

Measuring High School Students' Science Identities, Expectations of Success in Science, Values of Science and Environmental Attitudes: Development and Validation of the SIEVEA Survey

Rosa Aghekyan*

Rutgers Graduate School of Education, New Brunswick, New Jersey, USA

*Corresponding Author: rosa.aghekyan@gse.rutgers.edu

ABSTRACT

This article describes the development and validation of an instrument called science identities, expectations of success in science, values of science, and environmental attitudes (SIEVEA). The instrument is a convenient online survey that can be used to measure students' SIEVEA. First, 1764 high school students from 11 school districts in New Jersey, Pennsylvania, and Connecticut took the SIEVEA. The collected data were analyzed using descriptive statistics and exploratory factor analysis (EFA). The EFA results provided useful insights into the factor structure of the data and led to the formation of three candidate models. All three models were evaluated based on their fit to data, their alignment with the research constructs, and their factor loadings. Because of this evaluation, the three-factor model was selected as the final model. Next, the instrument was validated through confirmatory factor analysis using a new sample data made up of responses of 1495 high school students. The developed instrument is valuable for both researchers and educators, as it allows for measurement of constructs that play an important role in successful science learning.

KEY WORDS: science identities; expectations of success; environmental attitudes; values of science; exploratory factor analysis; confirmatory factor analysis

INTRODUCTION

Skills required of workers in the 21st century differ significantly from those needed in preceding centuries (Bybee & Fuchs, 2006; Silva, 2009). Indeed, to be successful in today's economy, one must be technologically savvy, skillful in data analysis, and scientifically literate. Even though science, technology, engineering, and mathematics (STEM) have played a prominent role for decades in preparing students to become part of the modern workforce, STEM education may be at its zenith. One important reason behind this scrutiny is the continued lag of the United States of America (U.S.) students behind their international peers (OECD, 2012). The national assessments like the 2009 National Assessment of Educational Progress (NAEP) report also indicate low proficiency level (National Center for Education Statistics, 2011). In addition, many students lack interest pursuing STEM-related careers (U.S. Department of Education, n.d.).

What can be done to improve students' attitudes toward science learning and nudge them toward choosing science careers? Motivation and interest are some of the key factors in initiating and sustaining students' interest in science learning (Eccles et al., 1998; Pintrich, 2003). In addition, students with strong science identities are more likely to participate and succeed in science

and math classes compared to their peers with lower science and math identities (Gresalfi, 2009; Sford & Prusak, 2005). Since students' science identities, expectations of success in science, and values of science play such an important role in their science learning, it is imperative to have instruments that can be used to measure these constructs in a simple, yet reliable way. This was the primary motivation behind this study. In addition, one of the intended uses of the designed survey was to measure how motivated students were in learning environmental science. Therefore, the survey also needed to include items for measuring students' environmental attitudes, as environmental issues relate to students' science learning.

The review of literature regarding current and past research uncovered that various instruments have been developed for measuring either motivational constructs in science learning (Glynn & Koballa, 2006; Mubeen & Reid, 2014; Plante et al., 2012; Swarat et al., 2012) or environmental identities, attitudes or literacy (Blatt, 2014; Eilam & Trop, 2012). These findings are summarized in Table 1.

Unfortunately, these instruments do not assess motivation in science learning and environmental attitudes in a single survey. Why is it useful to have a survey that can be used to measure these constructs together, rather than using multiple surveys? First, the logistics of administering one combined survey is much easier than

Table 1: Questionnaire or interview-based studies of motivational and environmental constructs

| Study | Discussion |
|--------------------------------|---|
| Blatt (2014) | The study measured 10 high school students' environmental identities while conducting qualitative analysis. This study used student interviews (at the beginning, middle, and end of the semester) and student reflections. |
| Glynn and Koballa (2006) | Developed Science Motivation Questionnaire designed for college and high school students. The survey had 20 questions; there were no questions about students' environmental attitudes. |
| Swarat et al. (2012) | Administered questionnaire focused on students' interest in science in diverse instructional episodes. The questionnaire measured students' interest in specific biology topics such as cells, ecosystems, diversity of living things, and human body systems. |
| Glynn et al. (2011) | Updated SMQ-II. This questionnaire was the updated form of the SMQ. It added five more questions. Furthermore, it had no questions pertaining environmental attitudes. |
| Johnson and Manoli (2011) | Developed 2-MEV questionnaire, which measured adolescences' environmental attitudes. Survey did not have any constructs measuring students' motivation of science learning. |
| Eilam and Trop (2012) | Investigated the relationship between environmental attitudes and environmental behaviors of students and their parents. Questionnaire was used to understand the processes that influence environmental behaviors and environmental attitudes. The questionnaire is focused on environmental education. It does not have science education component measuring motivation. |
| Zecha (2010) | Analyzed students' environmental awareness. Questionnaire was focused on assessing students' environmental knowledge, attitudes, and actions. |
| Plante et al. (2012) | Explored the relationship between expectancy-value and achievement goal theories and their role in predicting achievement behaviors. Two questionnaires assessed various motivational constructs including students' expectancies and values in mathematics and language arts. |
| Bradley et al. (1999) | Studied the relationship between environmental knowledge and environmental attitude of high school students. Utilized questionnaires as a pre-test and post-test for assessing aforementioned constructs before and after intervention. |
| Fortus and Vedder-Weiss (2014) | Examined students' CM for science learning. Survey was developed to measure CM constructs. The survey had only two items for assessing students' environmental attitudes. The environmental attitudes construct was blended with science constructs, and the second item had the same question as the first one; it was merely negatively worded. |

SMQ: Science Motivation Questionnaire, CM: Continued motivation

that of using multiple instruments. Second, a short survey helps avoid “respondent fatigue” phenomenon (Lavrakas, 2008) by reducing the possibility of participants getting bored and skimming over questions or leaving the survey incomplete. Thus, it has a greater chance of being complete, accurate, and candid than larger surveys with lots of questions. Third, this instrument will especially be useful for researchers who study students' environmental attitudes in conjunction with their science learning motivation and science identities, as it allows for collection of pertinent data in a single dataset. Finally, unlike the Swarat et al. (2012) questionnaire, which measured students' interest in specific topics within a single domain, and the Matsui et al. (1990) study, that asked students to rate how much they liked math, this instrument aimed at measuring students' interest in various science subjects.

LITERATURE REVIEW

Motivation can be defined as the study of why people think and act the way they do. In an academic context, motivation plays an important role in explaining the reasons behind why some students complete the task despite its arduous nature, while others give up on easy tasks (Graham & Weiner, 1996). Motivational theories claim that motivational factors influence cognitive processes that, in turn, affect performance (Eccles

& Wigfield, 2002). Therefore, understanding high school students' (students aged 14 to 17) science identities and their motivation in science can aid in explaining phenomena such as disliking science or not doing well in science.

According to Eccles et al. (1998) Expectancy-Value motivational framework, students' achievement behavior is primarily influenced by two factors: Their expectation of success in completing the task and the task's value. In a science learning context, expectancies indicate how students assess their abilities to complete science tasks (Fredricks and Eccles, 2002). Likewise, the task value component of motivation considers how students view the science task in relation to their goals and aspirations. The value of the science task can be further broken into four components: Intrinsic value/interest, attainment value, utility value, and cost. Intrinsic value/interest is the enjoyment that one receives as a result of engaging in a scientific activity (Eccles, 2009; Wigfield et al., 2009). Attainment value is the intrinsic importance of the task to individuals (Eccles, 2005) or the importance of being good at something (Eccles & Wigfield, 1995).

Utility value is how an individual relates current activities to goals, such as future occupation and career (Wigfield & Eccles, 1992). Cost encapsulates many elements, including

lost opportunities due to making one choice instead of another, the negative aspects of engaging in a task, like fear of failure (Plante et al., 2012), and the amount of effort devoted toward accomplishing the task (Wigfield & Eccles, 1992).

