alpha value of 0.848 and an acceptable Tukey index value of 1.226.

**Table 4.** Factor 2 - Interest in Science Outside School

Item	Loading	Alpha	Tukey Index
I like watching science programs on TV	.556	0.848	1.226
I like to visit science museums	.548		
I would like to do more science activities outside school	.545		
I like reading science magazines and books	.535		
I would like to join a science club	.450		
I like to use scientific tools to study the world around me	.436		

The construct validity of the scale was investigated, as before, by computing a One-way ANOVA with post-hoc multiple comparisons and SNK protection. For this scale, it was hypothesized that a student who had an Interest in Science Outside of School would also be more likely to find their science lessons exciting. Thus, it was hypothesized that students who have an interest in engaging with science outside of school as measured by their scale score (DV) would also be more likely to agree that their science lessons are exciting (IV). Alternatively, students not having an interest in science outside of school would tend to disagree that their science lessons are exciting.

**Table 5.** Post-hoc multiple comparisons output (SNK) for the Interest in Science Outside of School Scale

IV= Science lessons are exciting	N	1	2	3	4	5	
Strongly Agree	13	4.0032					
Agree	22		3.3702				
Undecided	92			2.9438			
Disagree	140				2.4848		
Strongly Disagree	52					1.9615	
Sig.		1.000	1.000	1.000	1.000	1.000	

Consistent with expectation, the results of the One-way ANOVA demonstrate that there is a significant difference in the mean scores on the Interest in Science Outside of School scale of the students who indicated that their science lessons are exciting, compared to those who indicated that they were not exciting ( $F(4, 314) = 36.366, p < 0.001$ ). Results of the multiple comparisons test are presented in Table 5. Here, students who Strongly Agreed with the statement that science lessons are exciting had significantly higher mean scale scores than all other students. Indeed, each homogeneous subset across the IV possess significantly different mean scale scores. These results suggest that this particular scale possesses high construct validity and measures students' interest in engaging in science outside of school.

### Factor 3: Practical Work in Science

Factor 3 is interpreted to mean the motivation generated by engaging in "Practical Work in Science". The two strongest loading items are: "I would like more practical work in my science lessons" (0.893) and "I look forward to doing science practicals" (0.830). Items loading on Factor 3 are shown in Table 6 describe feelings about carrying out practical work in science lessons at school.

Reliability analysis for the six strongest loading items yielded a Cronbach's alpha of 0.910 demonstrating high internal consistency. Moreover, Tukey's test of additivity was not significant and the index to which the individual items should be raised to achieve such additivity was 0.992. Thus, the individual item-response scores can be added to produce a scale that seems to measure students' experiences of practical work in science.

**Table 6.** Factor 3 – Practical Work in Science

Item	Loading	Alpha	Tukey Index
I would like more practical work in my science lessons	0.893	0.910	0.992
I look forward to doing science practicals	0.830		
I would rather do a science experiment than read about science	0.810		
We learn science better when we do practical work	0.799		
Practical work in science is exciting	0.793		
I like science practical work because you don't know what will happen	0.777		

A One-way ANOVA with post-hoc multiple comparisons (SNK) was computed using the scale score as the DV and the item "Practical work in science is boring" as the IV in order to assess the construct validity of the scale. It was hypothesized that if the student possessed a positive disposition towards practical work in science (the DV), they would tend to disagree with the statement that practical work in science is boring (the IV). Similarly, should a student dislike practical work, they would tend to agree that they find practical work boring.

Consistent with expectation, the results of the One-way ANOVA demonstrate that there is a highly significant difference in the mean scores on the *Practical Work in Science* scale of the students who indicated they did not think practical work was boring in science, compared to those who indicated that they do, ( $F(4, 314) = 44.525, p < 0.001$ ). Results of the multiple comparisons test are presented in Table 8 where the significant differences amongst students' responses can be seen. Here, students who agreed with the statement of interest, had significantly lower

mean scores in comparison to those students who disagreed. The results demonstrate that this particular scale possesses high construct validity and measures students' attitude towards practical work in science.

**Table 7.** Post-hoc multiple comparisons output (SNK) for the Practical Work in Science scale

<b>IV=Practical work in science is boring</b>	<b>N</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
Strongly Agree	5	2.4333			
Agree	19	2.7368	2.7368		
Undecided	47		3.2021		
Disagree	135			3.7395	
Strongly Disagree	113				4.3791
Sig.		0.201	0.050	1.000	1.000

Subset for alpha = 0.05

#### **Factor 4: Teacher's Actions in Science**

This factor is interpreted to be what the student perceives their teacher does to help them during their science classes. The two strongest loading items shown in Table 8 are: "My science teacher shows us how new work relates to what we have already done" (-0.830) and "My science teacher takes notice of students' ideas" (-0.828).

