Spring 5-19-2016

Evaluating an Assessment Instrument for the Oregon Environmental Literacy Plan

Susan Mae Duncan
Portland State University

Evaluating an Assessment Instrument for the Oregon Environmental Literacy Plan

by

Susan Mae Duncan

A dissertation submitted in partial fulfillment of the
requirements for the degree of

Doctor of Education
in
Educational Leadership: Curriculum and Instruction

Dissertation Committee:
Swapna Mukhopadhyay, Chair
Samuel Henry
Dilafruz Williams
Cary Sneider
Marion Dresner

Portland State University
2016

Abstract

This mixed methods study evaluated the validity, and reliability of an instrument designed to assess a middle school student's proficiency in systems thinking as described in the 2010 Oregon Environmental Literacy Plan. In Stage 1, a forum of middle school students, formal, and non-formal educators used the Delphi technique to reach consensus regarding which skills were important to include in a scoring guide for systems thinking. In Stage 2, the scoring guide was field tested by formal and non-formal educators using a sample of students' work. The two groups' scores were compared using Cohen's kappa to make inferences regarding inter-rater reliability. Concurrently, an autoethnographic narrative was written to explore issues of equity related to the assessment of environmental literacy.

The commonalities between formal and non-formal educators revealed a high level of validity for the construct of proficiency with systems thinking, and a moderate level of reliability between the scores assigned by two groups of educators. In the words of the middle school students, formal, and non-formal educators, who volunteered to create the scoring guide, the ability to make responsible decisions with natural systems, community, and the future in mind involves: creating solutions for systems that are not in balance, presenting the complex inner workings of a system in a simple and succinct way, collaborating, exploring multiple solutions, and sharing ideas in a way that people will understand you.

Dedication

This work is dedicated to my parents, Elmore and Elizabeth Duncan, my brother Steven, my sister Kathleen and her husband Gary Lundquist, because we have constructed an understanding of family that honors the infinite possibility of each life in the universe with respect for our human frailty. I dedicate this work to Neera Malhotra and Ryan Gregory, whose kindness and compassion are second only to the generosity of the deep, fertile mollisols beneath our feet. The soils that feed the plants that love us. I dedicate this work to Rose High Bear and D'Ana Soto for the opportunity to be led by a child to a beaver's den, stand in the bed of a creek with young leaders, and heap soil around berry bushes in preparation for the change of seasons.

Acknowledgements

I would like to acknowledge Samuel Henry for venturing out to my students' neighborhood on a cold, rainy night to invite me to apply to the Graduate School of Education Doctoral Program at Portland State University. His willingness to think in systems and metaphors about education provided me with a strong example of political and social leadership. When I look at the stars, I will remember Cary Sneider's deep fascination with the universe and appreciation for engineers who taught me how science knowledge frees one's creative potential. I appreciate the example of Dilafruz Williams, whose ideals are firmly planted in the school gardens of our city. She is committed to putting the knowledge of food into children's hands. Marion Dresner was the first to value the resources for teachers that a displaced non-formal educator like me had to offer those in the university setting, and I am so grateful for her respect. I also wish to thank Swapna Mukhopadhyay for her kind regard and support of my growth as an educator. This work is just a token of my appreciation to the formal and non-formal education communities that welcomed me into their hearts when I needed a warm fire or a cool drink.

Table of Contents

List of Tables

List of Figures

Glossary

Environmental education. In 1977, the Tbilisi Declaration defined environmental education as the knowledge, skills, attitudes, and beliefs to act on behalf of the environment. Environmental education was preceded by conservation education and the study of natural history. For the purposes of advancing the field in Oregon, the term *environmental literacy* was used to encompass the learning objectives of conservation educators and environmental educators. It also included educational efforts described as forest literacy and ocean literacy. Vocational educators in Oregon developed strong relationships with the education and outreach programs of the forest products companies and agribusiness. The resource directory of the state environmental education association lists over 150 organizations across Oregon that provide environmental education to the public (EEAO, 2016).

Environmental literacy. The operational definition of environmental literacy at the 12th-grade level for proficiency is defined by the 2010 OELP as "an individual's understanding, skills and motivation to make responsible decisions that consider his or her relationships to natural systems, communities and future generations" (p. 4). Five environmental literacy strands outline what a student knows and is able to do as a demonstration of environmental literacy. *Strands* and *academic standards* describe learning objectives that can be assessed by educators as students demonstrate they have met them.

Assessment instruments. Different types of instruments are used to measure students' abilities, depending on the purpose of the assessment. National standards have been used to develop surveys, tests, and scoring guides to measure the level at which students are meeting learning objectives. The Next Generation Science Standards use the term *performance indicators* to describe what needs to be shown to the educator assessing a student for skills and content knowledge. The assessment instrument used in the proposed research is a scoring guide that includes statements that educators can use to recognize a student's level of proficiency in using systems thinking skills to describe their rationale for decisions.

Scoring guide. A scoring guide uses an ordinal scale to measures a student's growth over time using descriptive categories. It is used by a student to create and self-evaluate a work sample. Educators use the same scoring guide to plan instruction and measure proficiency. *Proficiency* is the term used to describe a work sample or performance that is at the level of expectation. This particular scoring guide will be designed to measure the *construct* of environmental literacy. A construct is a particular idea that a person has regarding a particular phenomenon.

Preface

This dissertation is original unpublished work of the author, S. Duncan. The data collected from middle school students and the educators who teach them, reported in Chapter 4, was covered under Human Subject Research Review Committee approval for Mukhopadhyay-Duncan #132891. Approval was also granted for the research from the public school district in which the students and teachers worked. Cary Sneider suggested the Delphi technique as a method for creating a scoring guide. The tools used in the Delphi technique were adapted with permission from Wiley, the publisher of *The Delphi Technique in Nursing and Health Research* by Sinead Keeney, Felicity Hasson, and Hugh McKenna (2011). The archetypical models used to state the problem and propose solutions for this dissertation were derived from the work of Donella Meadows (2008), specifically the chapter titled "Systems Traps . . . and Opportunities" in *Thinking in Systems: A Primer.* The Waters Foundation's *Systems Thinking in Schools Modules* created by WebEd (2006) at www.watersfoundation.org/webed were provided as a link to the middle school students, as well as the educators who developed the scoring guide, so they had a reference for deepening their understanding of systems thinking. The phrase "more than human world" came to my attention while reading *Braiding Sweetgrass: Indigenous Wisdom, Scientific Knowledge, and the Teachings of Plants* by Robin Wall Kimmerer (2013), published by Milkweed Editions. Her writing about "reciprocity" restored my worldview and healed my spirit as I organized my autoethnographic data.

CHAPTER 1

Introduction

Pipher (2013) writes, "The breaking of silence surrounding global climate change gives me hope that, at last, we as a society might have a conversation about our beloved planet" (p. 7). In her publication, *The Green Boat: Reviving Ourselves in Our Capsized Culture,* she suggests "it is the most practical among us who come out of denial first. . . . They must wake up in order to do their jobs" (p. 6). The job of educating the public, and students in public schools, involves providing the "knowledge, skills, and experiences" (Kleckner, 2010, p. 53) necessary to engage in conversations about complex interactions in natural systems and communities that impact future generations. In 2012, representatives from 98 countries gathered to reaffirm the original vision for environmental education in the *Tbilisi Communiqué: Educate Today for a Sustainable Future:*

> The objectives outlined at the 1977 Tbilisi Conference – namely awareness, knowledge, attitude, skills and participation – are still valid today, and the main goal – aligning human behaviours, actions, practices and social conditions towards a sustainable future – has yet to be achieved. (*Tbilisi Communiqué*, 2012, p. 1)

The purpose of this chapter is to value educators' efforts to realize the main goal of the *Tbilisi Conferences* using Gough's (2013) four principles from the *International Handbook of Research on Environmental Education*:

- to recognize that knowledge is partial, multiple, and contradictory;
- to draw attention to racism and gender blindness in environmental education;
- to develop a willingness to listen to silenced voices and to provide opportunities for them to be heard; and
- to develop understandings of the stories of which we are a part and our abilities to deconstruct them. (p. 10)

In this chapter, the following key information is provided in preparation for evaluating an assessment instrument to measure a middle school students' proficiency with systems thinking for environmental literacy:

1. The Oregon Environmental Literacy Plan's (OELP) systems thinking strand gives educators the opportunity to accept responsibility for teaching and assessing middle school students' decision-making skills. Since decision-making skills are an integral part of educators practice and students' instructional time, educators and students can learn to, not only make sense of experiences and data which are "partial, multiple, and contradictory," (Gough, 2013, p. 10), but act to "[align] human behaviors, actions, practices and social conditions towards a sustainable future" (Tbilisi Communiqué, 2012, p. 1).
2. Systems thinking concepts and tools give middle school students the skills to model systems and use decision-making processes in conversation based on one's own frame of reference. Like the keys to the car, the tools are meant to give middle school students a vehicle through which they can make their voices heard.
3. "Artificial boundaries" describes by former Oregon Governor Kitzhaber (2013) between Career and Technical Education (CTE) and Science, Technology, Engineering and Math (STEM), have influenced the perceived value of different kinds of educational experiences. Artificial boundaries between formal and non-formal educators are unintended consequences of national funding for vocational education though the Smith-Hughes Act of 1911. There is work for educators to do in reconciliation for the historic injustices to people assigned to vocational

education rather than both vocational and general education without respect for culture and natural systems.

4. Citizen science and teaching practices that value both inductive and deductive reasoning are motivating educators and students in reconstructing their communities. By using observational and logical methods of inquiry, both separately, and in tandem to make decisions, students realize their function in the systems of which they are a part.