Students' science identities also play an important role in their participation in science and math classes (Gresalfi, 2009; Sfarid & Prusak, 2005). Identity is a multi-faceted notion in educational research, with various definitions existing across a wide range of conceptual frameworks and constructs. Indeed, Gee (2000) defined identity as being seen as a certain "kind of person." This view of identity focuses on the individual's performance in society and how one is recognized by himself/herself and by others. According to Gee, individuals can have multiple identities. In a science context, this translates into the individual's science identity defined as that of a "science person."

Identities can also be defined as narratives enacted in time, space, and relationships (Moje et al., 2007). Tucker-Raymond et al. (2007) argued that narratives as enactments of identities are not "pure," meaning they are not just "discourse" and no "action," but that they rather include both. The interactions during these enactments are simply subject positionings; therefore, they are referred to as subjectivities. Since the enactments happen in different settings, identities can be referred to as 'multimodal' narratives (Tucker-Raymond et al., 2007).

Learning and identities are related to each other (Tucker-Raymond et al., 2007). Learning involves more than the construction of information and juxtaposing of different pieces of information to paint the whole picture. It is also about people thinking about themselves in relation to the process of learning concepts (Moje et al., 2007).

Carlone and Johnson (2007) developed a model of science identity by projecting students' science identities onto three overlapping dimensions: Science competence, science performance, and recognition as a science person. According to this model, to be considered as an individual with a strong science identity, a student should rate himself/herself highly and be rated highly by others in all three dimensions. Science identity is an important tool for understanding student science learning. By apprehending students' science identities, science educators can get a better understanding of their students' personalities and aspirations. Through this, they can develop meaningful and beneficial instruction (Kane, 2016).

Various institutions emphasize the importance of environmental literacy and educating environmentally literate citizens (Jowett et al., 2014). Indeed, the goal of environmental education is to increase students' environmental literacy and raise their awareness of environmental problems. This is especially because people's habits and choices of lifestyle in the 21st century may result in the destruction of the environment and the diminishment of resources (Erdogan et al., 2012). It is important for students to familiarize themselves with their local environment first, before making judgments regarding global environmental issues, such as global warming and water/air pollution (Sobel, 1996).

Coyle's (2005) report regarding environmental literacy collected from 10 years of surveys not only showed that most Americans were environmentally illiterate but also concluded that even small changes in human behavior can have a huge impact on the environment. Like Project 2061's (AAAS, 1989) emphasis on the growth of scientific literacy, several studies highlight the importance of cultivating environmental literacy, where individuals gain better perception regarding the interaction between the humans and their surrounding environment (Osbaldiston, 2004).

The Science Identities, Expectations of Success in Science, Values of Science, and Environmental Attitudes (SIEVEA) Instrument

With this instrument, high school students' science identities, motivation in science, and environmental attitudes can be explored, and science instruction can be reshaped based on students' perceptions and responses regarding their science learning. The survey used a 5-point Likert scale (Likert, 1932) for the measurement of attitudes. The survey was web-based and published on the <http://www.qualtrics.com> web site. The first two questions of the survey captured students' favorite subjects and gender. The remaining items were designed to measure students' science identities, their expectations of success in science classes, how they value science, and their environmental attitudes. The survey's design was based on existing research about constructs and instruments for their measurement. The design incorporated modifications to existing instruments to accommodate the researcher's interests. The survey was piloted with 30 non-science major students.

The development of items was guided by literature regarding science identities (Carlone & Johnson, 2007), expectation of success (Eccles et al., 1998), values of science (Eccles & Wigfield, 1995), and environmental attitudes (Dunlap et al., 2000). Science identities were assessed by three items constructed according to the Carlone and Johnson (2007) "Science Identity" initial model. This model includes three overlapping dimensions: Competence, performance, and recognition. This study used performance and recognition dimensions. The expectation of success in science construct was adopted from the Eccles et al. (1998) expectancy-value framework, which accentuates the expectations of success as an integral part of the motivation. The survey items assessing students' expectations of success in science were modeled after the works of Plante et al. (2012) and Glynn et al. (2011).

The value of science construct was adopted from the Eccles et al. (1998) expectancy-value framework. This construct measures the importance of learning sciences (Eccles & Wigfield, 1995), as well as how individuals value an academic task (Plante et al., 2012). Questions used for measuring this construct were created based on MSLQ (Pintrich and De Groot, 1990), Science Motivation Questionnaire (SMQ) (Glynn & Koballa, 2006), and the survey implemented by Eccles and Wigfield (1995). The environmental attitudes construct was developed after reviewing existing literature regarding

measuring environmental attitudes (Dunlap et al., 2000). The Dunlap et al. (2000) revised New Ecological Paradigm scale was used as a framework. Figure 1 depicts survey items and their categories.

Since constructs such as motivation of science learning and environmental attitudes are not directly observable, they are considered latent variables (Glynn et al., 2011) or composite variables (Thompson, 2004). Table 2 summarizes how this study's constructs were mapped to questionnaire items and what literature and existing instruments were used in support of the mappings. The Lexile Framework for Reading system (MetaMetrics Inc., 2012) was used to analyze the readability of the survey's questions. The analyses showed that the survey's reading complexity did not exceed the reading ability of a typical high-school (Grade 9, students aged 14) student.

EXPLORATORY FACTOR ANALYSIS (EFA)

Data

A total of 1764 high school students from 11 school districts in New Jersey, Pennsylvania, and Connecticut (grades 9–12) completed the SIEVEA. However, only 1660 students' responses were complete and used for data analysis. Of these participants, 930 were girls and 827 were boys. Seven students chose not to report their gender. The participant schools were urban, suburban, and private high schools. The survey's data indicated a 92.3% completion rate. The data collection took place during 2014–2015 and 2015–2016 academic years.

Procedure

The survey's data were analyzed using descriptive statistics. Even though Item #9 (“I can use technology for learning

science content.”) had statistically significant correlations with other items in its group ($p < 0.01$), its Pearson correlation coefficients were noticeably smaller than those of others (the greatest coefficient was 0.30). Moreover, a quick test-run of the factor extraction process on survey's data showed that Item #9 did not factor well into any extracted factor: Its factor loadings were low (about 0.300). These results indicated that this item did not fit well within other items. Therefore, it made sense to drop it from the instrument and exclude its data from subsequent analyses.

After Item #9 was removed, an EFA was conducted on the remaining 12 items. EFA was carried out using SPSS (Version 24) statistical package. Before conducting factor analysis, item-to-item correlation was examined by conducting the Kaiser-Meyer-Olkin (KMO) test and Bartlett's test for sphericity. The value of KMO was 0.893 and Bartlett's test produced $p < 0.001$. These results indicated that the data were appropriate for conducting factor analysis (Glynn et al., 2011).

The choice of the most suitable method for EFA was based on the recommendation of Fabrigar et al. (1999) who argued that the maximum likelihood method should be used for relatively normally distributed data. The descriptive statistics of data (Table 3) indicated that item level data distributions did not significantly depart from the normal distribution. In addition, Varimax orthogonal rotation was chosen to maximize variance (Tabachnick and Fidell, 2001).

