

## Development of an Energy Literacy Measure for Middle School Students

**R. Justin Hougham**, University of Wisconsin - Extension

**Chad Gotch**, Washington State University

**Jennifer A. Schon**, University of Idaho

**Karla Eitel**, University of Idaho

**Danica Hendrickson**, Western Washington University, danicahendrickson@gmail.com

**Abstract:** Energy literacy, defined by the DOE, “is an understanding of nature and role of energy in the universe and in our lives” and, “. . . is also the ability to apply this understanding to answer questions and solve problems” (U.S. Department of Energy, 2013). Energy literacy is continuously evolving with the development of new feedstocks, technologies, and processes - all of which contribute to the changing landscape of energy production and use. In order for energy education to evolve with the energy field, better tools are needed to assess educational programs. The assessment discussed here is a step towards developing such an assessment for bioenergy.

**Keywords:** assessment, energy literacy, environmental education, youth development

---

***Dr. Hougham** is faculty at the University of Wisconsin- Extension where he supports the delivery of a wide range of science education topics to K-12 students, graduate students, and in-service teachers. Justin’s scholarship is in the areas of place-based pedagogies, STEM education, and education for sustainability.*

***Chad Gotch** is an Assistant Professor in Educational Psychology and the Learning and Performance Research Center at Washington State University. Dr. Gotch works to maximize effective and appropriate use of educational and psychological measurements. To this end, he studies assessment literacy among teachers, score reporting, and building validity arguments from both technical and non-technical evidence.*

***Jennifer A. Schon** is a Bluum fellow working on starting a charter school in rural Idaho. She has a Ph.D. in Education from the University of Idaho, a M.A. in Education from Ashford University, and a B.S. in Environmental Science from Western Washington University.*

***Karla Bradley Eitel** is a Research Associate Professor in the Natural Resources and Society Department of the University of Idaho and the Director of Education for the University of Idaho’s McCall Outdoor Science School. She holds a PhD in Natural Resources, an M.S. in Conservation Social Sciences and an M.Ed. in Curriculum and Instruction all from the University of Idaho. She earned her BA in Studio Art and American Studies from Williams College.*

***Danica Hendrickson** is the Program Director of Instruction for the Aerospace Joint Apprenticeship Committee. Danica earned a M.Ed. from the Harvard Graduate School of Education’s Mind, Brain, and Education program, co-authored Facing the Future’s Fueling Our Future curriculum, and has 12 years of middle school teaching experience.*

Energy provides the fuel for our lives. It heats our homes, powers our electronics, and transports people and products around the world. The energy sector is also responsible for more than two-thirds of the emissions causing climate change, and energy is inextricably linked with complex global issues such as human health, national security, and international trade (Birol, 2015). Energy issues are challenging, frequently changing, and the individual and collective choices that we make about energy—such as which energy sources to harness or extract—impact the social, economic, and environmental health of a community and our world. Responding to these challenging energy issues and making beneficial individual and collective decisions is critically important and depends upon an energy-literate populace that understands the basics of energy production and transfer and can weigh costs and benefits associated with varied options for achieving a sustainable energy future. As the world transitions from being primarily dependent upon fossil fuels to utilizing alternative energy sources, a need to equip both children and adults with current and foundational levels of energy literacy becomes apparent.

Efforts to address this need must first define energy literacy in an actionable way; secondly, tools to examine how energy literacy is shaped by educational efforts must be developed and implemented. While the world transitions to developing alternative energy sources, significant enhancements have been made to implement more energy-based education in – and out- of the classroom. The success of these two trajectories depends upon an energy-literate populace that is able to make informed choices for a sustainable energy future and for careers in ever changing domains of the energy industry. Several projects, grants, and websites have focused on creating innovative, integrative and meaningful energy related lessons over recent years (e.g., Advanced Hardwood Biofuels, Bioenergy Alliance of the Rockies). Our own team, through the Northwest Advanced Renewables Alliance (NARA) has been involved in the development of educational resources and programming that promotes energy literacy with respect to biofuels. Less effort, however, has been focused on developing assessments to inform instruction and evaluate learning in these settings. What does energy literacy look like and how do we assess it? This question led our research team to review existing research on energy literacy and to develop a pair of new energy literacy assessment tools for that reflect current research in bioenergy/biofuels—one for the middle school population and one for high school students and adults. In this paper, we focus on the Energy Literacy Inventory—Middle School (ELI-M).