A Plan for Environmental Literacy in Oregon

Oregon's Environmental Literacy Task Force was created by House Bill 2544—the No Oregon Child Left Inside Act (NOCLI). The NOCLI Act was signed into law on July 22, 2009. Oregon's governor at the time, Ted Kulongoski, indicated that NOCLI would "provide youth with classroom instruction about our vital natural resources and an opportunity to conduct field investigations in an outdoor learning setting" (OELP, 2010, p. 3). Chairperson Traci Price of the Freshwater Trust and the Environmental Education Association of Oregon (EEAO) led the process of writing the OELP as mandate in H.B. 2544. The plan made Oregon eligible for federal dollars when the federal No Child Left Inside Act is approved as part of the reauthorization of the National Elementary and Secondary Education Act. The Oregon Department of Education (ODE) facilitated the public meetings of the task force, whose members were appointed from the Department of Environmental Quality, Department of State Lands, Department of Fish and Wildlife, Department of Forestry, Department of Agriculture, the Parks and Recreation Department, and Metro Regional Government, as well as Oregon State University's College of Science and Sea Grant program. Citizens and representatives, as well as formal and non-formal educators, chose to serve as members of the working groups created by the task force. Formal educators included those who provide students with credit through school districts, the Oregon University System, and community colleges. Non-formal educators included those who provided students and their families with educational experiences through entities such as the Oregon Forest Resources Institute, Oregon Zoo, Oregon Coast Aquarium, and World Forestry Center. The working groups included: Educational Standards and Diploma Requirements, Teacher Professional Development, and Implementation and Assessment. The 2010 OELP calls for the kind of science literacy described by Osborne (2007) that "considers plural alternatives" and exemplifies "system-wide commitment" as suggested by Cutter-Mackenzie and Smith (2003).

The objectives of the international accord described in the 2012 *Tbilisi Communiqué* resonate with the No Oregon Child Left Inside (NOCLI) task force's definition of environmental literacy, which was adopted by the Oregon Legislature in 2010:

> An individual's understanding, skills and motivation to make responsible decisions that considers [one's] relationships to natural systems, communities and future generations" (Oregon Department of Education, 2010, p. 4).

The NOCLI task force identified Environmental Literacy Strands that "articulate a comprehensive content and skills learning framework" (p. 16). Upon graduation from high school, an environmentally literate student would be able to use systems thinking concepts and tools for making decisions:

> Understand and apply systems thinking concepts and tools
> a. Systems as context for thinking and action
> b. Implications and consequences
> c. Strategic responsibilities of systems thinking
> d. Shifting mental models and paradigms. (p. 16)

The tools for using systems as a "context for thinking and acting" included:

- Behavior over time graphs
- Connection circles
- Causal loops
- Stock flow diagrams
- Modeling (p. 19)

By including systems thinking concepts and tools in the learning framework, the Oregon Environmental Literacy Plan guides educators in designing learning experiences and assessments that match middle school students' ability to imagine possibilities and act on their beliefs (see Appendices A-F for templates of systems thinking tools from ©2015 Waters Foundation, Systems Thinking in Education, watersfoundation.org). Systems thinking is, not simply a problem-solving tool, but encourages each middle school student to voice the reasoning behind one's willingness to be a thriving solution acting in community with natural systems.

Systems Thinking Concepts and Tools

In a first read of the learning framework in the OELP, proficiency with systems thinking might seem too complex for middle school students to master, but considering the magnitude and complexity of natural phenomena and social structures in which we live, educators have a responsibility to share approaches to decision-making known to cross academic boundaries in the sciences. Golley (1998), who wrote *A Primer for Environmental Literacy,* explains,

> In our [Western] culture the context is either assumed (and thus of little interest) or is considered outside the area of responsibility of the problem solver. Because of this bias, the methods of analysis are exceptionally well developed in all the sciences. The only synthetic method developed in science has been systems dynamics. It did not develop rapidly until the advent of the computer; now it is a widely used methodology. (p. 16)

With, or without computers, systems thinking concepts and tools help individuals synthesize information and articulate the kind of reasoning for decisions characteristic of environmental literacy.

It was a middle school student's Iceberg Model recommending we take action to reduce carbon dioxide emissions that inspired me to continue learning more about systems thinking and do this investigation (see Figure 1). The Iceberg Model is a composite of several salient systems thinking tools: Behavior-Over-Time graphs (BOTGS), Connection Circles, Causal Loops, and the Ladder of Inference. These tools turn one's attention to the patterns between elements in systems that underlie events. Peter Senge, who co-authored, *The Triple Focus: A New Approach to Education* with Daniel Golman in 2014 and stressed why systems thinking matters:

> There are certain cognitive abilities that are anchored in a broader and deeper awareness. That awareness of caring and. . . . [an] awareness, you know, that my actions really matter—that awareness of self. (Senge, 2014)

Senge's explanation for awareness of caring, and self, is also affirmed by Seymour (2004) in his interview with Nel Noddings who edited, *Justice and Caring: The Search for Common Ground in Education.* Seymour asked Noddings:

> You speak of care for self, for intimate others, for associates and acquaintances, for distant others, for nonhuman animals, for plants, and the physical environment for the human-made world of objects and instruments. In your own words, and at the risk of belaboring the obvious, what suggests that caring should be the essential purpose of education?

Noddings, responded:

> Life itself suggests this. When you look at what is important to people—what really matters to them—all these things you just mentioned are things that really matter to people. Without acknowledging that, without having care at the center of our lives, I think life becomes superficial. . . . We can do better than that and can help kids to connect with themselves. (p. 90)

The Iceberg Model is the systems thinking tool that helps students share and synthesize thoughts and ideas that originate deep in their hearts, especially those most difficult to articulate, where actions speak for themselves.

The Iceberg Model can be used to illustrate, common mental models, or archetypes, so one's thinking can be understood and refined. The Ladder of Inference is found at the base of the iceberg model and is used to show one's own beliefs. As Janice Jackson shared in her presentation of the *Cultural Iceberg* at Camp Snowball in Portland, Oregon in July 2015, decision-making shares a sacred place with belief at the bottom of the iceberg model. Decision-making happens below familiar patterns of cause and effect and some of the archetypical mental models associated with them. Archetypical mental models are patterns of thinking that perpetuate the behavior of systems, even behaviors that we might wish to change like drifting to low performance due to pressures from forces that cause us to lower our goals (as shown in Appendix E). The questions shown in the Iceberg Model in Figure 1, and the individual systems thinking tools, which are described in more detail in the following pages, can help each of us work together to bring systems into balance.

Figure 1. Iceberg Model with Guiding Questions for Systems Thinking

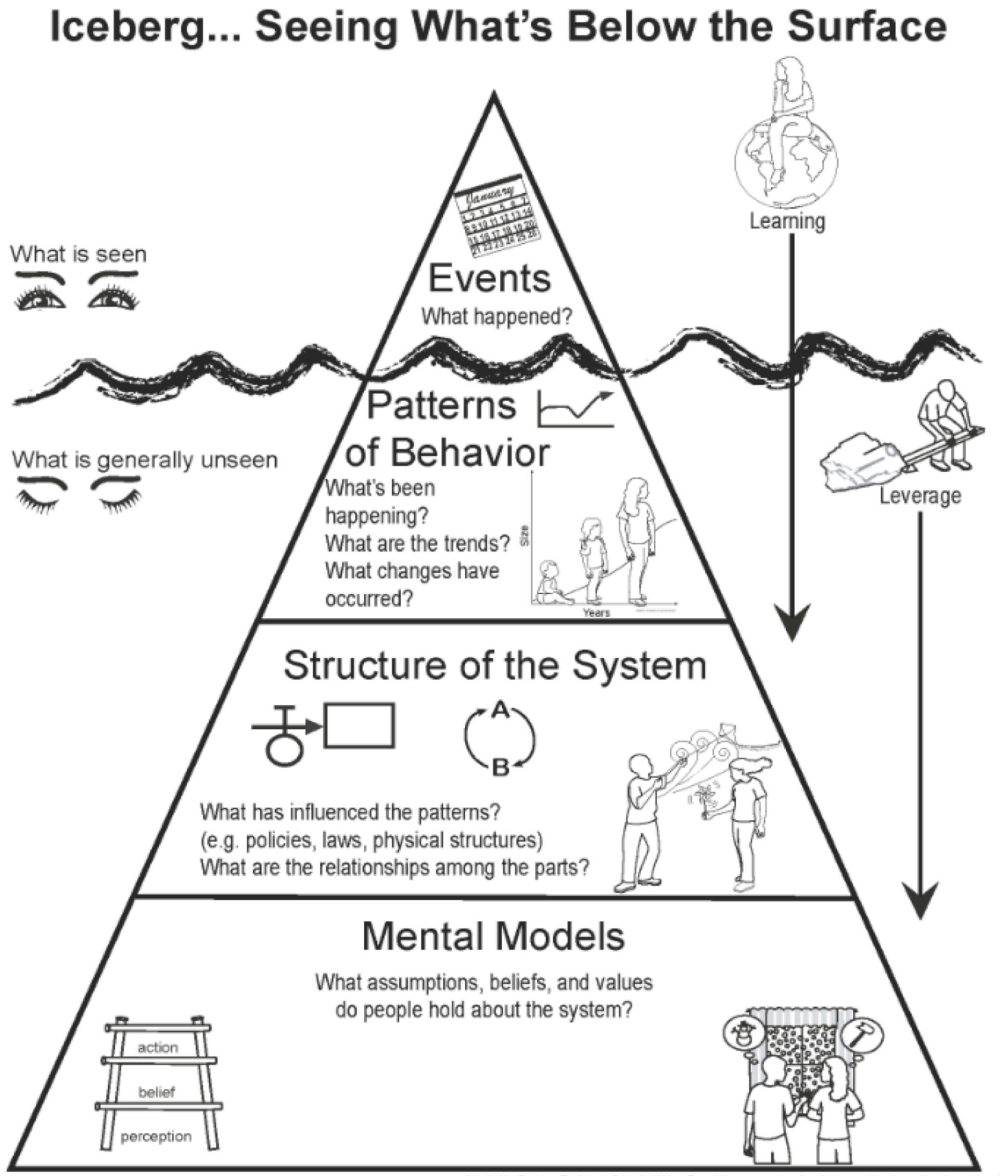


Figure 1. Iceberg model illustrating the organization of information from the application of systems thinking tools. People notice events and learn about the patterns in systems that generate those kinds of events. By understanding the one's underlying beliefs, one can take actions that leverage systems into balance. The iceberg model template without guiding questions is included in Appendix A. Adapted by Systems Thinking in Schools, Waters Foundation, www.watersfoundation.org, from Innovation Associates, Inc. Reprinted with permission.