**Table 8.** Factor 4 – Teacher's Actions in Science

<b>Item</b>	<b>Loading</b>	<b>Alpha</b>	<b>Tukey Index</b>
My science teacher shows us how new work relates to what we have already done	- .830	0.900	1.128
My science teacher takes notice of students' ideas	- .828		
My science teacher makes it clear what we have to do to get good marks	- .814		
My science teacher tells me how to improve my work	- .807		
My science teacher talks to me about how I am getting on in science	- .782		
My science teacher marks our work and gives it back to us quickly	- .564		

The first Reliability Analysis was computed with the six highest loading items indicated in Table 1 (shown in the Factor 3 column). Although it produced a high Cronbach's alpha value (0.916) it also produced a high Tukey index outside of the acceptable range of 0.7 to 1.3, which indicates that the items cannot simply be added together to create a scale. Rather than undertaking the suggested mathematical transformations and given that there were several other lower loading items available, alternative combinations were investigated.

The problematic item appeared to be "My science teacher uses language that is easy to understand". Consequently, a decision was made to remove this item. Further computations revealed that replacing it with the lower loading item "My science teacher marks our work and gives it back to us quickly" produced a scale with better properties.

Reliability analysis for the six items shown in Table 8 yielded a Cronbach's alpha of 0.900 demonstrating high internal consistency. Moreover, Tukey's test of additivity was not significant and the index to which the individual items should be raised to achieve such additivity was 1.128. Thus, the individual item response scores can be added to produce a scale that seems to measure what the student perceives their teacher does to help them.

To assess the construct validity of this scale, we computed One-way ANOVA with post-hoc multiple comparisons as before. The 6-item Teacher's Actions scale was computed as the DV and the item "My science teacher gives us quizzes that we mark to see how we are doing" was used as the IV.

It was hypothesized that should students agree that their teacher carries out particular actions to help them in science as measured by their scale score, they would tend to agree that their teacher gives them quizzes that they mark to see how they are doing. Likewise, should students believe that their teacher does not carry out particular actions in their science classes to help them, they would be more likely to disagree with the statement.

Consistent with expectation, the results of the One-way ANOVA demonstrate that there is a highly significant difference in the mean scores on the Teacher's Actions scale of the students who indicated that their teacher gives them quizzes, compared to those who indicated that their teachers do not, ( $F(4, 314) = 54.201, p < 0.001$ ). Results of the post-hoc multiple comparisons are presented in Table 9 where the significant differences amongst students' responses can be seen. Here, students who Strongly Agreed and Agreed with the statement about quizzes had significantly higher mean scores in comparison to those students who were Undecided or Disagreed. These results suggest that this particular scale possesses high construct validity and can reliably measure students' perceptions of their teacher's actions to help them in their science classes.

**Table 9.** Post-hoc multiple comparisons output (SNK) for the Teacher's Actions Scale

IV= My science teacher gives us quizzes that we mark to see how we are doing	N	1	2	3	4
Strongly Agree	16	4.5282			
Agree	32		3.8333		
Undecided	65			3.5333	
Disagree	141			3.2865	
Strongly Disagree	56				2.1250
Sig.		1.000	0.100	1.000	1.000

Subset for alpha = 0.05

#### Factor 5: Perception of Ability in Science

This factor (Factor 4, light grey in Table 1) is interpreted to mean the student's perception of their ability in science. The highest loading reverse coded item was "I find science difficult" (0.834) which can be read as "I find science easy". The second highest loading item was "Science is one of my best subjects" (0.801). The six highest loading items primarily describe students' perception of their ability and efficacy with regards to science at school.

Reliability analysis of the original six highest loading items in this factor produced a high Cronbach's alpha (0.889) but an inadequate value for the Tukey Index. Inspection of the Reliability Analysis output revealed that "I would like to study science at university" was a problematic item. Although it might be expected that a student who perceives themselves to be capable in science would be more likely to go on to further their science education than one who struggles with the subject, this item was considered to related more to their future aspirations than to be related to their ability. Consequently, we replaced the item with a lower loading: "In my science class, I understand everything" (0.629). The Reliability analysis with this item included produced both a higher Cronbach's alpha (0.898) and an acceptable Tukey index value as shown in Table 11. Thus, the individual item response scores may be added to produce a scale that appears to measure their perception of their ability in science.

**Table 10.** Factor 5 – Ability in Science

Item	Loading	Alpha	Tukey Index
I find science difficult (reversed)	0.834	0.898	0.776
Science is one of my best subjects	0.801		
I am just not good at science (reversed)	0.781		
I get good marks in science	0.719		
I learn science quickly	0.707		
In my science class, I understand everything	0.629		

As before, the construct validity of the scale was assessed by computing a One-way ANOVA with post-hoc multiple comparisons (SNK). The 6-item scale of Ability in Science was employed as the DV and the item "I feel helpless when doing science" was used as the IV. It was hypothesized that should a student feel confident in their ability in science as measured by the scale, they would likely disagree with the statement that they feel helpless when doing science. Similarly, should a student feel they have low ability in science, they would tend to agree with the statement that they feel helpless.