## **Method**

### **Assessing Student Energy Literacy: The Need for More Tools/Resources**

As energy literacy has become increasingly important for our communities and our world, a need to accurately determine and compare change in energy literacy has also increased. Vetted energy literacy assessment tools are needed to provide various levels of data and feedback for educators, curriculum developers, programs, and policy-makers. These tools can provide snapshots of an individual or group's current understanding of energy, can show change in energy literacy when administered before and after educational programming, and can provide insight about an individual or group's understanding of specific energy concepts. In particular, a concise tool informed by contemporary energy literacy research and tailored to different age groups can be a convenient way to advance energy education efforts.

## Tomorrow's Energy Literate Citizenry

As sustainability education and global climate science have become increasingly ubiquitous in contemporary science education, energy education has increased in importance and focus (Hendrickson, Corrigan, Keefe, Shaw, Jacob, Skelton, Schon, Eitel, & Hougham, 2015). While federally funded grants in the energy sector are numerous and at regional as well as national levels, if they are to fully realize their potential towards supporting a pipeline from grant to education to industry application or employment, much needs to be developed in the way of assessing and evaluating that pipeline (Schon, Eitel, Hougham, & Hendrickson, 2015). Initial approaches to energy education included energy types and pathways, later including energy conservation and efficiency as well as usage (DeWaters, Jan, Hougham, Hintz, & Frolich, 2015). Today, both the energy landscape and curricular environment have become more diverse and sophisticated. To drive and inform energy literacy education, a vetted assessment tool is needed to accurately and quickly provide valuable information to educators. Such a tool can inform teachers of the level and need of their students' energy literacy, allowing them to scaffold challenging topics, monitor student progress, and summarize student understanding. Yet, current assessment tools are largely created to be specific to a lesson or short unit, such as a short content quiz, and do not function as an overall gauge of general energy understanding.

One example of a more comprehensive energy literacy assessment tool was created by DeWaters, Qaqish, Graham, and Powers (2013). This test was created for middle school students with a separate version for high school students. The middle school version includes 61 questions that includes multiple choice and scale questions (strongly agree to strongly disagree or expert to novice on a topic). The high school version is made up of 69 questions. The test was designed to follow a four-week unit on energy to examine the impact the project-based work.

The DeWaters test includes questions pertaining to cognitive, affective, self-efficacy and behavior domains and energy literacy. While assessment of the affective, self-efficacy, and behavior domains of energy literacy is important, practitioners like teachers, informal educators, and curriculum designers are unable to realistically ask their students to take a 60+ question test.

Additionally, the DeWaters test was created prior to the U.S. Department of Energy (DOE) releasing of the *Energy Literacy: Essential Principles and Fundamental Concepts for Energy Education* (DeWaters, Qaqish, Graham, & Powers, 2013). The principles provide a guideline of energy topics and progression for educators to focus in on for the vast subject of energy education.

We therefore propose a shorter, more current test, using the Next Generation Science Standards and DOE principles that focuses on the cognitive domain of energy literacy. A short assessment with this focus is needed to determine students' entry energy literacy level and change in energy literacy, as it relates to content understanding, following energy-related lessons. The proposed tool therefore may serve as a stand-alone instrument for practitioners most interested in understanding of energy concepts, or it may be paired with the longer DeWaters instrument for a full assessment of energy literacy.

## Test Validation

In contemporary notions of educational testing, *validity* refers to “the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests” (American Educational Research Association, American Psychological Association, National Council on Measurement in Education, 2014, p. 11). The process of test validation involves an

explicit statement of the proposed interpretations that are to be drawn from a student's outcome on a test. These interpretations must be relevant to the intended use of the test. Therefore, validity concerns not just whether the test measures what it is supposed to measure (i.e., an intended construct), but that the way in which someone interprets or acts upon test results is logical and defensible. For example, suppose a researcher designed a test to measure an individual's knowledge of the ecological functions of wetlands. The design of this test may align strongly with science in the field and what ecologists cite as the primary benefits of wetlands. It may have been pre-tested with the target population and revised to provide accurate and consistent information about test-takers' level of understanding. Based on this evidence, we may conclude the test measures what it was intended to measure. It would be inappropriate, however, to interpret that individual's outcome on the test is indicative of one's willingness to advocate on behalf of wetland restoration projects. This would be an invalid interpretation. Likewise, one could not defensibly use scores from this test on their own to assign grades in a general biology course. In such a case, the content of the test would not likely align to the breadth of learning objectives for the course, and use of the test for this purpose would lack validity. In the same way, a content-focused energy literacy assessment would not indicate one's willingness to advocate for energy conservation or reduce their carbon footprint. Appropriate interpretation and use of scores from the assessment would stay aligned to conclusions about what the student understands about the costs and benefits of various options for energy production and transfer.