Behavior-Over-Time graphs (BOTGs). BTOG's show general trends in relationships between elements in systems based on observed data, measurements, or experiences so individuals can imagine how a systems' behavior changes over time (see Figure 2). By using time as a common factor in BOTGs, changes in multiple elements can be compared and extrapolated into the future based on one's understanding of multiple trends. BOTGs are a systems thinking tool that can be used to make decisions that consider connections in natural systems and communities together.

Figure 2. Behavior Over Time Graph—Change in Atmospheric Carbon Dioxide

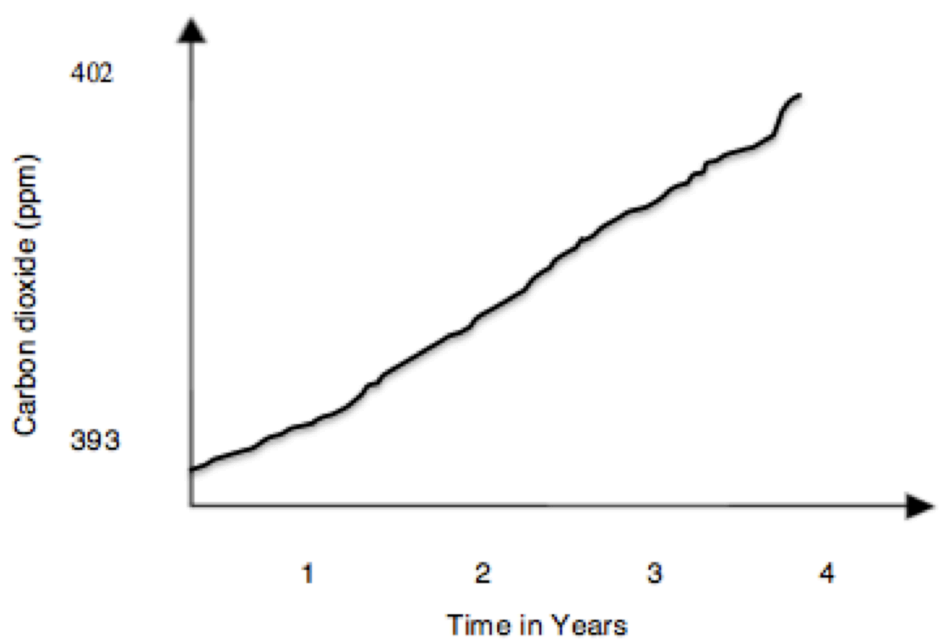


Figure 2. The important element that has changed over time is the amount of carbon dioxide in the atmosphere. The line shows the change in the level of atmospheric carbon dioxide from May 2012 to May 2016 using real time measurements in parts per million. The amount of atmospheric carbon dioxide has increase by 9 ppm over a 4-year period. The rate of increase is 2.25 ppm per year. In order to decrease the slope, action need to be taken to reduce the amount of carbon dioxide added to the atmosphere at its source. The carbon dioxide readings were collected from current and past real-time data from *Global Climate Change: Vital Signs of the Planet*, NASA. Data for graph retrieved from http://climate.nasa.gov/. Created using the guiding questions in the BOTG template in Appendix B. ©2015 Waters Foundation, Systems Thinking in Education, watersfoundation.org.

Connection Circle. Connection Circles show interrelationships between elements in systems. Each system has a function supported by the flow between elements in the system. In Figure 3 the flow is heat energy, and the function of the system is to transfer heat from one element to another to maintain balance. The elements in a system

Figure 3. Connection Circle Showing Flow of Heat Between Elements on Earth

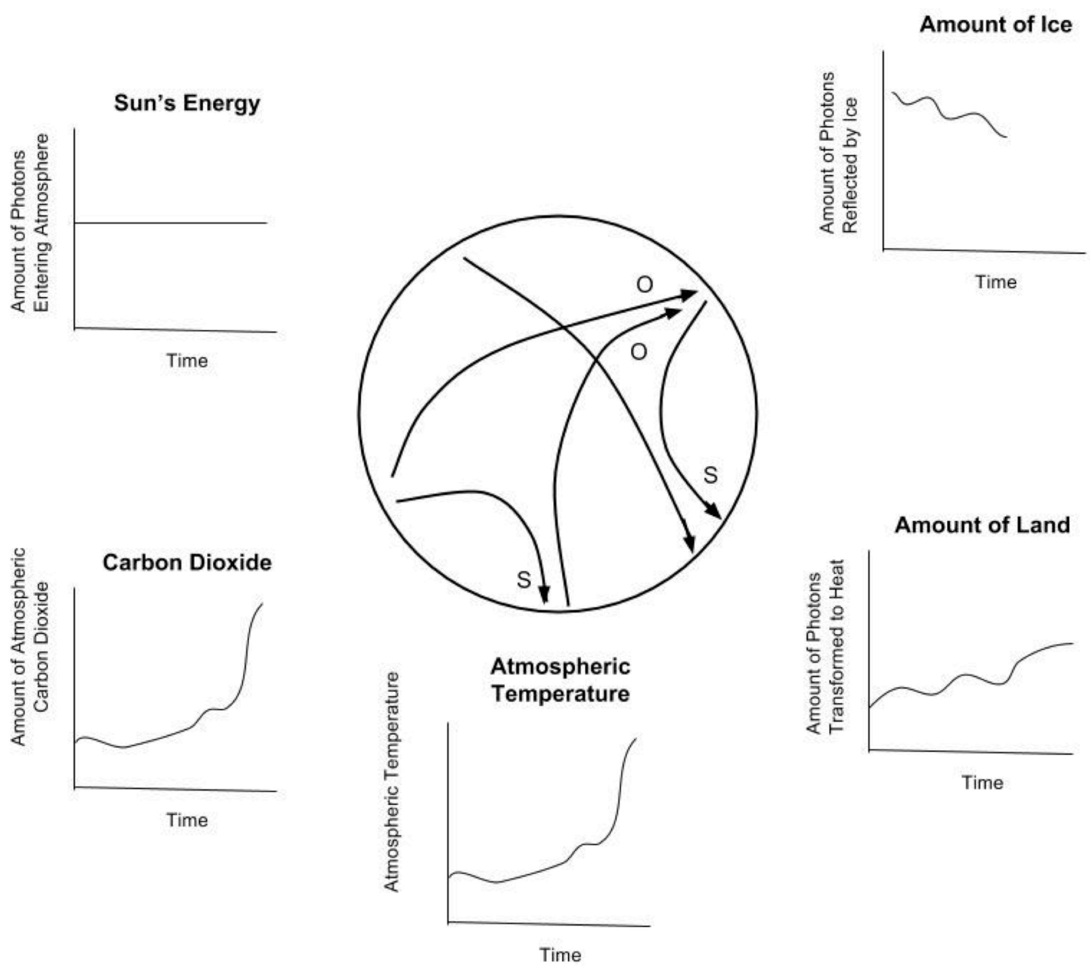


Figure 3. Connection circle showing the transfer of heat between the sun, land, atmosphere, and ice caps. The symbol "S" on the line connecting carbon dioxide and temperature indicate that as one increases so does the other. Alternatively, as temperature increases, the amount of ice decreases, which is represented b the symbol "O". The BOTGs are included so the relationships between the elements can be easily seen. Created using the Connection Circle template available through Waters Foundation, Systems Thinking in Education, watersfoundation.org.

are joined around a circle with arrows indicating flows. Arrows that have the symbol 'S' indicate that the elements are either both increasing, or both decreasing. The symbol "O" indicates that as one element increases, the other element decreases. The efficacy of a Connection Circle is that one can begin anywhere and follow a path between elements in the system to gain a deeper sense of the influence of each element in the system on the function of the whole system.