**Table 11.** Post-hoc multiple comparisons output (SNK) for the Ability in Science Scale

IV= I feel helpless when doing science	N	1	2	3	4
Strongly Agree	7	2.2381			
Agree	17	2.4216	2.4216		
Undecided	59		2.8220		
Disagree	137			3.3297	
Strongly Disagree	99				4.0522
Sig.		.398	0.066	1.000	1.000

Subset for alpha = 0.05

Consistent with expectation, the results of the One-way ANOVA shown in Table 11 demonstrate that there is a highly significant difference in the mean scores on the Ability in Science scale of the students who indicated they do not feel helpless in science compared to those who indicated that they do, ( $F(4, 314) = 44.243, p < 0.001$ ). Here, students who agreed with the statement of helplessness had significantly *lower* mean scores in comparison to those students who disagreed. These results suggest that this particular scale possesses high construct validity and can measure students' perceptions of their ability in science.

### Factor 6: Future Aspirations in Science

This factor (Factor 4, darker grey in Table 1) is interpreted to describe students' intentions towards using science in the future. The highest loading item was "I would like to study science at University" (0.665) and the second highest loading item was "I would like to have a job working in science" (0.647). The five items listed in Table 12 primarily describe the levels to which students intend to pursue science.

Reliability analysis of these items produced a high Cronbach's alpha (0.923) and an adequate value for the Tukey Index (0.986). Thus, the items demonstrate a high internal consistency and Tukey's test of additivity was not significant, therefore, response scores of the individual items can be added together to produce a scale that appears to measure what students intend to do with science in the future.

**Table 12.** Factor 6 – Future Aspirations in Science

Item	Loading	Alpha	Tukey Index
I would like to study science at University	0.665	0.923	0.986
I would like to have a job working in science	0.647		
I would like to study more science in the future	0.555		
I would like to become a scientist	0.514		
I would like to do more science at school	0.491		

To assess the construct validity of this scale, a One-way ANOVA with post-hoc multiple comparisons was again computed. The 5-item Future Aspirations in Science scale was computed as the DV and "the science we learn at school is relevant to my future" was used as the IV.

It was hypothesized that should students agree that they would like to continue with science in the future, they would tend to agree that the science they learn at school is relevant to their future. Likewise, if students display no intentions of continuing with science in the future, they would tend to disagree with the IV statement.

**Table 13.** Post-hoc multiple comparisons output (SNK) for the Future Aspirations in Science Scale

IV= The science we learn at school is relevant to my future	N	1	2	3	4	5
Strongly Agree	80	4.2225				
Agree	124		3.5919			
Undecided	73			3.0785		
Disagree	22				2.6545	
Strongly Disagree	14					2.0857
Sig.		1.000	1.000	1.000	1.000	1.000

Subset for alpha = 0.05

Consistent with this prediction, One-way ANOVA results demonstrate that there is a highly significant difference in the mean scores on the Future Aspirations in Science scale of the students who indicated that science at school is relevant to their future, compared to those who indicated it is not, ( $F(4, 314) = 40.311$ ,  $p < 0.001$ ). Results from the post-hoc multiple comparisons are presented in Table 13 where the significant differences amongst students' responses is clearly shown. Students who agreed with the IV statement about the relevance of science to their future, had significantly higher mean scores in comparison to those who disagreed. These clear cut results are indicative that this scale possesses high construct validity and can be used to reliably measure students' intentions for science in the future.

#### Factor 7: The Benefits of Science

Factor 7 (Factor 5 in Table 1) is interpreted to mean what students perceive to be the “Benefits of Science” to both themselves and to the wider society. The two strongest loading items are: “Science and technology makes our lives easier and more comfortable” (-0.805) and “Science and technology is important for society” (-0.772).

Originally, this factor only had four items. The fourth low-loading item was “I copy notes the teacher gives me”. This was considered to be unrelated to the nature of this factor and so was not included in the Reliability Analysis. Nonetheless the scale yielded a high Cronbach's alpha, thus demonstrating high internal consistency as well possessing an adequate Tukey index value of 0.861 as shown in Table 14.

**Table 14.** Factor 7 – Benefits of Science

Item	Loading	Alpha	Tukey Index
Science and technology makes our lives easier and more comfortable	-0.805	0.810	0.861
Science and technology is important for society	-0.772		
The benefits of science are greater than the harmful effects	-0.620		

To assess the construct validity of the scale formed by adding the item scores, a One-way ANOVA with post-hoc multiple comparisons (SNK) was computed. The 3-item scale formed was computed as the DV, and the item “the science we learn at school will help me understand environmental issues” was used as the IV.

It was hypothesized that should a student agree that science is beneficial to society as indicated by their scale score, they would also tend to agree that the science they learn at school will help them to understand environmental issues. Similarly, should a student believe that science is not beneficial, they would tend to disagree that the science they learn helps them to understand environmental issues.