Factor Extraction

When implementing EFA, one of the most important decisions to make is the number of factors to retain. Horn (1965) proposed a method for factor retention called parallel analysis (PA). His approach computes a new set of eigenvalues

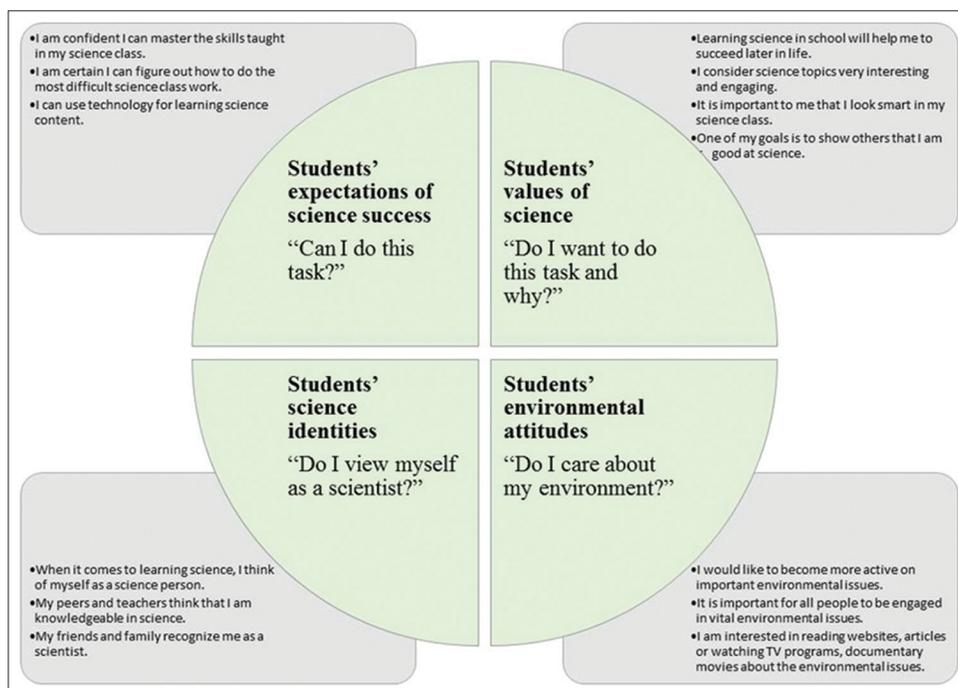


Figure 1: Survey categories and items

Table 2: Instrument items and developed constructs matrix

| Item | Construct | How the item was developed |
|--|------------------------------------|--|
| 3. Learning science in school will help me to succeed later in life. | Values of science | This question assesses the intrinsic value of science. It is modeled after items from MSLQ (Pintrich and De Groot, 1990). |
| 4. I am confident I can master the skills taught in my science class. | Expectations of success in science | Modeled after items from Plante et al. (2012) questionnaire and Glynn et al. (2011) SMQ-II. |
| 5. I consider science topics very interesting and engaging. | Values of science | This question assesses the intrinsic value of science. It is modeled after items from MSLQ (Pintrich and De Groot, 1990) and Eccles and Wigfield (1995). |
| 6. When it comes to learning science, I think of myself as a science person. | Science identity | Science Identity Model designed by Carlone and Johnson (2007). The question assesses the model's "recognition" dimension. |
| 7. My peers and teachers think that I am knowledgeable in science. | Science identity | Science identity model designed by Carlone and Johnson (2007). The question assesses the model's "recognition" and "performance" dimensions. |
| 8. I am certain I can figure out how to do the most difficult science classwork. | Expectations of success in science | Modeled after items from Plante et al. (2012) questionnaire and Glynn et al. (2011) SMQ-II. |
| 9. I can use technology for learning science content. | Expectations of success in science | Modeled after items from Glynn et al. (2011) SMQ-II. |
| 10. My friends and family recognize me as a scientist. | Science identity | Science identity model designed by Carlone and Johnson (2007). The question assesses the model's "recognition" dimension. |
| 11. It is important to me that I look smart in my science class. | Values of science | This question assesses the attainment value of science (Eccles and Wigfield, 1995). It was modeled after items from SMQ (Glynn and Koballa, 2006) questionnaire. |
| 12. I would like to become more active on important environmental issues. | Environmental attitudes | Modeled after survey items from Stern et al. (2011). This survey combined all questions assessing students' environmental concerns. |
| 13. One of my goals is to show others that I am good at science. | Values of science | This question assesses the attainment value of science. It is researcher-developed and is based on the existing literature and expectancy-value theory. |
| 14. It is important for all people to be engaged in vital environmental issues. | Environmental attitudes | Modeled after survey items from DiEnno and Hilton (2005). |
| 15. I am interested in reading websites, articles, or watching TV programs, documentary movies about the environmental issues. | Environmental attitudes | This item is researcher-developed. It is based on the revised NEP scale. |

NEP: New ecological paradigm, MSLQ: Motivated strategies for learning questionnaire, SMQ: Science motivation questionnaire

Table 3: Summary of item level descriptive statistics

| Item | Range | Mean | SD | Skewness | SE | Kurtosis | SE |
|------|-------|------|-------|----------|-------|----------|-------|
| 3 | 1-5 | 3.88 | 0.980 | -0.718 | 0.058 | 0.112 | 0.117 |
| 4 | 1-5 | 3.89 | 0.847 | -0.792 | 0.058 | 0.867 | 0.117 |
| 5 | 1-5 | 3.79 | 1.013 | -0.804 | 0.058 | 0.272 | 0.117 |
| 6 | 1-5 | 3.18 | 1.179 | -0.178 | 0.058 | -0.833 | 0.117 |
| 7 | 1-5 | 3.6 | 0.928 | -0.445 | 0.058 | -0.049 | 0.117 |
| 8 | 1-5 | 3.33 | 1.050 | -0.434 | 0.058 | -0.387 | 0.117 |
| 10 | 1-5 | 2.52 | 1.110 | 0.405 | 0.058 | -0.502 | 0.117 |
| 11 | 1-5 | 3.27 | 1.075 | -0.268 | 0.058 | -0.469 | 0.117 |
| 12 | 1-5 | 3.41 | 0.998 | -0.339 | 0.058 | -0.228 | 0.117 |
| 13 | 1-5 | 3.08 | 1.103 | -0.133 | 0.058 | -0.688 | 0.117 |
| 14 | 1-5 | 3.65 | 0.949 | -0.515 | 0.058 | 0.056 | 0.117 |
| 15 | 1-5 | 3.33 | 1.083 | -0.409 | 0.059 | -0.440 | 0.117 |

using randomly generated dataset with the same numbers of observations and variables as the original data. Then, it compares these eigenvalues to the ones computed by EFA. Last, if the eigenvalues based on the simulated data are greater than the matching eigenvalues from EFA, then corresponding EFA factors are dropped from the model. Otherwise, they are retained. Many studies confirmed that PA is the best method

for determining the number of factors during EFA (Zwick and Velicer, 1986). Its results are more accurate and reliable than those produced by other methods, including Kaiser's eigenvalue greater than one rule (Glorfeld, 1995; Kaiser, 1960).

For this study, SPSS program rawpar.sps developed by O'Connor (2000) was used to conduct PA. This procedure produced a scree plot of three data sets: Original data's eigenvalues, means of eigenvalues of random data, and their 95% confidence limits (Figure 2).

As the scree plot indicated, up to four factors could be extracted. Therefore, two more extractions were done using the maximum likelihood method forcing the extraction of three and four factors.

Model Selection

The factor extraction resulted in three candidate models: 2-factor model, 3-factor model, and 4-factor model. Table 4 contains the factor loadings for all three models.