Causal loop. By choosing an element in a Connection Circle as a starting point, causal loops can be found by tracing the arrows between elements. One of the unexpected outcomes of identifying causal loops, is that missing connections can be found! For example, in Figure 3, as the atmospheric temperature increases, the amount of glacial ice decreases, and the amount of land increases. More land will absorb more of the sun's energy, which increases atmospheric temperature. Figure 4 shows a reinforcing loop in

Figure 4. Causal Loop Diagrams for Flow of Heat in Earth's Systems

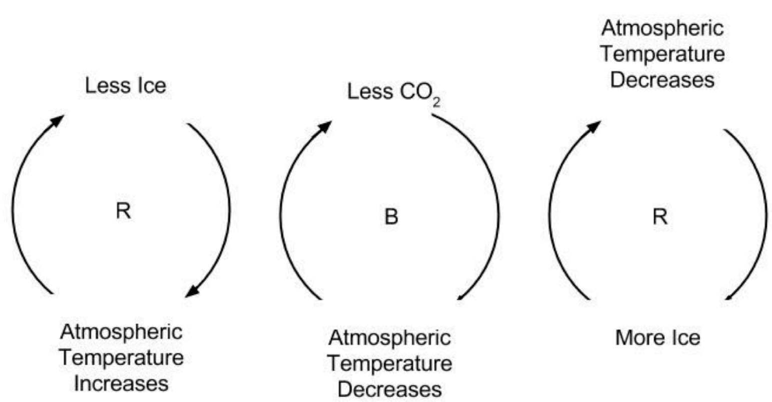


Figure 4. Series of reinforcing and balancing loops regulating heat flow. These loops describe a series of behaviors in Earth's system that reinforce or balance flow over time. They can be combine to create a model of the Earth's system for regulating the transfer of heat between the sun, land, ice, and the atmosphere. Created using the Causal Loop template in Appendix C. ©2015 Waters Foundation, Systems Thinking in Education, watersfoundation.org.

the flow of heat through Earth's Systems. If measurements are available, the changes can be modeled using equations, but in either case, these causal loops can be combined to show mutual causation. An example of mutual causation for global climate change might include heat absorbed by the earth's surface and heat absorbed by atmospheric carbon dioxide from a number of different biological and technological processes. Causal loops can be combined to show common mental models used to understand systems interactions.

Archetypes. Mental models, or archetypes, such as the Tragedy of the Commons, generate their own behaviors. For example, carbon dioxide emitting decisions without consideration for the actions of others can cause temperature increases from overshooting the limits of Earth's system to transfer heat energy. Simulations found in environmental education curriculum often use archetypical models to show interconnections, identify potential leverage points, and refine possibilities. The dynamic balance of wildlife populations and harvesting of renewable resources requires understanding of the regeneration limits of populations (see Appendix D for the Tragedy of the Commons template that could be used in a variety of contexts). The Waters Foundation (2010) suggests using computer modeling software, such as STELLA, to predict what will happen if different decisions are made to change the level of accumulation and amount of flow in a system (p. 6). Archetypes are models that can be used to find leverage points regarding actions that can optimize the system. Donella Meadows (2008) suggests ways to get out of archetypical ways of thinking and avoid "systems traps" like the Tragedy of the Commons (see Appendix D): "1. Educate and exhort, 2. Privatize the commons, and

3. Regulate the commons" (p. 119). For the archetype labeled "Drifting Goals" (see Appendix E), her suggestion is to:

> Make goals sensitive to the best performance of the past, instead of the worst. If perceived performance has an upbeat bias instead of a downbeat one, if one takes the best results as the standard, and the worst results as only a temporary setback, then the same system structure can pull the system up to better and better performance. (p. 123)

These two archetypes are useful to middle school educators and students who care about environmental literacy and creating assessments that represent proficiency with systems thinking.

Ladder of Inference. Climbing the Ladder of Inference is a tool that students and educators can use to make their reasoning visible. It illuminates the meaning one assigns to the information that comes to one's attention in the process of making decisions (refer to Appendix F). It is a helpful tool for comparing one's own understanding of existing mental models. The Ladder of Inference acknowledges as Gough (2013) did that "knowledge is partial" (p. 10) so each person can own one's experience as they collaborate and share ideas with others.

Crossing Boundaries in order to Assess Environmental Literacy

To support formal and non-formal educators, the North American Association of Environmental Education (NAAEE) began the Guidelines for Excellence Project in 1993 (McCrea, 2010). NAAEE commissioned Bora Simmons from Northern Illinois University to develop standards and write a series titled *Guidelines for Excellence in Environmental Education*. The guidelines included evaluations of curricular materials, K–12 learner guidelines, professional development, non-formal programs, and early childhood environmental education programs. The project documented clear connections between environmental education and national academic learning standards in all subjects. Over the past decade, educators in Oregon have taken an active role in advancing environmental education guidelines in their practice in educational settings, including schools, zoos, museums, field study sites, and natural history areas. Two common agreements have surfaced: the need for citizens to understand science and the need to personally interact with natural systems. Educators are encouraged to support students by encouraging them to collect and analyze data, especially data from multiple sources. The role of personal experience when making decisions as a member of the public who "understands the relation between its well-being and the health of natural systems" were also acknowledged (Orr, 1997, p. 90). At the national level, the *Excellence in Environmental Education—Guidelines for Learning (K-12),* which set expectations for performance and achievement, was last published in 2010, clearly explains the purpose of "thinking in terms of systems":

> We are asking individuals to go beyond the fact by fact, piece by piece examination of our environment and begin to understand and think in terms of systems bound together…. Environmental education must play an integral role throughout our educational systems – at the national level, at the state level, and in each an every classroom. (North American Association of Environmental Education [NAAEE], p. 3)

Systems thinking, as it is outlined in the OELP, with its concepts and tools, provides a background for measuring environmental literacy that is not specifically defined in the Next Generation Science Standards (NGSS), and were adopted by the Oregon State Board of Education in 2014. The Waters Foundation (2016) indicates:

> Systems thinking concepts are strongly infused within the Next Generation Science Standards (NGSS) for both grade specific contexts as well as crosscutting concepts. (para. 7)

According to NGSS (2013), the crosscutting concepts "provide students with connections and intellectual tools that are related across the differing areas of disciplinary contents" (p. 1). The Waters Foundation (2016) established a match between systems thinking concepts and tools and the following crosscutting concepts in the NGSS:

> *Patterns*
> *Cause and effect*: Mechanism and explanation
> Scale, proportion, and *quantity*
> *Systems and system models*
> Energy and matter: *Flows*, cycles, and *conservation*
> *Structure and function*
> *Stability and change* (para. 10)

Sneider (2014) established substantive matches between the standards in the NGSS and the *Excellence in Environmental Education—Guidelines for Learning (K-12.)* However, the Environmental Education Strand 3.2 for Decision-making and Citizenship Skills was an exception. It constitutes:

> A) Framing and evaluating personal views
> B) Evaluating the need for citizen action
> C) Planning and taking action
> D) Evaluating the results of actions. (p. 8)

Fortunately, Environmental Education Strand 3.2 Decision-making and Citizenship Skills are skills associated with the skills of systems thinking.

The systems thinking environmental literacy strand in the OELP moves educators toward common goals associated with science literacy, environmental science literacy and environmental literacy by accepting responsibility for teaching and assessing decision-making skills for middle school students (see Appendix G for a correlation of environmental literacy strands and the NGSS). So, although the NGSS specifically notes, "Science does not make the decisions for the actions society takes" in *Earth and Human*

Activity MS-ESS3-4 (NGSS Lead States, 2013, para. 1), the 2013 OELP systems thinking environmental literacy strand emphatically asks students to "consider issues fully, resisting the urge to come to a hasty conclusion" (p. 17). The environmental literacy strand for systems thinking further asks students to reflect on their actions:

(1) "Check results and change actions as needed (successive approximation)

(2) "Monitor system outcomes, and make adjustments where necessary to maintain or improve desirable conditions" (p. 17).

The systems thinking strand of the OELP supports educators pursuing instructional goals "aligning human behaviors, actions, practices and social conditions towards a sustainable future" (Tbilisi Communiqué, 2012, p. 1).

Since this investigation focuses on the work of formal and non-formal educators committed to teach middle school students, it is of the utmost necessity to bring the voices of the two groups together because they care for the same students, and at the same time students care for them. This group of educators worked together to draft and publish the Oregon's Environmental Literacy Plan (OELP), which introduced a learning framework and assessment structure for implementing environmental literacy with educational policy that redefined the role of formal educators, such as science teachers, and non-formal educators, such as Outdoor School instructors. According to Elder (2003) the tradition in environmental education had been to separate formal, non-formal, and informal sectors. Elder (2003) defined formal environmental educators as those working in educational systems that give students credit (p. 7). The non-formal sector represents an educational setting such as "nature/environmental centers, camps and resident outdoor programs, museums, zoos and aquaria, gardens and herbaria, and parks and other natural areas protected or managed by government agencies" (p. 52). Informal sectors are "not tied to any specific setting, and involve electronic and print media, the Internet, materials distributed by non-governmental organizations (NGOs) and community events" (Marcinkowski, et al., 2012, p. 52). The OELP suggests that non-formal educators are in a position to help students reach proficiency levels that may translate into credit by using assessment instruments for environmental literacy:

> The Oregon State Board of Education voted to adopt new high school graduation requirements…designed to better prepare each student for success in college, work, and citizenship. To earn a diploma, students…will also have the option to earn credit for proficiency. (ODE, 2010, p. 5)

The earlier categorization of educators as formal and non-formal as defined by Elder shifted when credit for proficiency gave students the opportunity to generate evidence either "inside the classroom [or] outside the classroom" (ODE, 2011, p. 6). The OELP makes it clear that "credit for proficiency is acutely suited to support education for environmental literacy as a vehicle in student pursuit of the Oregon Diploma" (p. 6). Since credit can be earned by demonstrating proficiency either "inside the classroom or outside the class room," educators in Oregon now share the responsibility for assessing students' proficiency with environmental literacy.