**Table 15.** Post-hoc multiple comparisons output (SNK) for the Benefits of Science Scale

<b>IV= The science we learn at school will help me understand environmental issues</b>	<b>N</b>	<b>1</b>	<b>2</b>
Strongly Agree	9	4.4145	
Agree	61	4.2424	
Undecided	160	4.1896	
Disagree	11	3.8907	
Strongly Disagree	78		3.2593
Sig.		1.000	0.103

Subset for alpha = 0.05

Consistent with expectations, the results of the One-way ANOVA demonstrate that there is a significant difference in the mean scores on the Benefits of Science scale by the students who indicated that the science they learn in school is useful in everyday life compared to those who indicated that it is not ( $F(4, 314) = 7.701, p < 0.001$ ). Results of the multiple comparisons test are presented in Table 15. Here, students who Strongly Disagreed with the statement of interest, had significantly lower mean scores in comparison to all other students. Table 15 also shows that the remaining mean scores for the other responses in the IV, Strongly Agree, Agree, Undecided and Disagree are not significantly different from each other. The pattern of decreasing mean scores indicate that this particular scale possesses a degree of construct validity and can measure students' perception of the benefits of science.

#### **Factor 8: Personal Relevance of School Science**

The final factor (Factor 6 in Table 1) is interpreted to mean the "Personal Relevance of School Science". The two strongest loading items are: "The science we learn at school is useful in everyday life" (0.518) and "The science we learn at school helps me make decisions about my health" (0.476). The factor analysis originally generated six items with the lowest loading one being "I watch the teacher do an experiment" (0.344). As before, it was eliminated since it was not considered to be related to the same construct as the other items.

Consequently, the reliability analysis was computed for the five items shown in Table 16. These items yielded a Cronbach's alpha of 0.833, demonstrating a high internal consistency. While the Tukey index of 1.374 was greater than the criterion set (1.3), it was not considered to be of a large enough magnitude to merit undertaking after the mathematical transformations of the raw Likert scale scores that were investigated. That is to say, the mathematically transformed scale possessed the same characteristics as the one formed by adding the raw scores.

**Table 16.** Factor 8 – Personal Relevance of Science

<b>Item</b>	<b>Loading</b>	<b>Alpha</b>	<b>Tukey Index</b>
The science we learn at school is useful in everyday life	0.518	0.833	0.1.374
The science we learn at school helps me make decisions about my health	0.476		
The science we learn at school is relevant to my future	0.467		
The science we learn at school deals with things I am concerned about	0.411		
The science we learn at school helps me understand environmental issues	0.354		
The science we learn at school is useful in everyday life	0.518		

The construct validity of the scale was assessed using the same method as before, viz., a One-way ANOVA with post-hoc multiple comparisons. The item "We learn interesting things in science lessons" was employed as the IV with the scale score for the Benefits of School Science scale as the DV. It was hypothesized that should a student feel that school science is personally relevant to them as measured by their scale score, they would also agree that they learn interesting things in their science lessons. Likewise, should a student not perceive any personal relevance of their science lessons, they would tend to disagree that the things they learn in school science are interesting.

Consistent with expectation, the results of the One-way ANOVA demonstrate that there is a highly significant difference in the mean scores on the Personal Relevance of Science scale of the students who indicated that they feel they learn interesting things in their science lessons compared to those who indicated that they do not, ( $F(4, 314) = 37.753, p < 0.001$ ). Results of the post-hoc multiple comparisons test are presented in Table 17. Here, students who Strongly Agreed with the statement of interest had significantly higher mean scores in comparison to all other students. Similarly, those who Agreed had significantly higher mean than those students who were Undecided or Disagreed. These results demonstrate that this particular scale possesses high construct validity and measures students' perceptions of the personal relevance of the science that they learn at school.

**Table 17.** Post-hoc multiple comparisons output (SNK) for the Personal Relevance to Science scale

<b>IV= We learn interesting things in science lessons</b>	<b>N</b>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>
Strongly Agree	9	4.1735			
Agree	14		3.5806		
Undecided	27			3.1852	
Disagree	186				2.6000
Strongly Disagree	83				2.3333
Sig.		0.171	1.000	1.000	1.000

Subset for alpha = 0.05

### **DISTRIBUTIONS OF THE SCALES, SKEWNESS AND KURTOSIS**

In order to assess the normality of the distributions more thoroughly and identify whether the factors are susceptible to floor or ceiling effects, the skewness and kurtosis of each factor was measured. However, it must be emphasized that the participants for the data collection upon which the Factor Analysis was computed do not present a random sample. The respondents were typically from top classes (highest ability) and had good teachers who commonly engage with astronomy and inquiry-based learning. It is therefore not surprising that students had predisposed positive attitudes towards science and their learning experiences.