The practice of test validation is operationalized through an argument framework (Kane, 2013). Based on Toulmin's model of inference (Toulmin, Rieke, & Janik, 1979), a series of claims are specified and linked together. Claims may focus on the integrity of how student responses are translated into a numeric judgment, how fair a test is in terms of providing equitable opportunity across different categories of students, the extent to which an underlying construct has been well represented, or how predictive a student's test score will be of future behaviors of interest, to name a few examples. In general, stated claims are oriented toward the interpretations one can draw from the outcome of a test and appropriate uses of that outcome.

The interpretive claims evaluated in this paper, which constitute a portion of the body of validation work that can and should be performed on the test, were as follows: 1) Test items target students' level of energy literacy in a way that is conceptually consistent with contemporary thinking in the fields of environmental and energy education. 2) Responses to test items allow for the differentiation of high, medium, and low levels of energy literacy among students. Evidence for the first claim is provided in this paper in the form of a documentation of test development procedures. Evidence for the second claim is provided via psychometric analyses of individual test items. A positive evaluation of these interpretive claims would provide partial support for a use claim that the energy literacy instrument can be used in an educational research context, to identify programmatic needs, and to assess impact of programming on individuals and groups (i.e., as an outcome measure).

### **Development of the Energy Literacy Inventory**

In this section, we document the efforts that went into the development of the Energy Literacy Inventory--Middle School (ELI-M) form. In doing so, we attempt to provide partial support for the intended uses of the tool based on relevant and representative test content. As seen in the review of the literature, a concise assessment tool for middle school students was previously lacking. Several members of the NARA Education team gathered to address this gap in understanding and to address the need of a more specific assessment tool. The NARA

Education members created a development panel to create a bio-energy focused assessment tool grounded in existing research with an accepted framework that could be scaled appropriately for various age groups (elementary, middle, and high school). The development panel consisted of several formal and informal (e.g., outdoor education) teachers, energy industry experts, and an educational assessment expert.

To ground the ELI-M in contemporary thinking in the area of energy literacy, the development panel selected the U.S. Department of Energy’s (DOE) framework presented in *Energy Literacy: Essential Principles and Fundamental Concepts for Energy Education* (U.S. Department of Energy, 2013) as the basis for its test blueprint (Fives & DiDonato-Barnes, 2013). In this document, the DOE defines energy literacy as “an understanding of nature and role of energy in the universe and in our lives” and “energy literacy is also the ability to apply this understanding to answer questions and solve problems”. The document breaks down energy concepts into seven principles that hope to empower individuals and communities to make informed energy decisions. Being the national resource for energy information, the DOE principles provide a unifying foundation for educators nationwide. The principles were themselves created with over 20 educational partners and 13 federal agencies, providing an educated, diverse, and dedicated focus to the topic of energy literacy (U.S. Department of Energy, 2013).

In an effort to create an assessment that addresses all 7 Essential Principles and emphasizes important bioenergy concepts that have not been included in other assessments, the development panel selected subtopics from each principle that were particularly relevant to bioenergy or renewables. With the foundation of the DOE Energy Literacy document the development panel worked to create and refine outcome statements that specified what students would be able to do as demonstration of mastering a particular energy literacy principle. The principles and resulting outcome statements were also correlated with the Next Generation Science Standards to seed the questions with relevant classroom requirements (an example of the process can be seen below in Table 1: DOE Energy Literacy Principle alignment with NGSS, biofuel connection and outcome statements)

DOE Energy Literacy Essential Principle	NGSS	Biofuel Connection	Draft outcome statements
<b>Physical Science</b>			
1.4 Energy available to do useful work decreases as it is transferred from system to system. During all transfers of energy between two systems, some energy is lost to the surroundings. In a practical sense, this lost energy has been “used up,” even though it is still around somewhere. A more efficient system will lose less energy, up to a	HS-PS3-3 Design, build, and refine a device that works within given constraints to convert one form of energy into another form of energy.	How much energy is lost in the process from tree to jet?	Students will be able to explain why the energy output of a system will always be less than the energy input.