Those who get up in the morning to do the jobs associated with science literacy, environmental science literacy, and environmental literacy share one characteristic in common. Regardless of context, one's perceptions, beliefs, and actions shape the communities in which middle school students live and learn decision-making skills. Science literacy emphasizes the need for teaching knowledge and skills to citizens so they can skillfully and critically examine multiple data sets to make responsible decisions. Environmental science inspires people to spend countless hours with nature collecting observational data that will help us make sense of natural phenomena such as the fragile, energy connections between sunlight, plants, insects, animals, combustion, and the concentration of atmospheric carbon dioxide. It is worth noting that the whole time we are observing and measuring natural phenomena, those same phenomena are functioning to sustain us with simple acts like plants producing oxygen and storing sugar that we can eat later as mangos and berries. These natural phenomena give us reasons to embrace the precautionary principle. Kriebel et al. (2001) outline four central aspects of the precautionary principle:

1. Taking preventative action in the face of uncertainty
2. Shifting the burden of proof to the proponents of an activity
3. Exploring a range of alternatives to possible harmful actions
4. Increasing public participation in decision-making (p. 1)

Previous researchers have documented, the benefits of learning from the environment, and educators with knowledge, skills, and experiences that have deepened both their understanding natural systems and a willingness to live sustainably. Environmental education researchers Lieberman and Hoody (1998) identify the specific benefits to students when using the Environment as an Integrating Context (EIC):

> In addition to traditional subject-matter knowledge and basic life skills, EIC students gain a wealth of added educational benefits including: a comprehensive understanding of the world; advanced thinking skills leading to discovery and real-world problem-solving; and, awareness and appreciation of the diversity of viewpoints with a democratic society. (p. 2)

Bartosh (2003) replicated Lieberman and Hoody's research from 1998 when she paired 100 schools in Washington State—those that used the environment as a context for learning, and those that did not. She found similar results when using scores on academic subject area tests assigned by formal educators who assign credit for academic standards. Science literacy, environmental science, and environmental education are components of conversations that educators use as they cross academic boundaries to, not only support students in real-world problem-solving, but also make environmental literacy an integral part of instructional time.

Rationale for the Study

The purpose of this investigation is to environmental literacy for middle school students from two different vantage points: formal and non-formal educators. In order to make valid and reliable claims regarding middle school students' level of proficiency in systems thinking for environmental literacy, the instrument needs to meet three criteria:

1. Construct validity: Match the instrument with the learning strands defined in the Oregon Environmental Literacy Plan.
2. Reliability: Use the instrument with a high level of reliability between formal and non-formal educators.
3. Equity: Provide individual middle school students with the freedom to demonstrate proficiency using one's understanding of community, natural systems, and the future in the places they live.

The probability of creating a valid understanding for the construct of proficiency with systems thinking for environmental literacy at the middle school level will be higher if it is defined together by formal and non-formal educators working in science literacy, environmental science literacy, and environmental education although environmental literacy is fundamentally and interdisciplinary endeavor. It can be field tested with a sample of middle school students' work to determine the reliability of scores between educators. Additionally, in order to be equitable, middle school students need to have a voice in the process used to assess their decision-making skills. This investigation, attempting to fuse academic and hierarchical relationships between students and educators, shows that middle school students are capable of making "responsible decisions that consider [one's] relationships to natural systems, communities and future generations" (Oregon Department of Education, 2010, p. 4).

Research Question

The more time individuals spend in communities that are in balance with natural systems, where decision-making processes are transparent, the more likely one is to learn and practice those skills. Because many middle school students in Oregon are scheduled to participate in instruction provided within the four walls of a classroom, the research question was written to evaluate an assessment tool that was already familiar to students. The question tests the assumption that a scoring guide would be an effective measure for environmental literacy based on experiences with natural systems and communities inside and outside the classroom. The commonalities between educators who work in both settings would reveal not only the feasibility of measuring students' proficiency in systems thinking for environmental literacy with a scoring guide, but how educators and students from formal and non-formal settings define the construct. The question guiding this investigation is:

> What do the commonalities between the evaluation of an assessment instrument by two groups of educators—formal and non-formal—reveal about using a scoring guide for systems thinking to measure a middle school student's environmental literacy?

The question works like and hours glass opening to the many possible understandings of environmental literacy, narrowing to choosing specific measures, and then opening again to the many possible expressions of environmental literacy. The findings will be used to make recommendation for educational policy and practices associated with environmental literacy and assessment.

Key Issues

The research question opens discussions related to assessment, environmental literacy, and equity. Gough's (2013) guiding principles point to four key issues associated with the development of assessments for environmental literacy:

1. Formal and non-formal educators need to know how to provide middle school students with knowledge, skills, and experiences that consider "contradicting data, multiple perspectives, and partial knowledge" (p.10).
2. Equity must be evident in the assessment instruments used to measure environmental literacy by calling attention to "racism and gender bias" (p.10).
3. Educators need to listen to "silenced voice" and be able to hear students' voices separately form their own (p.10).
4. System thinking tools, like the keys to a car, transfer responsibilities to students so they can make the decision-making processes behind their actions transparent.

Each of these issues influences what it means to be proficient in systems thinking for environmental literacy and how proficiency can be measured in ways that ensure equity.

The skills needed to address contradicting data. Gough (2013) explains that one of the principles for educators to put into practice is to "recognize that knowledge is partial, multiple, and contradictory" (p. 10). In 1987, Stevenson characterized environmental education and science education as adversaries. He framed the discourse by contrasting "the socially critical and political action goals of environmental education…with the uncritical role of schooling in maintaining the present social order" (p. 139). His statement distinguishing the "action goals of environmental education" from the "uncritical role of schooling" mirrored the polarization of citizens in Oregon, where those who proposed litigation against government agencies, businesses, and industries in order to protect natural resources were labeled as environmental activists. Fear of violence and litigation fueled polarization and misunderstanding.

One example of mistrust occurred when free copies of *An Inconvenient Truth*, a film with data concerning global warming, were not distributed to educators attending the National Science Teachers Association Conference in Salt Lake City, Utah, on December 8, 2006. The inference could be made that conference organizers were trying to discern whether an educator could be trusted to provide environmental education rather than "environmentalist" education. In fact, the Washington State School Board in January 2007 "voted to require approval by the principal and the superintendent for teachers to show the film…and that teachers must include the presentation of an approved opposing view" (Absolute Astronomy, 2013). Aware of polarization, non-formal educators have developed guidelines that acknowledge the importance of opposing views, conflicting data, mutual causation, and multiple data sets for enhancing conversations about the planet. In fact, Gough (2002) makes the following bold statement in her narrative suggesting educators rethink their relationships in the field of environmental education research as "a mutualistic one that meets the needs of both to continue to survive in a changing world" (p. 1203).

The need for citizen scientists. In 2007, Professor Jonathan Osborne from King's College London asserted that the future citizen requires "scientific literacy" (p. 174). He defines scientific literacy as "more than a knowledge of the basic concepts of science but rather a vision of *how* such knowledge relates to other events, *why* it is important, and *how* this particular view of the world came to be" (p. 174). Osborne (2007) describes a split in science education that distinguished "training the future scientists…in all the basic concepts of the discipline” and meeting “the needs of the future citizen” (p. 173–174). He describes the current practice of science educators “rather like introducing a child to jigsaws by giving him or her bits of a thousand piece puzzle and hoping that they have enough to get the whole” (p.174). He suggests that this deductive approach toward understanding the world carries some students deeper into specialized fields of study where the details of their findings are difficult for citizens to comprehend (p. 177). Osborne alludes to using an inductive approach to reasoning where the narratives shared by educators “give the message first and the details second” (p. 178). For example, rather than beginning a lesson on digestion with dissecting the alimentary canal of a three-legged amphibian genetically mutated by pesticides, an educator might begin by describing the transfer of energy from the sun to the algae in the pond near the wheat field that feeds the flies consumed by the frog.

Osborne (2007) identifies the unintended consequence of primarily using deductive reasoning in science as an overuse of the “precautionary principle” by citizens making decisions on issues that influence society (p. 177). The precautionary principle simply describes a tendency for one to choose an option with a low level of risk when one is excluded from understanding the science related to the issue. Osborne’s solution is for educators to use fewer “puzzle pieces” and to focus on “develop[ing] the ability to think critically about scientific evidence” (p. 179). Osborne (2007) identifies one key action educators can take to include all future citizens in decisions: give students the “opportunity to consider data which has no clear interpretation and to consider plural alternatives” (p. 179). To paraphrase Osborne’s earlier definition for science literacy, one skill demonstrated by a scientifically literate citizen is that they consider how data relates to events in their community, discuss why the data is so important, and decide how the data influences their view of the world.

Importance of personal experience in decision-making. While working to improve the implementation of environmental education in schools, Amy Cutter-Mackenzie, aassociate professor at Southern Cross University in Australia, and Richard Smith from Australia (2003), identified the concept of "ecological literacy" as the "missing paradigm in environmental education" (Cutter-McKenzie & Smith, 2003, p. 497). They explain that definitions of environmental literacy have changed since first used in 1968 and redefined in 1992 by Charles Roth, a former general science and biology teacher and the director involved in education at the Massachusetts Audubon Society (Paul F-Brandwein Institute, 2013). In his guiding publication, *Ecological Literacy: Education and the Transition to a Postmodern World,* Orr (1992) suggests, "Ecological literacy is becoming more difficult…because there is less opportunity for direct experience of it (p. 89). He says that educators need to consider the "process of education at all levels" (p. 90). Specifically, he asks educators to move away from education that "happens mostly as a monologue of human interest, desires and accomplishments" and move toward "education that occurs in part as a dialogue with a place and has the characteristics of good conversation" (p. 90). Similarly, Cutter-Mackenzie and Smith (2003) describe an ecologically literate individual as "knowing how the world works, and therein knowing how to preserve and maintain the environment" (p. 502).

Cutter-Mackenzie and Smith's (2003) research utilized interviews with 26 primary school teachers, and indicated that 62.8% approached environmental education when teaching science and 5.1% taught it separately (p. 511). They found that educators used "personal experience, creativity and imagination as a *means* [emphasis added] for understanding the world" (Cutter-Mackenzie & Smith, 2003, p. 500). As a result, they conclude, "the introduction of ecological literacy (eco-literacy) in educational policy may advance the goals for environmental education" (p. 520). Individuals redefine, reorganize, elaborate, and change their initial concepts through interaction with their environment, other individuals, or both. The learner "interprets" objects and phenomena and internalizes the interpretation in terms of the current experience encountered (Bybee et al., 2006, p. 11). To be successful in introducing ecological literacy requires not only personal commitment as noted by Cutter-Mackenzie and Smith (2003), but also a "system-wide commitment to environmental education…on the part of governments, education departments, pre-service education providers, primary schools and teachers themselves" (p. 520). Evidence for the kind of system-wide commitment recommended by Cutter-Mackenzie and Smith is found in the 2010 Oregon Environmental Literacy Plan (OELP).