**Table 18.** Skewness and Kurtosis Values for each of the Eight Factors

<b>Factor</b>	<b>Skewness</b>	<b>Std Error of Skewness</b>	<b>Kurtosis</b>	<b>Std Error of Kurtosis</b>
Interest in Astronomy	-0.610	0.137	0.441	0.272
Science Outside of School	-0.235	0.137	0.138	0.272
Practical Work in Science	-0.784	0.137	0.635	0.272
Teacher's Actions	-1.106	0.137	1.761	0.272
Ability in Science	-0.224	0.137	-0.747	0.272
Future Aspirations in Science	-0.502	0.137	0.020	0.272
Benefits of Science	-1.319	0.137	2.653	0.272
Personal Relevance to Science	-0.636	0.137	1.459	0.272

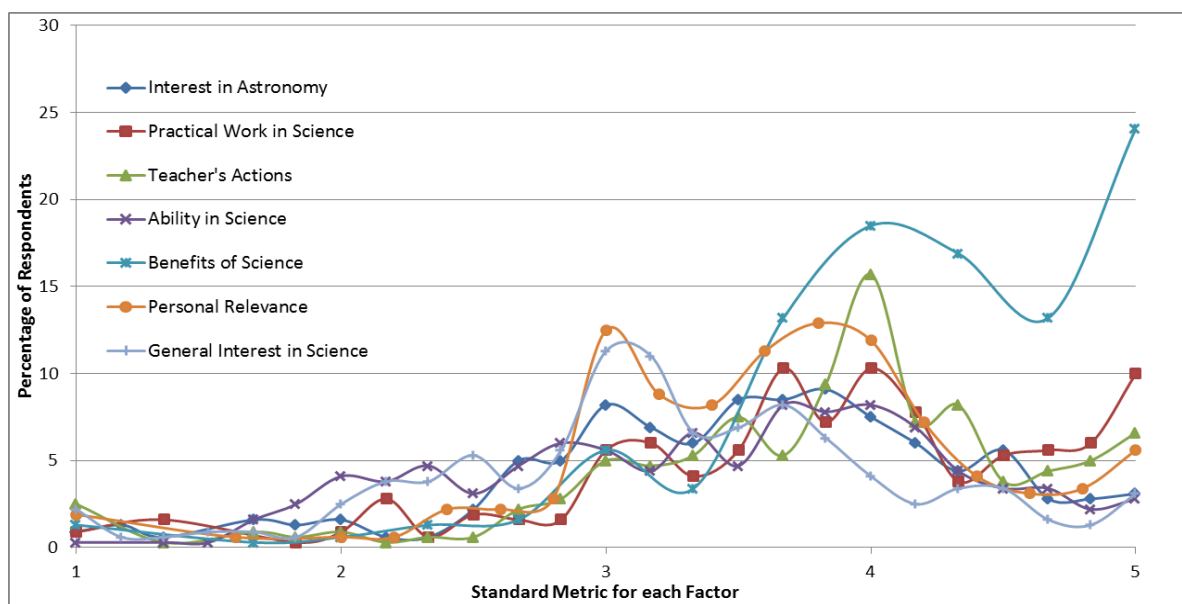
The results displayed in Table 18 and plotted onto the histograms in Figure 3 indicate that the data for the eight factors was negatively skewed, demonstrating that higher scores were more common, and not unexpected, given the characteristics of the respondents.

The six best scales with lower departures from normality are: Science Outside of School (mean score = 3.23, SD = 0.860), Ability in Science (mean score = 3.39, SD = 0.877), Interest in Astronomy (mean score = 3.51, SD = 0.845), Personal Relevance of School Science (mean score = 3.62, SD = 0.764), Practical Work in Science (mean score = 3.81, SD = 0.864) and Future Aspirations in Science (mean score = 3.49, SD = 0.965). These factors all possess skewness values within the range of  $\pm 1$  and can be considered not to be limited by floor or ceiling effects. That is to say, each factor has room for respondents to move upwards and/or downwards on the scales if an intervention is employed that is designed to shift their perceptions.

The more skewed distributions were the Benefits of Science (mean score = 4.16, SD = 0.773) and Teacher's Actions (mean score = 3.77, SD = 0.843), where the skewness values exceeded the  $\pm 1$ . These are the more problematic scales within the ASSA but the reasons for these distributions are explicable in terms of the respondents being from higher ability classes and probably being taught by better teachers as is the normal case in Australian schools.

Figure 3 presents a series of line graphs for each of the scales in terms of the percentages of participants' responses that rated across 1 (Strongly Disagree) through 5 (Strongly Agree). Here, the distribution of the Benefits of School Science scale in particular is noticeably skewed with almost 25% of respondents strongly agreeing about the benefit. This scale is the weakest one comprising only three items. If used with higher ability groups, it is unlikely to show any significant improvement.