To evaluate the model's fit to data, the study computed absolute and incremental close-fit indexes and compared them to established cutoff values (Gignac, 2009). For absolute close-fit, this study used the Root Mean Square Error of Approximation

Table 4: Factor loadings of SIEVEA models using maximum likelihood and varimax rotation

| Item | 2-Factor model | | 3-Factor model | | | 4-Factor model | | | |
|------|----------------|--------------|----------------|--------------|--------------|----------------|--------------|--------------|--------------|
| | C1 | C2 | C1 | C2 | C3 | C1 | C2 | C3 | C4 |
| 3 | 0.618 | 0.285 | 0.605 | 0.268 | 0.152 | 0.369 | 0.449 | 0.216 | 0.261 |
| 4 | 0.653 | 0.076 | 0.630 | 0.054 | 0.146 | 0.670 | 0.257 | 0.087 | 0.109 |
| 5 | 0.658 | 0.341 | 0.686 | 0.352 | 0.023 | 0.347 | 0.607 | 0.289 | 0.124 |
| 6 | 0.775 | 0.309 | 0.794 | 0.304 | 0.087 | 0.380 | 0.768 | 0.212 | 0.183 |
| 7 | 0.626 | 0.076 | 0.582 | 0.028 | 0.277 | 0.667 | 0.216 | 0.075 | 0.148 |
| 8 | 0.640 | 0.095 | 0.607 | 0.063 | 0.201 | 0.620 | 0.267 | 0.085 | 0.166 |
| 10 | 0.663 | 0.279 | 0.650 | 0.260 | 0.164 | 0.345 | 0.550 | 0.188 | 0.261 |
| 11 | 0.373 | 0.304 | 0.246 | 0.197 | 0.751 | 0.300 | 0.090 | 0.209 | 0.494 |
| 12 | 0.127 | 0.774 | 0.120 | 0.753 | 0.148 | 0.068 | 0.135 | 0.748 | 0.188 |
| 13 | 0.475 | 0.431 | 0.409 | 0.373 | 0.437 | 0.140 | 0.320 | 0.252 | 0.801 |
| 14 | 0.114 | 0.693 | 0.104 | 0.677 | 0.140 | 0.088 | 0.098 | 0.693 | 0.140 |
| 15 | 0.238 | 0.605 | 0.247 | 0.613 | 0.054 | 0.108 | 0.267 | 0.588 | 0.102 |

SIEVEA: Science Identities, Expectations of Success in Science, Values of Science, and Environmental Attitudes

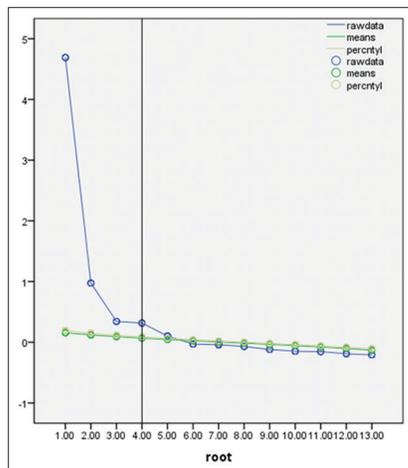


Figure 2: Scree plot of parallel analysis

(RMSEA; Browne & Cudeck, 1993). The incremental close-fit indexes used in this study were the Normal Fit Index (NFI; Bentler & Bonnett, 1980), the Tucker-Lewis Index (TLI; Tucker & Lewis, 1973), and the comparative fit index (CFI; Bentler, 1990).

The cutoff values of close-fit indexes are as follows. For absolute close-fit index RMSEA, values of 0.08 or less suggest an acceptable model fit (Hu & Bentler, 1999), whereas values >0.10 indicate a poor fit (Browne & Cudeck, 1993). For NFI, TLI, and CFI, an index value of 0.95 or higher indicates a good fit (Hu & Bentler, 1999; Thompson, 2004), whereas the cutoff value for an acceptable fit is 0.90 (Bentler & Bonett, 1980).

During the factor extraction, the Chi-square statistics and the degrees of freedom of the null and implied models were calculated (Table 5).

These statistics and degrees of freedom were used to calculate absolute and incremental close-fit indexes for all models (Table 6).

As these results showed, the 2-factor model failed to produce acceptable fit index values (RMSEA was too high and

Table 5: Chi-square statistics and degrees of freedom of null and implied models

| | Null model | 2-Factor model | 3-Factor model | 4-Factor model |
|----------|------------|----------------|----------------|----------------|
| χ^2 | 8053.806 | 761.683 | 470.022 | 172.123 |
| df | 66 | 43 | 33 | 24 |

Table 6: Calculated values of close-fit indexes

| Model | Absolute close-fit | | Incremental close-fit | |
|--------------|--------------------|------|-----------------------|------|
| | RMSEA | NFI | CFI | TLI |
| Two-factor | 0.100 | 0.91 | 0.91 | 0.86 |
| Three-factor | 0.089 | 0.94 | 0.95 | 0.89 |
| Four-factor | 0.061 | 0.98 | 0.98 | 0.95 |

incremental close-fit indexes were too low), whereas both the three-factor and the four-factor models had indexes within or close to their appropriate range/threshold values.

Next, all three candidate models were evaluated based on how well they were aligned with the research constructs. All models correctly placed the environmental attitude items into a separate factor. However, the models produced a less than perfect match for the remaining items measuring motivational and science identity constructs. The two-factor model bundled all these items together into a single factor. Even though the four-factor model produced one factor per construct, it placed several items in factors different from their predicted locations. For example, items three and five were placed into the “science identity” factor even though they were motivational items. The 3-factor model was able to place two values of science items into their own factor correctly. The remaining items were placed into a single factor, therefore, combining science identity and motivational items into one factor. Thus, even though no model produced a perfect match to the predicted factor structure, the 3-factor model did a better job than other two models.

Finally, factor loadings from Table 4 showed some uncertainty in choosing the factors for some items. In the three-factor model, Item #13 had comparable factor loadings for all three factors (0.409, 0.373, and 0.437). It was decided to pick factor #3 for this item since it was a science value item, and it made sense to combine it with the other science value item (Item #11), which had a high factor loading (0.751) on that factor. The four-factor model had problems with two items: Item #3 and Item #11. Item #3 had factor loadings of 0.369, 0.449, 0.216, and 0.261 indicating possibility of placing this item into either factor 1 or factor 2. Likewise, Item #11 had comparable loadings on factors 1 and 4 (0.300 and 0.494, respectively). Once again, the three-factor model appeared as a better choice than the four-factor model.

Based on the above-mentioned analyses, it made a perfect sense to choose the three-factor model as the final model and use it in subsequent analysis. Indeed, this model had an acceptable fit to data and aligned well with the research constructs. Figure 3 depicts the main steps of EFA of SIEVEA.

CONFIRMATORY FACTOR ANALYSIS

Confirmatory factor analysis (CFA) is conducted when the models are already established from the theory and then need to be tested (Eccles & Wigfield, 1995). CFA allows the researchers to test their initial theory about the constructs present in data (Chamberlin et al., 2017). As EFA results showed, the three-factor model had acceptable fit to data and was chosen as the final model. Consequently, CFA was conducted on this model using a new sample data for conformity of its suggested factor structure to the data.

Data

The sample data were made up of survey responses of 1495 high school students out of which 814 (54.44%) were females and 677 (45.28 %) males. Four students (0.28 %) did not report their gender. Data were collected from three schools (urban and suburban) in New Jersey and Connecticut. These data were specifically gathered for conducting a CFA.

Before performing CFA, then list wise deletion, which is the most common method in dealing with missing data (Schafer

& Graham, 2002), was used. As a result, 56 incomplete responses out of the total 1495 responses were dropped. CFA was performed on the remaining data, totaling 1439 responses.

Procedure

CFA was executed on the three-factor model using the Analysis of Moment Structures (AMOS) software (Version 24). The maximum likelihood (ML) method, the most common model estimation procedure for conducting CFA, was utilized. The study closely followed the guidelines and procedures recommended by Matsunaga (2010) and Jackson et al. (2009). This helped in avoiding common mistakes that happen during factor analysis. While testing the model's fit to data and calculating relevant close-fit indexes, special attention was paid to evaluated factor loadings, correlations, and their statistical significance. The model was repeatedly adjusted and reevaluated to improve the fit.