theoretical limit.			
2.5 Movement of matter between reservoirs is driven by Earth's internal and external sources of energy. These movements are often accompanied by a change in the physical and chemical properties of the matter. Carbon, for example, occurs in carbonate rocks such as limestone, in the atmosphere as carbon dioxide gas, in water as dissolved carbon dioxide, and in all organisms as complex molecules that control the chemistry of life. Energy drives the flow of carbon between these different reservoirs.	HS-ESS2-2 Analyze geoscience data to make the claim that one change to Earth's surface can create feedbacks that cause changes to other Earth systems. HS-ESS2-4 Use a model to describe how variations in the flow of energy into and out of Earth's systems result in changes in climate. HS-ESS3-5 Analyze geoscience data and the results from global climate models to make an evidence-based forecast of the current rate of global or regional climate change and associated future impacts to Earth systems.	Foundational energy concept.	Students will explain the role of energy in the flow of carbon between different reservoirs.
2.6 Greenhouse gases affect energy flow through the Earth system. Greenhouse gases in the atmosphere, such as carbon dioxide and water vapor, are transparent to much of the incoming Sunlight but not to the infrared light from the warmed surface of Earth. These gases play a major role in determining average global surface temperatures. When Earth emits the same amount of energy as it absorbs its average temperature remains stable.	HS-ESS2-2 Analyze geoscience data to make the claim that one change to Earth's surface can create feedbacks that cause changes to other Earth systems. HS-LS2-7 Design, evaluate, and refine a solution for reducing the impacts of human activities on the environment and biodiversity.	Discuss disruption in plant populations?	Students will explain how greenhouse gases affect energy flow through the Earth system.

Life Science			
3.1 The Sun is the major source of energy for organisms and the ecosystems of which they are a part. Producers such as plants, algae and cyanobacteria use the energy from sunlight to make organic matter from carbon dioxide and water. This establishes the beginning of energy flow through almost all food webs.	HS-LS1-5 Use a model to illustrate how photosynthesis transforms light energy into stored chemical energy.	Foundational energy concept.	Students will trace energy flows from the sun through food webs that include plants and algae.
3.2 Food is a biofuel used by organisms to acquire energy for internal living processes. Food is composed of molecules that serve as fuel and building material for all organisms as energy stored in the molecules is released and used. The breakdown of food molecules enables cells to store energy in new molecules that are used to carry out the many functions of the cell and thus the organism.	HS-LS1-5 Use a model to illustrate how photosynthesis transforms light energy into stored chemical energy.	This is biofuel can also serve as the power needed for transportation.	Students will explain how energy is transferred between, released by and/or stored by organisms.

Table 1: DOE Energy Literacy Principle alignment with NGSS, biofuel connection and outcome statements.

From these statements, the panel then began drafting test items in a multiple-choice format, following guidelines for best practices (Haladyna & Rodriguez, 2013). A total of 22 items--one for each outcome statement--with three response options was drafted. Over the course of 14 meetings, the panel scrutinized the wording of each item for clarity and reading level, and checked for accuracy of the intended correct answer options. Questions were then vetted by classroom teachers, biofuels energy staff, and age-level representatives.

After revisions made based on feedback from these individuals, a talk-aloud method was employed to gather data on flow and reliability from a student’s perspective. Talk-alouds (or think-alouds) are used to determine the quality and clarity of the questions used to test students. Talk-aloud protocol includes asking the participants to “talk-aloud constantly from the time the problem was presented until they had given their final answer” (Kenny, Marks, & Wendt, 2007,

p. 22). The aim of a talk-aloud is to capture “what is held in the short-term memory” which “results in a sequence of thought that reflect what occurs cognitively during completion of a given activity” (Young, 2005, p.20). The use of a talk-aloud with assessment questions can help to determine if the questions make sense and/or are appropriate for the associated age population (e.g., middle school students) as well as point out any gaps in addressing the targeted construct (e.g., energy literacy).

Four groups of five to eight students per group were asked to participate in a talk-aloud when reading the assessment to determine areas of confusion or define any limitations, as suggested by the National Quality Council (2009). Results from the talk-alouds, feedback from teachers, and comments from industry experts were compiled and analyzed to validate the content of the questions. The incorporation of the results and feedback led to the next round of edits for the questions and answer choices.