**Figure 3.** Line graphs of percentages of students responses at 1=Strongly Disagree, through 5=Strongly Agree



Thus, only two of the eight scales present themselves as slightly problematic and although they exceeded a preferred skewness of  $\pm 1$ , they were positioned within  $\pm 2$ . Under circumstances where a random sample is selected, we may expect that these two scales will present better distribution characteristics for measuring students' attitudes and not be limited by floor or ceiling effects rendering them more sensitive to change. This aspect, however, will have to be investigated further.

These results indicate that the ASSA provides a stronger set of scales than the ones derived by Zeilik. In addition, the ASSA also offers an instrument that is suitable for use in high school classrooms and not restricted to undergraduate students. The next phase of our research will involve random sampling of participants so that results are not so constrained by selection effects. This will allow more powerful analyses capable of detecting changes attributable to interventions. Such procedures include Analysis of Variance with repeated measures on the occasion of testing. It should be noted by those who uses this instrument in this way that there are eight scales that can be analyzed separately or together depending on the characteristics of the distributions. If analyzed separately, a protection should be employed to safeguard against a Type I error by dividing the usual p-value of 0.05 by eight (8) to yield a value of 0.00625 below which significance could be claimed, or alternatively modifying this p-value if, as indicated above, the scales are correlated. A useful location for calculating the modified p-value using either a Bonferroni, modified Bonferroni or Sidak correction is here:

(<http://www.quantitativeskills.com/sisa/calculations/bonfer.htm>).

In contrast, Multivariate normality will allow an omnibus Multivariate Analysis of Variance with repeated measures on the occasion of testing. This approach is much more powerful and allows for interactions amongst the scales to be investigated. The mathematical assumptions of MANOVA, however, are much more restrictive. That is to say, on each occasion of testing, each of the scales must be normally distributed together with the distributions of eight scales all collapsed together on each occasion of testing being normally distributed plus, the distribution of all of the variables on all occasions of testing also being normally distributed. Fortunately, one statistic can be requested in SPSS that tests this homogeneity of covariance matrices in one swoop: Box's M. This statistic should not be significant given that it tests the null hypothesis that the covariance matrices are equal. If this is achieved (i.e., the statistic is not significant) an omnibus MANOVA may be computed. If, however, it is significant, one is either left with mathematical transformations of the ill-behaved individual scales or resorting to computing multiple univariate ANOVAs with repeated measures and foregoing any investigations of possible interactions.

## CONCLUSION

In this paper, we have presented a short history of student attitude measurement in astronomy, the rationale for and the construction of a new astronomy student attitude instrument. One clear factor is implied in the exploratory factor analysis that appears to be able to be split into eight sub-domains of interest and attitude yielding reliable and valid scales. This instrument is intended to be retested and revalidated with a variety of different groups of users to test its external and ecological validity as the current version was trialed purely on high school students. Most of the projects that claim to want to shift student attitudes do tend to focus on high school though, although there is likely a use for this attitude tool at university level, especially for high enrolment Astro101 courses in the USA.

If you intend using the instrument, we encourage you to contact us for a copy and the procedures for using it.

## AUTHOR BIOGRAPHIES

**Sophie Bartlett** is a PhD candidate at Cardiff University in Wales, United Kingdom. Her research focuses on promoting high school students' engagement and confidence in school physics through contexts of astronomy. Email: [sophie-bartlett@outlook.com](mailto:sophie-bartlett@outlook.com) (contact author)

**Michael Fitzgerald** is an ARC DECRA Research Fellow at the Edith Cowan Institute for Education Research in Perth, Western Australia. His research and practical interests lie in the incorporation of authentic astronomy research, particularly that involving use of robotic telescopes, into the high school classroom as a means to improve science student engagement and retention. Email: [mfitzasp@gmail.com](mailto:mfitzasp@gmail.com)

**David McKinnon** is a Professor of Education in the School of Education, Edith Cowan University at Perth, Western Australia. He has a long esteemed history in Science Education research with a particular focus on Astronomy Education combined with a successful career in pre-service teacher education that actually makes change happen. Email: [d.mckinnon@ecu.edu.au](mailto:d.mckinnon@ecu.edu.au)

**Lena Danaia** is an Associate Professor in the School of Teacher Education at Charles Sturt University in Bathurst, New South Wales, Australia. She specializes in science and technology curriculum and pedagogy, and is experienced in mixed-methods research. Her work has involved investigating the impact of many large-scale education projects on jurisdictions. Email: [ldanaia@csu.edu.au](mailto:ldanaia@csu.edu.au)

**Jasmina Lazendic-Galloway** is a Lecturer in the School of Physics & Astronomy at Monash University in Melbourne, Victoria, Australia. Broadly her research interest covers astronomy and physics education, STEM education with a particular focus on assessment and feedback techniques, work integrated learning and teachers' education. Email: [jasmina.lazendic-galloway@monash.edu](mailto:jasmina.lazendic-galloway@monash.edu)