The ML estimation procedure assumes that data are normally distributed (Jackson et al., 2009). The descriptive statistics of data showed that data distribution was close to normal. Hence, data were adequate for conducting CFA.

For evaluating candidate models, this study used the most frequently used goodness-of-fit indexes. Based on Matsunaga's (2010) recommendations, the following indexes were computed: RMSEA, SRMR, NFI, CFI, and TLI.

To do CFA in AMOS, the model was set up as a path diagram with circles representing the latent concepts and squares representing observed variables. Since this model has three factors and 13 items, three circles were drawn for three latent variables along with eight, three, and two squares for observed variables, each set of squares connected to its respective latent variable as suggested by EFA. One-directional arrows used to connect latent variables to observed variables (survey items), indicated assumed causal influence of latent variables on corresponding items. Double-directional arrows were used to represent covariance among three latent variables.

In addition, since factor analysis assumes that the observed variables were measured with errors (Matsunaga, 2010), the diagram needed to capture the measurement errors. This was

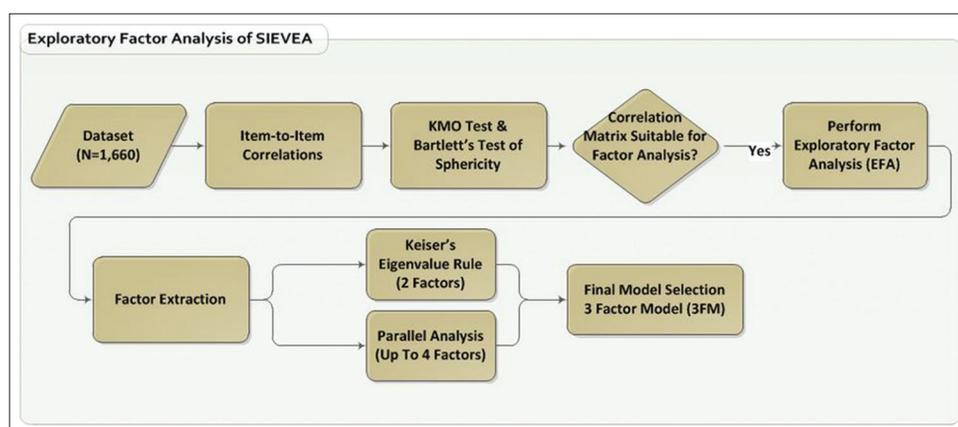


Figure 3: Exploratory factor analysis of SIEVEA

accomplished by adding unobserved variables, one error variable per item (i.e. e3 for Item #3). The resulting measurement model for the three-factor model is shown in Figure 4.

All standardized factor loadings were significant with $p < 0.001$ and ranged in magnitude from 0.608 to 0.873. As data in Table 7 shows, these factor loadings closely resemble those of EFA even though the data set used for CFA was different from the one used for EFA. Inter-factor correlations were 0.48, 0.51, and 0.63, indicating some correlation between factors. This was not surprising because, as Matsunaga (2010) emphasized, most phenomena that are studied in social sciences are interrelated, and CFA models usually specify inter-relations among latent variables/factors.

In addition, the inspection of modification indices for covariances demonstrated that the treating of covariances between some error variables as free parameters, provided the variables were within the same factor, could further improve the model's fit.

The final measurement model for the three-factor model resulted in the following close-fit indexes values: RMSEA = 0.071, SRMR = 0.044, NFI = 0.95, CFI = 0.95, TLI = 0.94. Figure 5 depicts the final measurement model for the three-factor model after the above-mentioned changes were applied.

RESULTS

CFA generated values of the close-fit indexes of the three-factor model showed an acceptable fit, with some indexes approaching the thresholds of a good fit. Both the RMSEA (0.071) and the SRMR (0.044) were below the 0.08 threshold of the acceptable fit. The incremental close-fit indexes were either equal to or slightly less than the borderline value of 0.95 (NFI = 0.95, CFI = 0.95, TLI = 0.94).

Instrument's Reliability and Validity

Tavakol and Dennick (2011) asserted that two important elements, validity, and reliability, are crucial for evaluating

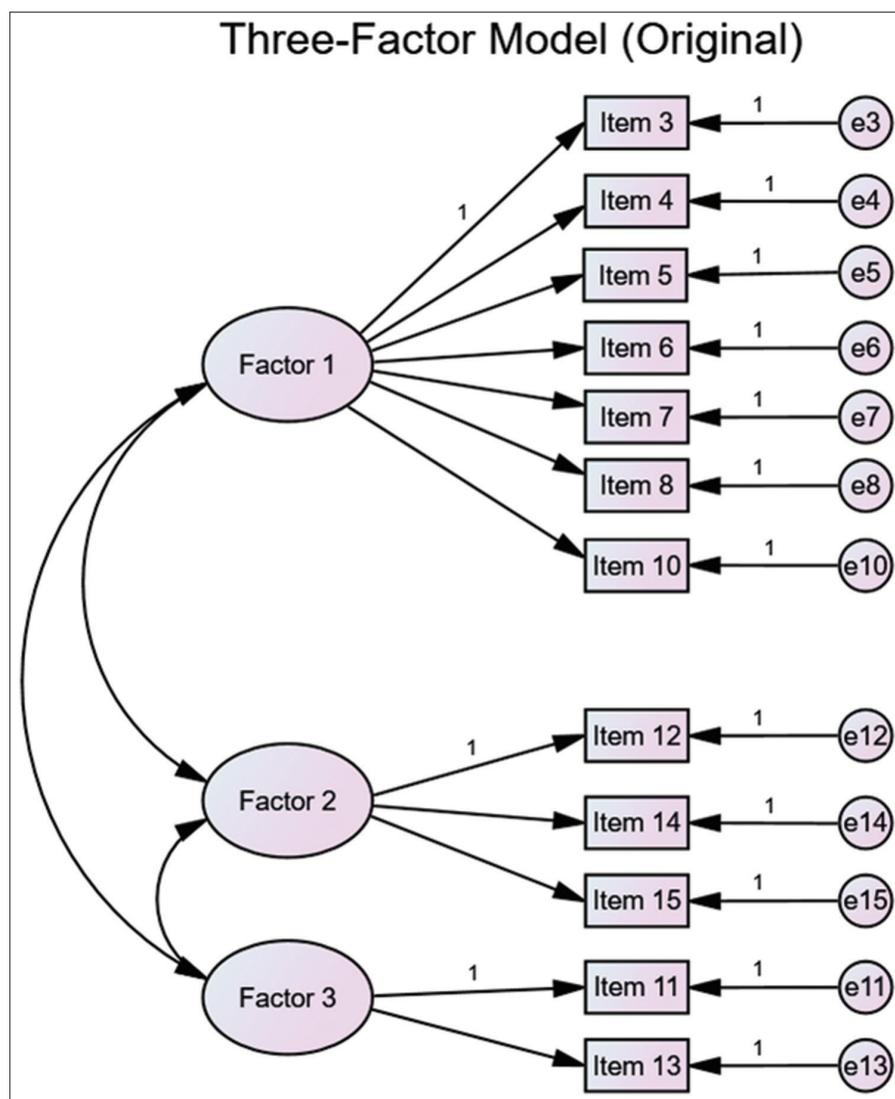


Figure 4: The representation of the three-factor model as a path diagram. Oval shapes represent unobservable latent factors, rectangles are representations of observed items, and arrows indicate loadings