## **Results**

### **Item Analysis**

In order to be useful in educational programming, the energy literacy test needs to be able to capture a broad range of energy literacy levels, and differentiate between students who are at high, medium, and low in their levels of energy literacy. In this section, we present the results of an item analysis (Livingston, 2006) designed to assess the functioning of each test question relative to these needs. Analyses employed a classical test theory (Crocker & Algina, 1986) approach to determine if items had appropriate levels of difficulty and discrimination.

### **Sample and Data Collection**

A sample of 508 students ( $n_{\text{male}}=273$ ) at a residential outdoor science school in the northwest United States participated in a pilot administration of the instrument. This sample represented all of the Grade 5 and 6 students who attended the school during a four-month period during the Winter and Spring of 2015. (Approximately 2,500 students participate in the resident program each school year.) Students came from 11 schools. One school was located in a rural area; the rest were located in a suburban setting. All schools drew from a population that is predominantly White and of moderate socio-economic status. Prior exposure to energy education content was limited among the students. Most schools had not engaged in any formal energy curriculum, but basic energy principles had been covered relating to motion, heat, potential energy, and hydropower.

Students completed the instrument at the beginning of their week-long experience at the school. Students were placed into small groups of six to ten per field group instructor. In small classrooms around campus, graduate-student instructors provided the students paper-based test forms containing the following instructions: “Please read each statement carefully and complete it as honestly as possible. You can indicate your answer by circling the letter that corresponds to your answer.” Students were instructed to take as long as they needed and to remain silent until everyone was done. Typical test administration lasted about 20 minutes.

### **Item Difficulty**

An analysis of item difficulty was conducted in order to assess the value provided by the test questions in an applied setting. Items answered correctly by too many or too few students

prior to the introduction of a learning program do not provide much value. That is, if an item is answered correctly by almost everyone, there is little opportunity for the students to show growth in their energy literacy as a result of educational programming. Conversely, if an item is answered incorrectly very often, not enough information is contributed toward knowing the baseline levels of energy literacy the students can demonstrate. Therefore, when the interest is to examine student growth over time, it is ideal for items to demonstrate a moderate level of difficulty (i.e., around 50%). Item difficulty in this study was assessed through the percent of students answering the item correctly.

On average, items were answered correctly by the students 52% of the time. Individual item-correct percentages ranged from 13% to 82% (Table 2). Ten items were answered correctly by 40-60% of the students. Another 10 items were answered correctly by 30% to just above 70% of the students. The vast majority of items were centered around that ideal moderate level of difficulty. Among those outside that range, one was moderately easy (82%) and one was difficult (13%). An inspection of the content and response patterns (i.e., how many students selected each answer option) of the difficult item found it addressed a fundamental nuance representing a high level of energy literacy. The moderately easy item addressed the topic of photosynthesis at a basic level, which is covered in early grade levels (K-4). Both items were judged to be necessary for adequately representing the energy literacy construct. For these reasons, no item was flagged for revision or deletion based on item difficulty analyses.

---

Item	Difficulty (% correct)	Discrimination ( <i>D</i> )
1	44	0.37
2	54	0.38
3	30	0.20
4	67	0.48
5	45	0.25
6	64	0.27
7	61	0.47
8	50	0.37
9	30	0.26
10	82	0.19
11	72	0.39

12	46	0.33
13	71	0.36
14	46	0.38
15	38	0.39
16	45	0.46
17	13	0.14
18	52	0.40
19	36	0.36
20	51	0.48
21	52	0.64
22	33	0.24

Table 2: *Item Difficulty and Discrimination Estimates on the Energy Literacy Inventory-Middle School*

These outcomes suggest the instrument is well positioned to serve in a research or program assessment context. The lack of item difficulties very close to 0 or 100% means student change over time (e.g., before and after an educational intervention) can be assessed without much concern that some level of growth has been “missed”. In other words, performance of the pilot sample suggests most students will be able to demonstrate an accurate baseline level of energy literacy and will have room to grow on the instrument. Furthermore, correlational analyses that investigate the energy literacy instrument in relation to other variables (e.g., broad educational outcomes) will likely not suffer from a restriction of range influence.