Equity in Environmental Literacy

The initial premise of equity used for this research project was informed by Singleton and Linton's (2006) book *Courageous Conversations about Race*: "[Equity] is an operational principle that enables educators to provide whatever level of support is needed to whichever students require it" (p. 47). Singleton and Linton (2006) describe the role of language for meeting the needs of each individual, using the terms "White talk" and "Color commentary" (p. 23). They describe White talk as "verbal, impersonal, intellectual, [and] task-oriented," and Color commentary as "nonverbal, personal, emotional, [and] process oriented" (p. 123). Equity is especially important in the creation of a scoring guide for middle school students because individual students will need to understand what it is they need to show they can do. The nonverbal nature of the systems thinking tools and the emotions associated with personal experience are key components for meeting the needs of each student in experiencing proficiency with environmental literacy.

Leaders in communication. Outdoor Science Schools, directed by the Northwest Regional Educational Service District, are examples of non-formal education providers, with expertise in creating social environments that encourage inclusion and conversations across racial differences (Friends of Outdoor School, 2013, para. 1). During its inception, Outdoor Schools directed by the Multnomah County Educational Service District intentionally scheduled schools' visits so students from diverse socio-economic backgrounds would share the learning experience together. Pangrac and Christensen's (2012) report, *Fifth and Sixth Grade Student Participation in Outdoor School Programs in Oregon,* estimates that "52.8% of Oregon's current sixth-grade students attended an outdoor school program in the fifth or sixth grade" (p. 4). Outdoor school educators build common vocabulary with students from the moment students arrive by creating community and exploring natural systems. High school age counselors shepherd about eight students through each day's schedule. Students are regrouped into larger groups for field studies and homeroom time. They are divided individually as they enter the dining area so they sit at tables with different counselors and students from other cabins. The opportunity to meet and engage in conversation with unfamiliar people is intentional. A staff member reads a quote and assigns a question to discuss as food is brought back to the table. Singleton and Linton (2006) refer to this process as establishing "guidelines for exactly what participants speak about, how long they speak and listen, and who is and is not speaking and listening" (p. 131). High school age counselors serve as cultural brokers, teaching communication strategies and skills that were taught when they first became part of the Outdoor School community. Nowhere are conversations better structured than during campfires where instructions are provided about when to sing, stamp, clap, dance, listen, show appreciation, and walk quietly out to the field for a night hike. Since students understand the process for communication, they are free to contribute as they might wish.

The most prized tokens at Outdoor School are symbolic. Students add beads of various colors for participating in each field study, exemplary cabin behavior, picking up litter, scrub club, leading a song at campfire, or going to sleep on time. In the language of Outdoor School, roles are assigned through ritual and beads are the measure of proficiency. Students do not need to take tests or have their field study journals scored to earn beads. Instead, they sing songs about transpiration, otherwise known as tree sweat. The non-formal educators' communication style favors what Singleton and Linton (2006) refer to as "color commentary" (p. 123). Instructors and counselors use non-verbal techniques and teach directly from their own personal experiences, with enthusiasm and integrity, to the five or six students they are with, all the while walking the river together, holding animal skulls, drawing a map of the geographic features of Oregon in the sand, and unscrewing a soil profile with an augur. Using beads of various colors to measure the proficiency of students learning in a highly verbal, constantly moving, often wet culture is an excellent example of how to structure conversations for environmental literacy.

Interconnected through non-traditional communication. Singleton and Linton (2006) suggest encouraging "non-traditional ways of communicating" (p. 131). By using the tools of systems thinking such as connection circles, inference ladders and iceberg models, educators can meet each student's need "to be respected, validated and affirmed" (Singleton and Linton, 2006, 123). As students discover more about the community and natural systems of which they are a part, their needs for clean water, air and land will naturally appear. Assessment instruments for environmental literacy need to be evaluated to ensure "people of all races are valued, appreciated and heard" (Singleton and Linton, 2006, p. 125).

By using systems thinking, the OELP opens conversations about environmental literacy and natural resources by asking students to consider themselves as part of community and natural systems. Singleton and Linton (2006) describe two perspectives on the relationship between people and the environment. They say "White individualism (representative of prevailing U.S. culture) . . . understands the physical world as knowledge apart from its meaning for human life. . . . Color group collectivism (representative of many immigrant cultures) . . . understands the physical world in the context of meaning for human life" (p. 191). The latter assumption supports the fundamental concept of interconnections in systems thinking as the root for the construct of environmental literacy.

Interconnected through unexpected consequences. Rezendes (1999) describes how a scientific perspective that tends to categorize the world in order to make sense of it can have unexpected consequences:

> Nature is not made up of separate enclaves—predators in this corner, prey in that corner—but a totality, predator and prey living together. . . . We have labeled and separated the moose and the wolf, and in so doing, we have lost sight of their essential unity. We have also misunderstood ourselves, for the biggest separation we have imposed on the world is between ourselves and nature. . . . When we encounter nature, we also encounter ourselves. (p. 20)

Similar unintended consequences can results when educators use data from environmental literacy assessments instruments to measure the power of a program to communicate a message that changes behavior, or the power of a science course to credit students for explaining a concept. Educators in formal and non-formal settings have categorized and separated students' learning experiences into "separate enclaves" while "not noticing or hearing a thing" about how a student uses knowledge, skills, and experience to make decisions. By working together to evaluate assessment instruments to measure environmental literacy, it is much more likely educators will begin to see signs of "an individual's understanding, skills and motivation to make responsible decisions that consider his or her relationships to natural systems, communities and future generations" (ODE, 2010, p. 4).

Artificial boundaries. During Portland Public School Board's Town Hall on March 18, 2013, Oregon governor John Kitzhaber made recommendations for improving vocational and technical education by getting "rid of the artificial boundaries between Career and Technical Education (CTE) and Science Technology Engineering and Math (STEM)" (Portland Public Schools, 2013). Artificial boundaries may have been an unintended consequence of the Smith-Hughes Act of 1911, which provided federal funding for states to hire vocational teachers with skills in agriculture, forestry, canning, sewing, and industry to implement a new vision of progress in the United States. The new vision involved learning marketable skills for an industrialized economy. Students were not only trained to mechanically harvest peas, can food, and sew their own clothes, but to serve the community as nurses, teachers, and secretaries (E. S. Duncan, personal communication, July 5, 2009). If a school received funding through the Smith-Hughes Act, students spent half their day in general education classes and half their day in vocational classes. Du Bois spoke for families enslaved to work in agriculture and grow the U.S. economy while "shut out from their world by a vast veil" (Du Bois, 1903, p. 128). He writes:

> In those somber forests of his striving, his own soul rose before him, and he saw himself,—darkly as through a veil; and yet he saw in himself some faint revelation of his power, of his mission. He began to have a dim feeling that, to attain his place in the world, he must be himself, and not another. (Du Bois, 1903, p. 128)

Du Bois articulated how people had been denied education because they were seen "to be servants and nothing more" (Dubois, 1903, p. 130). Du Bois' writing offered this answer to the question of the purpose of education:

> Work, culture, liberty—all these we need, not singly but together. . . . Not in opposition to or contempt for other races, but rather to the great ideals of the American Republic. . . . Give each to each those characteristics both so sadly lack. (Du Bois, 1903, p. 130)

Booker T. Washington supported the Smith-Hughes Act, suggesting that it provided "opportunity" for the whole population of the southern United States, including former African slaves used in agriculture, by ushering in a "new era of industrial progress" (Washington, 1901, p. 133). He famously said:

> Cast down your bucket where you are—cast it down by making friends in every manly way of the people of all races by whom we are surrounded. . . . No race can prosper till it learns that there is just as much dignity in tilling a field as in writing

> a poem. . . . Cast down your bucket to . . . the education of head, hand, heart. (Washington, 1901, p. 133)

The phrase "head, hand and heart" is still echoed today in Oregon's 4-H program, which has recently diversified its historically vocational curriculum to include STEM education opportunities such as robotics. The four words in 4H are: head, hands, heart, and health. Oregon's former chief education officer, Rudy Crew, unpacked the idea of "artificial boundaries" mentioned by Kitzhaber. Crew indicated that "we have to start laying out a new set of lanes in schools . . . [so] everybody has a way to see themselves as being both gainfully employed and gainfully employable, and learned, and smart and capable and confident" (Portland Public Schools, 2013). He affirmed the need to "embrace vocational training, CTE opportunities, internships, externships, [and] community-based learning" (Portland Public Schools, 2013). The Environmental Literacy Plan has the potential to stitch together vocational education, environmental literacy, and STEM by taking a lead in providing community-based learning opportunities with instruction linked to valid and reliable assessment instruments.