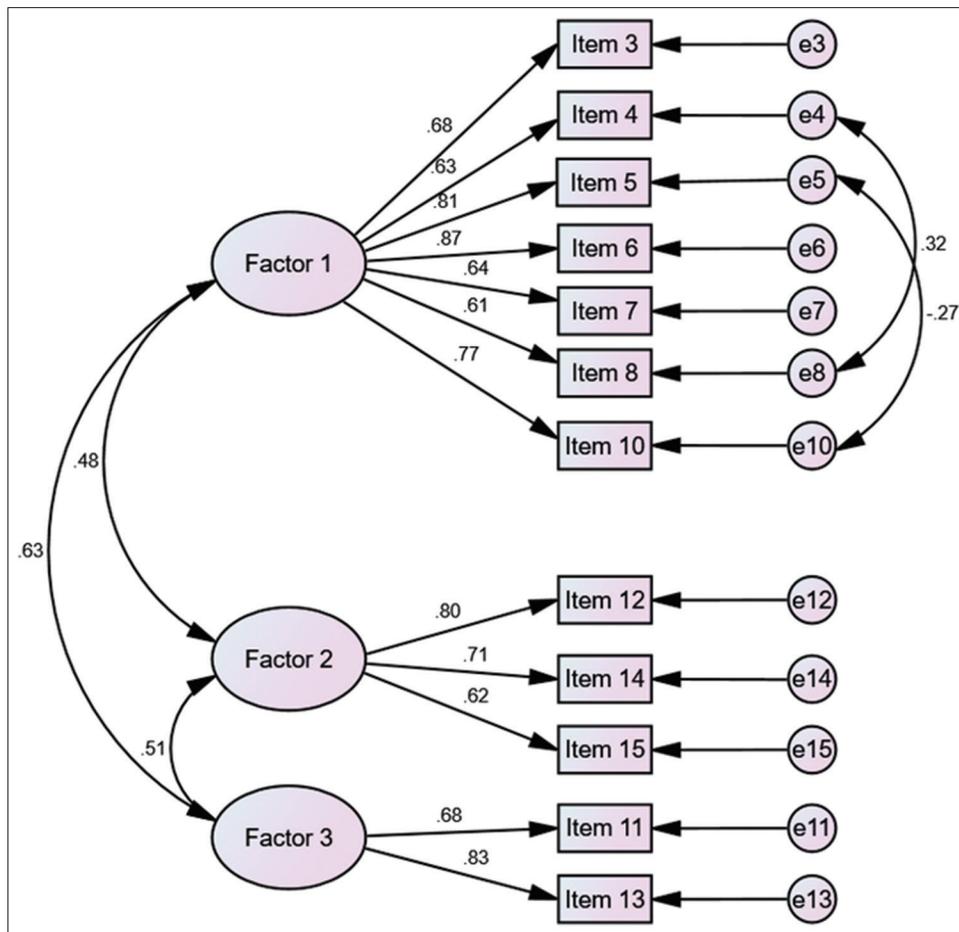


Figure 5: The representation of final three-factor model factor loadings and correlations between factors are displayed. It also shows additional free parameters between error variables: e4 and e8; e5, and e10

Table 7: Unrestricted (EFA) and restricted (CFA) three-factor model: standardized factor loadings

| Item | Unrestricted (EFA) | | | Restricted (CFA) | | |
|------|--------------------|-------|-------|------------------|-------|-------|
| | C1 | C2 | C3 | C1 | C2 | C3 |
| 3 | 0.605 | 0.268 | 0.152 | 0.677 | * | * |
| 4 | 0.630 | 0.054 | 0.146 | 0.634 | * | * |
| 5 | 0.686 | 0.352 | 0.023 | 0.807 | * | * |
| 6 | 0.794 | 0.304 | 0.087 | 0.873 | * | * |
| 7 | 0.582 | 0.028 | 0.277 | 0.636 | * | * |
| 8 | 0.607 | 0.063 | 0.201 | 0.608 | * | * |
| 10 | 0.650 | 0.260 | 0.164 | 0.775 | * | * |
| 11 | 0.246 | 0.197 | 0.751 | * | * | 0.680 |
| 12 | 0.120 | 0.753 | 0.148 | * | 0.799 | * |
| 13 | 0.409 | 0.373 | 0.437 | * | * | 0.828 |
| 14 | 0.104 | 0.677 | 0.140 | * | 0.713 | * |
| 15 | 0.247 | 0.613 | 0.054 | * | 0.619 | * |

EFA was performed on the original data set (1660 responses), whereas CFA was performed on the new data set (1439 responses)

survey instruments. Furthermore, researchers stated the importance of making sure the instrument passes content and construct validities (Moskal et al., 2002). Therefore, ensuring the high degree of validity was the focus of the SIEVEA instrument.

The validity indicates the extent to which the survey measures whatever construct it intends to measure (Groves et al., 2009). Since the SIEVEA instrument was intended for measuring multiple constructs, it was important to verify the construct validity by measuring whether convergent and discriminant validities were satisfied (Agarwal, 2011). According to Trochim (2006), to ensure its construct validity is satisfied, both convergent and discriminant validities should be confirmed. For that purpose, correlations between constructs' items were measured.

Jöreskog (1967) highlighted the importance of assessing the instrument's construct validity. Measuring and confirming the instrument's validity are a crucial step in analyzing the instrument's psychometric properties and must be carried out before the instrument's data can be used for other statistical tests. Two major types of validity, convergent, and discriminant, need to be estimated for supporting the evidence of construct validity (Campbell & Fiske, 1959). According to these scholars, convergent validity indicates whether the constructs are well measured by their respective items or not. Likewise, discriminant validity assesses the degree to which measures of different constructs are unrelated.

Table 8: Validation measures and inter-construct correlations (the three-factor model)

| Construct | CR | AVE | MSV | ASV | C1 | C2 | C3 |
|-----------|-------|-------|-------|-------|--------------|--------------|--------------|
| C1 | 0.882 | 0.521 | 0.393 | 0.311 | 0.722 | – | – |
| C2 | 0.755 | 0.510 | 0.260 | 0.245 | 0.479 | 0.714 | – |
| C3 | 0.727 | 0.574 | 0.393 | 0.327 | 0.627 | 0.510 | 0.758 |

The bolded numbers on the diagonal on the right part of the table are the square roots of the AVE. The non-diagonal numbers are inter-construct correlations. AVE: Average variance extracted, CR: Composite reliability, MSV: Maximum shared variance, ASV: Average shared variance

Fornell and Larcker (1981) proposed several convenient and widely used measures for establishing validity and reliability. The composite reliability (CR) and the average variance extracted (AVE) are used for assessing reliability and convergent validity. CR is similar to Cronbach's alpha, but is less biased, as Cronbach's alpha tends to underestimate true reliability (Peterson & Kim, 2013). CR has an acceptable value of 0.7 and above. The AVE is similar to explained variance in EFA as it measures the average variance in items that a construct manages to explain. In other words, the AVE expresses the level of variance captured by a construct versus the level due to measurement errors. AVE values 0.5 and above are acceptable for establishing convergent validity.

The discriminant validity can be evaluated using the maximum shared variance (MSV) and the average shared variance (ASV), which, respectively, measure maximum and average variances among constructs. Both measures should be lower than the AVE for all constructs for confirming discriminant validity (Hair et al., 2010). Furthermore, the square root of the AVE should be greater than inter-construct correlations for all constructs. The CR, AVE, MSV, and ASV were calculated for the three-factor model to assess its convergent and discriminant validities. AMOS software (Arbuckle, 2016) was used to produce input data (factor correlations and standardized regression weights) for calculating these metrics. Table 8 contains the results of these calculations.

According to these results, the three-factor model satisfied all reliability and validity requirements. Indeed, since the CR values of all three constructs were greater than the threshold value of 0.7, the reliability requirement was met. Likewise, convergent validity was confirmed as well, because all AVE values exceeded 0.5. Finally, all discriminant validity requirements were satisfied: Both the MSV and the ASV were less than the AVE for all three constructs, and the square roots of the AVE (the bolded numbers on the diagonal on the right part of the table) were greater than the corresponding inter-construct correlations.