### **Item Discrimination**

An analysis of item discrimination was applied to assess the extent to which each question on the ELI-M differentiates between those who know the material and those who do not (Crocker & Algina, 1986). A common method for investigating item discrimination is through the calculation of an index of discrimination ( $D$ ). To calculate  $D$ , the proportion of low-performing students (i.e., those in roughly in the lowest quartile in terms of overall score on the test) answering an item correctly is subtracted from the proportion of high-performing students (i.e., upper quartile) answering the item correctly. Values of  $D$  can range from -1 to +1. Negative  $D$

values indicate more low-performing students answered the item correctly than high-performing students. This situation is, clearly, problematic, and any item associated with a negative  $D$  needs to be checked for accuracy in the answer key and then possible revision or deletion. For positive  $D$  values, a general rule of thumb is that values greater than 0.30 indicate little need for revision (Ebel, 1965). Items with  $D$  values between 0.20 and 0.29 are performing at the margin of acceptability, and should be inspected for basic revisions. Items below 0.20 require a thorough revision or deletion.

Students participating in the pilot administration of the energy literacy instrument were ranked by their total score on the instrument, and two groups were formed. A high-performing group was composed of the 133 students with the highest scores (59% and higher). A low-performing group was composed of the 114 students with the lowest scores (36% and lower).  $D$  values derived from the performances of these two groups ranged from 0.14 to 0.64. The 0.30 threshold was achieved by 15 items. Six items fell in the range of  $D$  values between 0.20 and 0.29. These items may need further inspection as the ELI-M is administered across diverse settings. Two items had  $D$  values below 0.20. These two items represented the easiest and most difficult items on the instrument. As stated above, these items were retained for their contributions to representing the energy literacy construct.

### **Discussion**

The purpose of this paper was to share the development of the NARA Energy Literacy Inventory for middle aged children (ELI-M), and provide support for proscribed interpretations and uses with evidence based on test content and item functioning. To this end, based on a thorough development process that was grounded in a widely acknowledged framework for energy literacy and results from analyses of item difficulty and discrimination, initial evidence for the ELI-M is encouraging. Sufficient support exists for use of the ELI-M in low-stakes research and programmatic settings. Specifically, users of the ELI-M can be confident that 1) test items target students' levels of energy literacy in a way that is conceptually consistent with contemporary thinking in the fields of environmental and energy education, and 2) responses to test items allow for reasonable differentiation of high, medium, and low levels of energy literacy among students. Taken together, the field may begin to apply the ELI-M in an educational research context, to identify programmatic needs, and as an outcome measure to assess impact of programming on individuals and groups. In the latter case, we encourage the use of other forms of outcome measures in conjunction with the ELI-M, especially when decisions are at the level of the individual student.

The ELI-M appears to hold promise for advancing scholarship in energy education. Results from the item analyses show the instrument could function well in a research setting. The ELI-M could therefore be used to improve understanding of how malleable a construct energy literacy is. That is, does it appear short programming in an informal education context can quickly improve students' energy literacy or does it take a more thorough and extended effort to make a difference? In the ELI-M, researchers also have a tool to allow them to explore how energy literacy associates with other constructs such as science content knowledge, logical reasoning, creativity, or general intelligence.

### **Future Directions**

While the present paper does provide evidence in support of the ELI-M, the process of test validation is ongoing, and additional evidence can be gathered across contexts and as energy

literacy assessment needs evolve (Kane, 2013). As mentioned previously, item difficulty and discrimination estimates can be dependent upon the sample from which they were calculated. Further efforts should replicate the analyses across a wide array of settings and student backgrounds, especially students representing minority groups and racially diverse groups. Support for the ELI-M should also be sought from additional analyses. Given the test was developed from a framework that implies different dimensions of energy literacy (U.S. Department of Energy, 2013) an analysis of test structure via factor analysis (e.g., Kieffer, 1999) is warranted. Such an analysis would begin to inform the decision of whether student scores should be disaggregated along the seven energy literacy principles. If support for such uses is found, the implications for scholarship and practice are multiplied, and the ELI-M becomes more valuable. Another source of support could come from an investigation of how the test items function across demographic variables such as gender and ethnicity. A differential item functioning analysis would assist users in their interpretations of energy literacy outcomes observed across groups.