Listening to silenced voices. The significance of generating a scoring guide for systems thinking that can be used to measure student proficiency in environmental literacy is that it gives students a voice. Systems thinking tools meet a critical need in science education by teaching students the skills they need to set goals and leverage actions unique to the community and natural systems of which they are a part. The most accessible and familiar natural system to the students provide a context for learning, and it becomes part of a student's répertoire to understand scientific principles and patterns well enough to make decisions about their interrelationships. Systems thinking teaches students a process that Oregon's existing scoring guides for science inquiry and engineering design do not—a process that not only illuminates leverage points and interactions, but generates an experience of what Pipher (2013) labels "hope." She wrote, "Hope is not about outcome, but about process" (p. 210). Lieberman and Hoody (1998) claimed that "using the environment as an integrating context for learning holds great promise," but the burden remains on the educator to "build bridges between theory and reality, school and communities, children and their futures" (p. 11). Lieberman and Hoody's (1998) research was influential in the field of environmental education because it legitimized students and teachers working with non-formal educators outside the classroom using the "Environment as an Integrating Context (EIC)" (p.2). Their use of academic measures to support EIC helped environmental educators weather the standards-based test-heavy reform that came with the 2002 reauthorization of the Elementary and Secondary Education Act known as No Child Left Behind. Process encourages students to define context in conversation with an educator based on their current understanding of natural systems and communities living in their particular corner of the planet.

In Einstein's 1941 radio address to the British Academy for the Advancement of Science, he asked, "What hopes and fears does the scientific method imply for mankind?" (Springer, 2013). According to Springer's podcast (2013), Einstein's rhetorical response was as follows:

> Whatever this tool in the hand of man will produce depends entirely on the nature of the goals alive in this mankind. Once these goals exist, the scientific method furnishes means to realize them. Yet, it cannot furnish the very goals. The scientific method itself would not have led anywhere—it would not even have been born without a passionate striving for clear understanding. . . . If we desire

> sincerely and passionately for the safety, the welfare, and the free development of the talents of all men, we shall not be in want of the means to approach such a state.

The habits of systems thinking, gaining knowledge and skills in science, sharing experiences, and designing sustainable technology can be the means by which people work toward living in balance with one another and natural systems.

Had Einstein lived today, he most certainly would have appreciated a student with the “means” to use scientific principles to transform peoples’ interactions with the planet. One student collected data on indoor air pollutants, which has led to the development of an online tool that doctors can use to reduce patients’ symptoms of disease due to chronic exposure to gases and particulates. She had the opportunity to teach President Obama, Environmental Protection Agency staff, and the public about the scope of indoor air pollution problems. She shared her proposed solution by winning the online Google Science Fair and giving a TED talk titled *Award-Winning Teenage Science in Action*. The student’s passion for the health of one’s family provided immeasurable motivation. The student’s work is an example of having the means to identify and leverage possibilities. Christian Long indicates that students “change the world in real time and get us to invest” in their plans for the future (C. Long, personal communication, June 25, 2013). So, environmental literacy assessment instruments need to recognize “the means” individuals have to leverage actions in various community and natural systems so we can invest in their plans now.

Systems Thinking Tools Unpack the Decision-Making Process

Meadows (2008), a professor in the environmental studies program at Dartmouth College, worked on one of the first computer modeling programs for “population, economic growth and a finite planet" (p. 11). She deconstructs a system as "an interconnected set of elements that is coherently organized in a way to achieve something" (p. 11). For example, a tree's system is made up of roots, trunk, branches, fungi, leaves, sunlight, and water organized in a way that transforms sunlight into the starch that is used as food for the cells of the tree. One of the first skills of systems thinking is to describe the boundaries of the system under consideration before finding the interconnections. Meadows (2008) defines interconnections "as the relationships that hold things together" (p. 13). She indicates that in the tree the interconnections would be the chemical reactions that form the water, oxygen, and glucose (p. 13). Interconnections represent the "actual physical flow" through a system (p. 14). With an understanding of how to use the tools of systems thinking, students and educators are equipped to critically analyze multiple systems acting together in the places they live. Assessing the environmental learning strands, specifically systems thinking, using a scoring guide could function to maintain the integrity of the goals of the OELP. Assessing students' decision-making skills in the context of the problems and possibilities they find in the community and natural systems, of which they are a part, unites science and environmental literacy.

Missing scoring guide. The OELP provides a curricular framework that formal and non-formal educators can use to organize how they invest in students' proposed solutions. According to Cloud (2002), systems thinking gives students a common "frame of reference" (para. 5). So, students from diverse ecoregions across Oregon, can use it as a common language through which to consider decisions. The habits of a systems thinker were selected by the Oregon Environmental Literacy Task Force to generate awareness and understanding so one can decide how to act to optimize community, natural systems, and the future. The Sustainable Oregon Schools Initiative also provided a forum for formal educators to create sustainable community, and supported natural systems in schools though architecture and design. The OELP learning strand for systems thinking teaches students how to use Connection Circles, Causal Loops, Stock/flow maps, computer modeling and simulations, a Ladder of Inference, and the Iceberg Model to identify a leverage point, goal, problem, or possibility. The systems thinking tools invite students and educators to learn a common language for connecting science concepts with economic, social, and political concepts (see Appendix A-F for templates of systems thinking tools).

A case study on decision-making. Scoring guides are examples of frames that teach students how to communicate using a particular process. Note that frames give each student a sense of agency. Each can be valid or true for the individual describing it. Rose and Barton (2012) affirm the use of frames in science education:

> The frames themselves provide a context in which teachers can support talk about why an issue matters to students, and the implications this has for how students evaluate socioscientific issues. At the same time, frames make the science in socioscientific issues more visible, and open to examination from multiple perspectives. (p. 563)

Rose and Barton (2012) completed a qualitative case study using social practice theory to investigate the role science played in the students' decisions regarding the proposed building of a biomass-coal-fired energy plant in their community as a replacement for the previous coal-fired energy plant. They interviewed "youth from non-dominant backgrounds—youth whose families struggle to pay their electricity bills, youth who have seen the impact first hand of the lack of opportunities for work" (p. 564). They made "sense of how these experiences shape when, how or why they might leverage their scientific understandings to make good and justifiable socioscientific decisions" (p. 564). Not only did their findings reiterate the responsibility for educators to provide multiple perspectives for solutions as outlined in the original Tbilisi Declaration, but they also provided evidence that suggests "the range of knowledge and experiences [middle school age] youth bring with them are powerful and legitimate resources for making sense of socioscientific issues" (Rose & Barton, 2012, p. 565). The students learned about the science of wind and solar power, but deemed these alternatives as inappropriate for their region. Rose and Barton (2012) argue that science knowledge, especially related to greenhouse gases and particulates that cause asthma, was only one of many frames students used to make their decision to choose a biomass and coal-fired energy plant to be built where they lived. Providing educators and students from diverse community and natural systems with equal access to framing the context of problems lends credibility to using the systems thinking skills as a strategy for decision-making.

Existing scoring guides. Oregon Department of Education has adopted two scoring guides for educators to use when assessing middle school students' proficiency in science inquiry and engineering design. A student who does research on indoor air pollution would be scored for the quality of the question, experimental design, data collection, and analysis using Oregon's science inquiry scoring guide. Proficiency in engineering design involves defining a problem, designing a solution, testing, and evaluating the design use criteria. The practice of using scoring guides to assess students' work samples in science is not new to teachers and students. The middle school students and educators who chose to participate in this study created and tested a scoring guide for systems thinking to assess a middle school student's ability to understand the implications of one's choices in the context of community and natural systems.

A formal educator in action. As an experienced graduate-level instructor of environmental ethics, Golley (1998) taught environmental literacy to all students, both non-science and science majors, using environmental science concepts. He writes,

> When my students and I go into the field, I tell them that our first task is to learn to read the landscape. I show them that the landscape is a text that informs us about its capacity to produce and support life, its history, and what organisms are likely to be present. But for me, at least, environmental literacy connotes more than knowing the names of the organisms and understanding geomorphology. I also emphasize feeling the landscape through the senses. This feeling of place distinguishes each site and makes a place special and memorable. Environmental literacy begins with experience of the environment. (p. ix)

Golley (1998) addresses the relationship between science and environmental literacy. "The scientist searches for patterns of relationships between natural objects and processes. . . . We are searching for the relationships among the patterns of nature. . . . Scientists know that their observations will be challenged" (p. xii). His unique approach encourages conversations between people by inviting them to consider their feelings while discussing science concepts, environmentalism (action), and environmental ethics. For example, after explaining the concept of mutualism, he suggests implications for social interactions in terms of competition and cooperation between humans. Individuals with an understanding of interrelationships between trees and mycorrhizal fungal hyphae, which provide "a potential link for the flow of chemical information among individuals in a forest," may gain a deeper appreciation for understanding variations in human communication as well (Golley, 1998, p. 182).

Summary

Without a system-wide commitment led by educators working together in a public process for implementing the OELP, the OELP is susceptible to "Drifting Goals" (Waters Foundation, 2010, p. 13). Even with a strong state-wide plan that champions the work of formal and non-formal educators, national legislation and competition for funding could continue to perpetuate the lack of parity between educators for their common work with middle school students in and out of school settings. By crossing boundaries posed by guidelines and standards, as well as national, state and local funding sources, educators and students have demonstrated they can "understand and apply systems thinking concepts and tools" (p. 16). The systems thinking tools can be used as a visual language to support educators and students with describing the interconnections, patterns, and principles that inform decisions. The importance of learning this language is that it uses diagrams to show how the parts of systems work together to function as a whole:

- Behavior-Over-Time graphs
- Connection circles
- Causal loops
- Stock flow diagrams
- Models (p. 19)

The jargon of systems thinking is not intended to silence voices, and the tools communicate best in the hands, hearts, and heads of those using them to explain their decision-making process. Systems thinking tools can help individuals organize the hundreds of "puzzle pieces" Osborne (2007) suggest are taught for science literacy, but they engage both deductive and inductive thought. They support Gough's (2013) guiding principles because they first, recognize that "knowledge is partial, multiple and contradictory," second, encourage students to make the unseen seen and the unheard

heard, and, most importantly, teach us to follow our mental models deep into the icebergs of their origins in order to “develop understandings of the stories of which we are a part” (p. 10).