DISCUSSION

The three-factor model is appropriate for measuring the following latent constructs: Students' science identities and motivation (C1), environmental attitudes (C2), and science values (C3). The model has a good fit and a simple factor structure. The three-factor model was also tested for reliability. In addition, this model was assessed for construct validity. These

tests indicated that the three-factor model was reliable and passed all the requirements of convergent and discriminant validities.

As a result of these analyses, the three-factor model was fully validated and was confirmed as an appropriate representation of the SIEVEA's factor structure. Indeed, the three-factor model had an acceptable fit to data, high levels of reliability, and construct validity.

The results of this study illustrated the usefulness of SIEVEA as a simple and expedient instrument for measuring important constructs related to science learning and environmental attitudes. Indeed, the online survey format made this instrument readily available for multiple schools, resulting in extensive student participation. The instrument has several valuable and convenient features. First, students can take the survey either from home or in the classroom. Second, according to the Lexile framework, the survey's questions are age-appropriate for high school students' reading level. Third, it takes <15 min to complete the survey. Fourth, the survey is short and has only 15 items. Finally, the survey is online, which makes it very convenient for response data extraction and analysis

It is recommended to improve the SIEVEA survey instrument by adding more items related to the science value construct. This will overcome the survey's current limitation by enhancing its reliability in capturing the science value latent variable. Since the survey was not a lengthy one, adding several (e.g. 2–3) items will continue to maintain the brevity of the instrument. Another recommendation is to expand the geography of data collection. Even though the survey was conducted in multiple states (New Jersey, Pennsylvania, and Connecticut), its online nature made it possible to collect data in various states, as well as outside of the U.S. This allows for more diverse data collection, leading to the improved external validity of future studies.

CONCLUSIONS

When researchers are interested in understanding what motivates students to learn science, they usually examine their reasons for science learning and their beliefs and feelings that influence this learning (Glynn & Koballa, 2006). It is believed that understanding the factors that contribute to students' motivation in science learning will help science education researchers and educators to improve science education. For example, various studies of theoretical models of motivation pointed out that when students hold strong beliefs and expectations about themselves, they are more likely to engage in tasks and persist in doing them despite any difficulties they may encounter while working on tasks (Weiner, 1992). Likewise, understanding students' attitudes toward environment can help environmental educators to take action for increasing students' environmental awareness and their environmental literacy.

The survey developed in this study can help researchers and educators in accomplishing the above-mentioned tasks. The instrument can also serve as a screening/diagnostic tool for science teachers and school administration who want to identify

students who lack motivation or have indifferent attitudes toward the environment. Finally, it can be used as an instrument for evaluating newly introduced science curriculum, teaching strategy or instructional methodology. The instrument can also be used to survey students' science and environmental attitudes in general. In addition, the researchers can use it prior and/or post their research related to science learning and environmental attitudes to measure the effectiveness of their researched methods.

REFERENCES

- Agarwal, N.K. (2011). Verifying survey items for construct validity: A two stage sorting procedure for questionnaire design in information behavior research. *Proceedings of the American Society for Information Science and Technology*, 48(1), 1-8.
- American Association for the Advancement of Science. (1989). *Science for all Americans*. New York: Oxford University Press.
- Arbuckle, J.L. (2016). *Amos User's Guide Version 24*. Chicago, IL: Small Waters.
- Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107, 238-246.
- Bentler, P.M., & Bonett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606.
- Blatt, E. (2014). Uncovering students' environmental identity: An exploration of activities in an environmental science course. *The Journal of Environmental Education*, 45(3), 194-216.
- Bradley, J.C., Waliczek, T.M., & Zajicek, J.M. (1999). Relationship between environmental knowledge and environmental attitude of high school students. *The Journal of Environmental Education*, 30(3), 17-21.
- Browne, M.W., & Cudek, R. (1993). Alternative ways of assessing model fit. In: Bollen, K.A., & Long, J.S., (Eds.), *Testing Structural Equation Models*. Newbury Park, CA: SAGE. p136-162.
- Bybee, W.R., & Fuchs, B. (2006). Preparing 21st century workforce: A new reform in science and technology education. *Journal of Research in Science Teaching*, 43(4), 349-352.
- Campbell, D.T. & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2), 81.
- Carlone, H.B., & Johnson, A. (2007). Understanding the science experiences of successful women of color: Science identity as an analytic lens. *Journal of Research in Science Teaching*, 44(8), 1187-1218.
- Chamberlin, S.A., Moore, A.D., & Parks, K. (2017). Using confirmatory factor analysis to validate the Chamberlin affective instrument for mathematical problem solving with academically advanced students. *British Journal of Educational Psychology*, 87(3), 422-437.
- Coyle, K. (2005). *Environmental Literacy in America: What Ten Years of NEETF/Roper Research and Related Studies say about Environmental Literacy in the U.S.* Washington, DC: The National Environmental Education and Training Foundation. Available from: <https://www.neefusa.org/file/234/download?token=XFwEZSxW>. [Last accessed on 2019 Sept 18].
- DiEnno, C.M., & Hilton, S.C. (2005). High school students' knowledge, attitudes, and levels of enjoyment of an environmental education unit on nonnative plants. *The Journal of Environmental Education*, 37(1), 13-25.
- Dunlap, R.E., Van Liere, K.D., Mertig, A.G., & Jones, R.E. (2000). New trends in measuring environmental attitudes: Measuring endorsement of the new ecological paradigm: A revised NEP scale. *Journal of Social Issues*, 56(3), 425-442.
- Eccles, J.S. & Wigfield, A. (1995). In the mind of the actor: The structure of adolescents' achievement task values and expectancy-related beliefs. *Personality and Social Psychology Bulletin*, 21(3), 215-225.
- Eccles, J.S. (2005). Subjective task values and the Eccles et al. model of achievement-related choices. In: Elliot, A.J., & Dweck, C.S. (Eds.), *Handbook of Competence and Motivation*. New York: Guilford. p105-121.
- Eccles, J.S. (2009). Who am i and what am i going to do with my life? Personal and collective identities as motivators of action. *Educational Psychologist*, 44, 78-89.
- Eccles, J.S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53(1), 109-132.
- Eccles, J.S., Wigfield, A., & Schiefele, U. (1998). Motivation to succeed. In: Damon, W., & Eisenberg, N. (Eds.), *Handbook of Child Psychology: Vol. 3. Social, Emotional and Personality Development*. New York: Wiley. p1017-94.
- Eilam, E., & Trop, T. (2012). Environmental attitudes and environmental behavior which is the horse and which is the cart? *Sustainability*, 4(9), 2210-2246.
- Erdogan, M., Bahar, M., & Usak, M. (2012). Environmental education in high school 9th-12th biology course curricula started to be implemented in 2007. *Educational Sciences: Theory and Practice*, 12(3), 2230-2235.
- Fabrigar, L.R., Wegener, D.T., MacCallum, R.C., & Strahan, E.J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272-299.
- Fornell, C., & Larcker, D.F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39-50.
- Fortus, D., & Vedder-Weiss, D. (2014). Measuring students' continuing motivation for science learning. *Journal of Research in Science Teaching*, 51(4), 497-522.
- Fredricks, J.A., & Eccles, J.S. (2002). Children's competence and value beliefs from childhood through adolescence: Growth trajectories in two male-sex-typed domains. *Developmental Psychology*, 38(4), 519-533.
- Gee, J.P. (2000). Identity as an analytic lens for research in education. *Review of Research in Education*, 25, 99-125.
- Gignac, G.E. (2009). Partial confirmatory factor analysis: Described and illustrated on the NEO-PI-R. *Journal of Personality Assessment*, 91(1), 40-47.
- Glorfeld, L.W. (1995). An improvement on Horn's parallel analysis methodology for selecting the correct number of factors to retain. *Educational and Psychological Measurement*, 55, 377-393.
- Glynn, S.M., & Koballa, T.R. (2006). Motivation to learn in college science. In: Mintzes, J.J., & Leonard, W.H. (Eds.), *Handbook of College Science Teaching*. Arlington, VA: National Science Teachers Association Press. p25-32.
- Glynn, S.M., Brickman, P., Armstrong, N., & Taasobshirazi, G. (2011). Science motivation questionnaire II: Validation with science majors and nonscience majors. *Journal of Research in Science Teaching*, 48(10), 1159-1176.
- Graham, S., & Weiner, B. (1996). Theories and principles of motivation. In: *Handbook of Educational Psychology*. Vol. 4. New York: Macmillan. p63-84.
- Gresalfi, M.S. (2009). Taking up opportunities to learn: Constructing dispositions in mathematics classrooms. *The Journal of the Learning Sciences*, 18(3), 327-369.
- Groves, R.M., Fowler, F.J. Jr., Couper, M.P., Lepkowski, J.M., Singer, E., & Tourangeau, R. (2009). *Survey Methodology*. Vol. 561. Hoboken, NJ: John Wiley and Sons.
- Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., & Tatham, R.L. (2010). *Multivariate Data Analysis*. 7th ed. New York: Pearson.
- Horn, J.L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30, 179-185.
- Hu, L.T., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55.
- Jackson, D.L., Gillaspay, J.A. Jr., & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: An overview and some recommendations. *Psychological methods*, 14(1), 6-23.
- Johnson, B., & Manoli, C.C. (2011). The 2-MEV scale in the United States: A measure of children's environmental attitudes based on the theory of ecological attitude. *The Journal of Environmental Education*, 42(2), 84-97.
- Jöreskog, K.G. (1967). A general approach to confirmatory maximum likelihood factor analysis. *ETS Research Bulletin Series*, 34(2), 183-202.
- Jowett, T., Harraway, J., Lovelock, B., Skeaff, S., Slooten, L., Strack, M., & Shephard, K. (2014). Multinomial-regression modeling of the environmental attitudes of higher education students based on the revised new ecological paradigm scale. *The Journal of Environmental Education*, 45(1), 1-15.