## REFERENCES

- Adams, W. K., Perkins, K. K., Dubson, M., Finkelstein, N. D., & Wieman, C. E., (2005). *Proceedings of AIP Conference: The design and validation of the Colorado Learning Attitudes about Science Survey*. 790(1).
- American Association for the Advancement of Science (AAAS), (1990). *Science for all Americans: Project 2061*, New York: Oxford University Press.
- Bailey, J. M., & Lombardi, D., (2015). Blazing the trail for astronomy education research, *Journal of Astronomy & Earth Sciences Education*, 2(2), 77.
- Beare, R., (2007). Investigation into the potential of investigative projects involving powerful robotic telescopes to inspire interest in science. *International Journal of Science Education*, 29(3), 279-306.
- Bektasli, B., (2013). The effect of media on preservice science teachers' attitudes towards astronomy and achievement in astronomy class, *The Turkish Online Journal of Educational Technology*, 12(1), 139-146.
- Bektasli, B. (2016). *The Relationship between preservice science teachers' attitude toward astronomy and their understanding of basic astronomy concepts*. *International Journal of Progressive Education (IJPE)*, 12 (1), 108-116. ISSN 1554-5210.
- Bilici, S. C., Armagan, F.O., Cakir, N., & Yuruk, N., (2012). Adaption of astronomy attitude scale into Turkish: The Validity and Reliability Analysis, *Journal of Turkish Science Education*, 9(2), 128-131.
- Danaia, L., McKinnon, D. H., Parker, Q., Fitzgerald, M., & Stenning, P., (2012). Space to Grow: LCOGT.net and improving science engagement in schools, *Astronomy Education Review*, 11, 010106.
- Danaia, L., McKinnon, D. H., & Fitzgerald, M. (2017). Ideal pictures and actual perspectives of junior secondary school science: comparisons drawn from Australian students in an astronomy education programme. *Research in Science & Technological Education*, 35(4), 445-460.
- Ferlet, R. (2008). *Hands-On Universe - Europe (EU-HOU)*. In: *Innovation in Astronomy Education*. Cambridge University Press, pp. 23-26.
- Fitzgerald, M. T., Hollow, R., Rebull, L. M., Danaia, L., & McKinnon, D. H., (2014). A review of high school level astronomy student research projects over the last two decades, *Publications of the Astronomical Society of Australia*, 31, e037.
- Fitzgerald, M., McKinnon, D. H., & Danaia, L. (2015). Inquiry-based educational design for large-scale high school astronomy projects using real telescopes. *Journal of Science Education and Technology*, 24(6), 747-760.
- Fitzgerald, M., McKinnon, D. H., Danaia, L., & Deehan, J. (2016). A large-scale inquiry-based astronomy intervention project: impact on students' content knowledge performance and views of their high school science classroom. *Research in Science Education*, 46(6), 901-916.
- Fitzgerald, M. T., McKinnon, D.H., Danaia, L., Cutts, R., Salimpour, S., & Sacchi, M. (2018). *Proceedings of the Robotic Telescopes, Student Research and Education Conference: Our Solar Siblings. A high school focused robotic telescope-based astronomy education project*. San Diego, Jun 18-21 2017.
- Gomez, E. L., & Fitzgerald, M. T. (2017). *Robotic telescopes in education*. *Astronomical Review*, 13(1), 28-68.
- Goodrum, D., Hackling, M., & Rennie, L. (2001). *The status and quality of teaching and learning of science in Australian schools*. Canberra: Department of Education, Training and Youth Affairs.
- Goodrum, D., Druhan, A., & Abbs, J. (2012). *The status and quality of year 11 and 12 science in Australian schools*. Canberra: Australian Academy of Science.
- Gogolin, L., & Swartz, F. (1992). A quantitative and qualitative inquiry into the attitudes toward science of nonscience college students. *Journal of Research in Science Teaching*, 29(5), 487-504.
- Havard, N. (1996). Student attitudes to studying A-level sciences. *Public Understanding of Science*, 5(4), 321-330.
- Hollow, R., Hobbs, G., Champion, D., Amy, S., Khoo, J., Chapman, J., Mulcahy, M., Alem, L., Krumm-Heller, A., McKinnon, D. H., Danaia, L. J., Jenet, F., & Carr, M., (2008). PULSE@Parkes: Pulsar Observing for High School Students. In: *ASP Conference Series: Preparing for the 2009 International Year of Astronomy: A Hands-On Symposium*, 1-5 June, 2008, St. Louis, Missouri, USA.
- Kadooka, M.A., Bedient, J., Hu, S., Hemphill, R., & Meech, K. J. (2008). US student astronomy research and remote observing projects. In: *Innovation in Astronomy Education*. Cambridge University Press, 219-224.
- Kareva, A., Miller, S., Foster, A. & James, C.R. (2014). The power of affect: A comparison of Astronomy/Science attitudes among students and secondary teachers. *Bulletin of the American Astronomical Society*, 223, 448.11 [poster].
- Kind, P., Jones, K., & Barmby, P. (2007). Developing attitudes towards science measures. *International Journal of Science Education*, 29(7), 871-893.
- Lazendic-Galloway, J., Fitzgerald, M. T., & McKinnon, D. H., (2016). Implementing a studio-based flipped classroom in a first year astronomy course. *International Journal of Innovation in Science and Mathematics Education*, 24 (5), 35-47.
- Millar, R., & Osborne, J. (1998). *Beyond 2000: science education for the future* (The report of a seminar series funded by the Nuffield Foundation), London: King's College London, School of Education.
- Miller, S. T., & Redman, S. L. (2010). Improving instructor presence in an online introductory astronomy course through video demonstrations, *Astronomy Education Review*, 010115-1.
- Newsam, A., Carter, D., & Roche, P. (2003). The Liverpool and Faulkes Telescopes. In: *The Future of Small Telescopes in the New Millennium. Astrophysics and Space Science Library*. Springer Netherlands (Kluwer Academic Publishers),