The analyses of item difficulty and discrimination across samples, factor structure, and differential item functioning can also help inform potential revisions of the ELI-M. Six items demonstrated marginal ability to discriminate between students with high and low levels of energy literacy. At this stage in the life of the ELI-M, we recommend retaining these items. Further analyses could show better functioning of these items, providing support for their retention in the instrument as is. Further analyses, however, could suggest additional attention to these items (or even items that performed well in the present study) is warranted. For example, if the marginally discriminating items also do not associate strongly with a single underlying factor or seem to favor boys over girls, that would signal a need to revisit the question stems and answer choices to see if item content can be improved.

Results from the research conducted with the ELI-M indicate that it is a valuable tool for assessing baseline energy literacy and change in energy literacy. In-depth validity testing and item analyses demonstrate the usefulness of the inventory for researchers and educators alike. As the need for energy literacy increases, so too does the need for adequate assessment tools that include consideration for bioenergy. As we continue to see energy economies shift and innovate, so too will the need persist to revise literacy assessments that address content knowledge for evolving energy sources and technologies.

### References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.
- Birol, F. [International Energy Agency]. (2015, June 15). *WEO Special Report on Energy & Climate Change: Part 2 - Presentation* [Video File]. Retrieved from <https://www.youtube.com/watch?v=OHV1bZ-5BLQ>.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. Belmont, CA: Wadsworth Group/Thompson Learning.
- DeWaters, J., & Powers, S.E. (2013). Establishing measurement criteria for an energy literacy questionnaire. *The Journal of Environmental Education*. 44:1 38-55

- DeWaters, J., Qaqish, B., Graham, M., & Powers, S. (2013). Designing an energy literacy questionnaire for middle and high school youth. *The Journal of Environmental Education*. 44:1, 56-78.
- DeWaters, J., Hougham, R. J., Hintz, C., & Frolich, L. (2015). Beyond Conservation: Reimagining the Purpose of Energy Education. *Journal of Sustainability Education*.
- Ebel, R. L. (1965). *Measuring educational achievement*. Englewood Cliffs, NJ: Prentice-Hall.
- Fives, H., & DiDonato-Barnes, N. (2013). Classroom Test Construction: The Power of a Table of Specifications. *Practical Assessment, Research & Evaluation*, 18(3). Retrieved from <https://pareonline.net/getvn.asp?v=18&n=3>.
- Haladyna, T. M., & Rodriguez, M. C. (2013). *Developing and validating test items*. New York, NY: Routledge.
- Hendrickson, D., Corrigan, K., Keefe, A., Shaw, D., Jacob, S., Skelton, L., Hougham, R. J. (2015). Global Sustainability: An Authentic Context for Energy Education. *Journal of Sustainability Education*, 8. Retrieved from [http://www.jsedimensions.org/wordpress/content/global-sustainability-an-authentic-context-for-energy-education\\_2015\\_01/](http://www.jsedimensions.org/wordpress/content/global-sustainability-an-authentic-context-for-energy-education_2015_01/)
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement*, 50, 1–73.
- Kenny, L.E., Marks, C., Wendt, A. (2007). Assessing Critical Thinking Using a Talk-Aloud Protocol. *CLEAR Exam Review*, 18:1, 18-27.
- Kieffer, K. M. (1999). An introductory primer on the appropriate use of exploratory and confirmatory factor analysis. *Research in the Schools*. 6:2, 75-92.
- Livingston, S. A. (2006). Item analysis. In S. M. Downing and T. M. Haladyna (Eds.), *Handbook of test development* (pp. 421-441). New York, NY: Routledge.
- Schon, J., Eitel, K., Hougham, R. J., & Hendrickson, D. (2015). Creating a research to classroom pipeline: Closing the gap between science research and educators. *Journal of Sustainability Education*, 8. Retrieved from [http://www.susted.com/wordpress/content/creating-a-research-to-classroom-pipeline-closing-the-gap-between-science-research-and-educators\\_2015\\_01/](http://www.susted.com/wordpress/content/creating-a-research-to-classroom-pipeline-closing-the-gap-between-science-research-and-educators_2015_01/)
- Toulmin, S. E., Rieke, R. D., & Janik, A. S. (1979). *An introduction to reasoning*. New York, NY: Macmillan.

United States, U.S. Department of Energy. (2013). *Energy literacy: Essential principles and fundamental concepts for energy education*. Washington, D.C.: U.S. Department of Energy.

Young, K.A. (2005). Direct from the source: the value of ‘think-aloud’ data in understanding learning. *Journal of Educational Enquiry*, 6:1, 19-33.