- Kaiser, H.F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20, 141-151.
- Kane, J.M. (2016). Young African American boys narrating identities in science. *Journal of Research in Science Teaching*, 53(1), 95-118.
- Lavrakas, P.J. (2008). *Encyclopedia of Survey Research Methods*. Thousand Oaks: SAGE Publications Ltd.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 140, 5-55.
- Matsui, T., Matsui, K., & Ohnishi, R. (1990). Mechanisms underlying math self-efficacy learning of college students. *Journal of Vocational Behavior*, 37(2), 225-238.
- Matsunaga, M. (2010). How to factor-analyze your data right: Do's, don'ts, and how-to's. *International Journal of Psychological Research*, 3(1), 97-110.
- MetaMetrics Inc. (2012). *The Lexile Framework for Reading*. Available from: <http://www.lexile.com>. [Last accessed on 219 Sep 18].
- Moje, E.B., Tucker-Raymond, E., Varelas, M., & Pappas, C.C. (2007). FORUM: Giving oneself over to science exploring the roles of subjectivities and identities in learning science. *Cultural Studies of Science Education*, 1(3), 593-601.
- Moskal, B.M., Leydens, J.A., & Pavelich, M.J. (2002). Validity, reliability and the assessment of engineering education. *Journal of Engineering Education*, 91(3), 351-354.
- Mubeen, S., & Reid, N. (2014). The measurement of motivation with science students. *European Journal of Educational Research*, 3(3), 129-144.
- National Center for Education Statistics. (2011). *National Assessment of Educational Progress (NAEP)*. Available from: <http://www.nces.ed.gov/nationsreportcard>. [Last accessed on 2019 Sep 18].
- O'Connor, B.P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instrumentation, and Computers*, 32, 396-402.
- OECD. (2012). *PISA 2012 Results in Focus*. Available from: <http://www.oecd.org/pisa/keyfindings/pisa-2012-results-overview.pdf>. [Last accessed on 219 Sep 18].
- Osbaldiston, R. (2004). *Meta-analysis of the Responsible Environmental behavior Literature*. (Unpublished Doctoral Dissertation). Columbia: University of Missouri-Columbia.
- Peterson, R.A., & Kim, Y. (2013). On the relationship between coefficient alpha and composite reliability. *Journal of Applied Psychology*, 98(1), 194.
- Pintrich, P.R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology*, 95(4), 667-686.
- Pintrich, P.R., & De Groot, E.V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, 82(1), 33-40.
- Plante, I., O'Keefe, P.A., & Théorêt, M. (2012). The relation between achievement goal and expectancy-value theories in predicting achievement-related outcomes: A test of four theoretical conceptions. *Motivation and Emotion*, 37(1), 65-78.
- Schafer, J.L., & Graham, J.W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7(2), 147-177.
- Sfard, A., & Prusak, A. (2005). Telling identities: In search of an analytic tool for investigating learning as a culturally shaped activity. *Educational Researcher*, 34(4), 14-22.
- Silva, E. (2009). Measuring skills for 21st century learning. *The Phi Delta Kappan*, 90(9), 630-634.
- Sobel, D. (1996). *Beyond Ecophobia: Reclaiming the Heart in Nature Education*. Great Barrington, MA: The Orion Society and Myrin Institute.
- Stern, J.M., Powell, B.R., & Ardoin, M.N. (2011). Evaluating a constructivist and culturally responsive approach to environmental education for diverse audiences. *The Journal of Environmental Education*, 42(2), 109-122.
- Swarat, S., Ortony, A., & Revelle, W. (2012). Activity matters: Understanding student interest in school science. *Journal of Research in Science Teaching*, 49(4), 515-537.
- Tabachnick, B.G., & Fidell, L.S. (2001). *Using Multivariate Statistics*. Boston, MA: Allyn and Bacon.
- Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, 2, 53-55.
- Thompson, B. (2004). *Exploratory and Confirmatory Factor Analysis: Understanding Concepts and Applications*. Washington, DC: American Psychological Association.
- Trochim, W.M. (2006). *The Research Methods Knowledge Base*. 2nd ed. Available from: <http://www.socialresearchmethods.net/kb>. [Last accessed on 2019 Sep 18].
- Tucker, L.R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, 38, 1-10.
- Tucker-Raymond, E., Varelas, M., Pappas, C., Korzh, A., & Wentland, A. (2007). "They probably aren't named Rachel": Young children's scientist identities as emergent multimodal narratives. *Cultural Studies of Science Education*, 1(3), 559-592.
- U.S. Department of Education. (n.d.). *Science, Technology, Engineering, and Math*. Available from: <http://www.ed.gov/stem>. [Last accessed on 2019 Sep 18].
- Weiner, B. (1992). *Human Motivation: Metaphors, Theories, and Research*. Newbury Park, CA: SAGE Publications.
- Wigfield, A., & Eccles, J.S. (1992). The development of achievement task values: A theoretical analysis. *Developmental Review*, 12(3), 265-310.
- Wigfield, A., Tonks, S., & Klauda, S.L. (2009). Expectancy-value theory. In: Wentzel, K.R., & Wigfield, A. (Eds.), *Handbook of Motivation at School*. New York: Routledge. p55-75.
- Zecha, S. (2010). Environmental knowledge, attitudes and actions of Bavarian (Southern Germany) and Asturian (Northern Spain) adolescents. *International Research in Geographical and Environmental Education*, 19(3), 227-240.
- Zwick, W.R., & Velicer, W.F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, 99, 432-442.