- Netherlands, 391-410.
- Platco, N. L. (2005). *A comparative study of the effectiveness of star show versus participatory oriented planetarium lessons in a middle school Starlab setting*, Doctoral Thesis, Temple University.
- Redish, E. F., Steinberg, R. N., & Saul, J. M. (1998). Student expectations in introductory physics. *American Journal of Physics*, 66, 212-224.
- Richwine, P. L. (2007). *The Impact of Authentic Science Inquiry Experiences Studying Variable Stars on High School Students' Knowledge And Attitudes About Science And Astronomy And Beliefs Regarding The Nature of Science*. Ph.D. Dissertation, University of Arizona.
- Schau, C., Stevens, J., Dauphinee, T. L., & Del Vecchio, A. (1995). The development and validation of the survey of attitudes towards statistics, *Educational and Psychological Measurement*, 55(5), 868-875.
- Slater, S. J., Slater, T. F., Heyer, I., & Bailey, J. M. (2015). *Conducting Astronomy Education Research: A Primer*, 2nd edition. Hilo: Pono Publishing, ISBN: 978-1515025320.
- Slater, S. J., Slater, T. F., & Shaner, A., (2008). Impact of backwards faded scaffolding in an astronomy course for pre-service elementary teachers based on inquiry, *Journal of Geoscience Education*, 56(5), 408-416.
- Slater, S. J., Tatge, C. B., Bretones, P. S., Slater, T. F., Schleigh, S. P., McKinnon, D., & Heyer, I., (2016). *iSTAR First Light: Characterizing astronomy education research dissertations in the iSTAR Database.*, *Journal of Astronomy & Earth Sciences Education*, 3(2), 125-140.
- Slater, T. F., Safko, J. L., & Carpenter, J. R. (1999). Long-term attitude sustainability from a constructivist-based astronomy-for-teachers course. *Journal of Geoscience Education*, 47(4), 366-368.
- Slater, T. F., & Jones, L.V. (2004). Assessment of an internet-delivered interactive approach to introductory astronomy for non-science majors, *Astronomy Education Review*, 1(3), 17-25.
- Turk, C., & Kalkan, H. (2015). Astronomy Attitude Scale: Development, validity and reliability, *Journal of Studies in Education*, 5(4), 23-50.
- Tyler, R. (2007). *Re-imagining science education: Engaging students in science for Australia's future*. Australian Education Review. Camberwell, VIC: Australian Council for Educational Research.
- Ucar, S., & Demircioglu, T. (2011). Changes in preservice teacher attitudes toward astronomy within a semester-long astronomy instruction and four-year-long teacher training programme. *Journal of Science Education and Technology*, 20(1), 65-73.
- Weingburgh, M. H., & Steele, D. (2000). The modified Attitudes Toward Science Inventory: Developing an instrument to be used with fifth grade urban students, *Journal of Women and Minorities in Science and Engineering*, 6(1), 87-94.
- White, B., Elby, A., Frederiksen, J., & Schwarz, C., (1999). *The Epistemological Beliefs Assessment for Physical Science*, AERA Annual Meeting (Montréal, 1999).
- Wittman, D. (2009). Shaping attitudes toward science in an introductory astronomy course. *The Physics Teacher*, 47(9), 591-594.
- Zeilik, M., Schau, C., Mattern, N., Hall, S., Teague, K. W., & Bisard, W. (1997). Conceptual Astronomy: A novel model for teaching postsecondary science courses, *American Journal of Physics*, 65, 987-996.
- Zeilik, M., Schau, C., & Mattern, N. (1999). Conceptual Astronomy. II. Replicating conceptual gains, probing attitude changes across three semesters, *American Journal of Physics*, 67, 923-927.
- Zeilik, M., Bisard, W., & Lee, C. (2002). Research-Based reformed astronomy: will it travel?, *Astronomy Education Review*, 1(1), 33-46.
- Zeilik, M., & Morris, V. K. (2003). An examination of misconceptions in an astronomy course for science, mathematics, and engineering majors, *Astronomy Education Review*, 1(2), 101-119.