CHAPTER 2

Literature Review

In order to deepen our understanding of how to assess the environmental literacy of middle school students based on our corresponding goals for systems thinking, environmental education guidelines, and science standards as described in Chapter 1, I plan to provide a synthesis of the relevant background literature. A number of educational leaders with experience in environmental literacy have forged a path for environmental literacy assessment instruments. Osborne suggests that the research community needs to give more energy to assessment as part of their practice**.**

> Practice is a combination of the triumvirate of curriculum, pedagogy *and* assessment. So far, the research community has displayed far less interest in this [assessment] component than the other two. However, in a context of increasing accountability, it is to assessment that teachers look for the intended curriculum, not the curriculum itself. (Osborne, 2007, p. 182)

Hollweg et al. (2011) call for research documenting students' confidence in decision-making, and measuring their progress over time. They suggest, "There is a clear need for national and international assessment data to better understand the status of environmental literacy, with data broken down by the components and by age/developmental levels" (p. 4). On the national level, Elder (2003) foreshadowed the need to develop assessments for environmental literacy using a common set of standards. He writes, "If environmental literacy is to gain a more substantial foothold within the nation's priorities, it is critical to establish a fundamental baseline through a thorough set of national goals, benchmarks and standards" (p. 93). By 2013, educators in Oregon succeeded in defining learning strands for environmental literacy, which can be used for assessment. Naturally, the next task involves using the strands of the OELP to measure each student's level of proficiency with appropriate instruments.

Chapter 2 has two major components. First, educational theories for assessing science literacy and environmental literacy are discussed in order to understand the reasoning and methods used to create assessments. Knowing that educators design assessments to measure specific aspects of environmental literacy as it is understood by a particular group of people in a particular place at a particular period in time, a number of existing assessment instruments that resonated with the environmental literacy strands in the OELP were evaluated. The second part of the chapter takes a closer look at eight existing assessment instruments, and compares them through the lenses of validity, reliability, and equity.

Theoretical Framework

The educational theory appropriate for developing an environmental literacy assessment instrument was constructivist because it "empowers" educators (Guba & Lincoln, 1998, p. 210). For Guba and Lincoln (1998) constructivism entailed "*understanding and reconstruction* of the constructions that people (including the inquirer) initially hold, [aiming] toward consensus, but still open to new interpretation as information and sophistication improve" (p. 211). A benefit of using a constructivist approach was that implications for policy and practice were based on validity and reliability. In this case, validity and reliability were "derived from community consensus regarding what is 'real,' and what has meaning, especially for future action" (Guba & Lincoln, 2005, p. 197). Constructivist theory was adopted so educators could speak with equal conviction to a board of directors, principal, school board, or funding agency regarding middle school students proficiency with systems thinking for environmental literacy as outlined in the 2010 Oregon Environmental Literacy Plan (OELP).

The constructivist approach continued as a public conversation that began when the No Child Left Inside Task Force was established by the Oregon legislature to draft an Environmental Literacy Plan in 2009. By developing mutual understanding of one another's constructs, the credibility of an environmental literacy assessment instrument can be improved. Maddock (1999), citing Ellsworth, noted that instruction was moving away from presenting all perspectives so a student can choose one over another, towards "the kind of talk that reflects the partial, interested and potentially oppressive character of all knowledge and which works at reshaping alliances in which 'difference can thrive'" (p. 49). Guba and Lincoln (1998) described the importance of measuring credibility in terms of "authenticity" (p. 196) as well as whether the findings serve as a "catalyst for action," (p. 212) and continued efforts to move towards consensus. Educators and students were assumed to "seek understanding of the world in which they work…through the meanings of their *experiences*" (Creswell, 2007, p. 20). The meanings assigned to experiences were used to describe the "processes of interactions between individuals" that were representative of "cultural norms that operate on individual's lives" (Creswell, 2007, p. 21). The rationale for adopting a constructivist approach was to measure the level of consensus and gain a deeper understanding of the constructs of systems thinking, proficiency, equity, and environmental literacy to inform the assessment practices of educators in Oregon. The Delphi technique was selected to create a scoring guide for systems thinking to "become more aware of the content and meaning of competing constructions" (Guba & Lincoln, 1998, p. 211). In the autoethnography, Vygotsky's social constructivist theory describing the Zone of Proximal Development (ZPD) was employed to explain how children learn. The ZPD represents “the difference between what a child can do by themselves, and what they can achieve with guidance and encouragement from a skilled partner” (McLeod, 2014, para. 25). Gopnik’s (2010) description of how children develop “counterfactual thinking” or “the ability to imagine a different world and act” was used to understand the developmental appropriateness of asking middle school students to explain the reasoning behind their action using systems thinking tools.

Research in the area of assessment for environmental literacy increased since Osborne (2007) identified increasing accountability for educators and the development of curriculum based assessments as two reasons to change educational practice. In an effort to be accountable for successful implementation, the OELP identifies possible instruments for measuring environmental literacy: the OAKS, student work samples, and an adaption of the 2008 NELA, which was used by the North American Association of Environmental Education (NAAEE) to established a baseline measure of middle school students’ environmental literacy. Assuming the purpose of educators in Oregon is to show evidence of each student’s proficiency in environmental literacy as defined in each of the five OELP learning strands, two other types of assessment instruments appear to be valid, reliable and equitable: work samples that are scored by teachers, and instruments, like MEERA designed to measure the specific of individual programs. To avoid confounding uncertainties, and provide honest information to the public about students’ proficiency with environmental literacy, these lessons from previous researchers can be applied:

- Consider a student's ability to show their reasoning and evidence in evaluating students' decision-making skills.
- Verify the understanding of the construct of environmental literacy among those using the scoring guide.
- Recognize that the understanding of scientific principles and personal experience influence the development of environmental literacy over time.

In addition, the level of generalization that needs to be made from the data from a particular assessment instrument needs to be considered. The research design of this study limits the generalization of findings to a small, representative group of middle school students, and educators with skills in the application of systems thinking, science concepts associated with natural phenomena, and experience interacting with community and natural systems from the diverse eco-regions of Oregon**.**

Review of the Methodological Literature

Construct validity. Based on Trochim's work (2006), two primary threats to construct validity needed to be considered: (1) a "preoperational explanation of the construct" of systems thinking for environmental literacy before it was measured, and (2) the "mono-operation bias" of using a single place, time and group to measure the construct (para. 4). Trochim (2006) explained "construct validity as an overarching quality with all the other measurement validity labels underneath it" (para. 2). He recommended Pattern Matching Theory for matching a "theoretical pattern" with "observed pattern" (para. 2). In short, he suggested a researcher investigate whether the "ideas" or "hunches" educators have about how a student demonstrates a particular environmental literacy strand match what they can "observe" and "measure" using a particular assessment instrument (para. 2). He used the example of Cronbach and Meehl's (1955) nomological network, which was used as evidence for research using psychological tests. He described a nomological network as “a representation of concepts (constructs) of interest in a study . . . their observable manifestations and the interrelationships” (para. 1). Mono-operational bias can be addressed by using the same scoring process with educators who provided middle school students with knowledge and skills for systems thinking in non-formal or formal settings, and providing the same scoring instructions to each person who participated in the process of field testing the scoring guide.

Face validity. Fortunately, the OELP Task Force completed foundational work to increase face validity for this study by clarifying the meaning of environmental literacy. Face validity increased as a more diverse group shared a common understanding of a particular construct. Face validity was at the forefront in drafting the OELP because H.B. 2544 (2009) required a Task Force made of representatives from agencies with different, yet related missions to work with the staff of the Oregon Department of Education to define: “The meanings of key terms required for developing the plan, including the meanings of the terms “environmental literacy,” “climate change” and “healthy lifestyles” (H.B. 2544, Sec. 2, 2009). By bringing together experts from agencies with diverse missions, environmental literacy was defined as: “An individual’s understanding, skills and motivation to make responsible decisions that consider his or her relationships to natural systems, communities and future generations” (Oregon Department of Education, 2010, p. 4). This was an example of the legislative process using Pattern Matching Theory, which was one of the best practices in social science research suggested by Trochim (2006). He also advised improving face validity by using “a carefully selected sample of experts” (para. 4). For the purposes of assessing environmental literacy in Oregon, the instrument needs strong construct and face validity to clearly articulate what it means for a middle school student to be proficient with systems thinking as it was outline in the 2010 Oregon Environmental Literacy Plan.

Threats to internal validity. A mixed methods design was selected to address primary threats to internal validity, which included the selection criteria for the educators and students on the forum, repeated measuring with the scoring guide, and investigator bias. The selection of the members of the forum used an opportunistic sampling technique, which was necessary in order for the individuals to represent a particular perspective at a particular point in time. The Delphi technique ensured that each person remained anonymous. Because the researcher and forum members use electronic communication, they could remain anonymous, and non-verbal communication did not influence their conversation. Each member had an equal voice, and including all the response from the surveys maintained internal validity. In addition, the researcher disclosed her bias by drafting a scoring guide before the Delphi began (see Appendix D).

External validity. The primary threat to external validity is that the educators involved in testing the scoring guide differ from the population. The snowball sampling method was used so if educators found testing the scoring guide useful, they might suggest participation to a colleague. The criteria for participating in testing the scoring guide was limited to those who work with middle school students as formal or non-formal educators irrespective of years of experience, or the amount of self-study in systems thinking, or practice with assessing students' work. The four steps of the scoring guide for systems thinking are simply a tool for a conversation between a student demonstrating one's skill, and an educator that recognizes that skill irrespective of whether learning occurs in the shrub-steppe of Oregon's high plateau, urban school gardens or tide pools at the coast.

Evaluation of Existing Assessment Instruments

A number of existing instruments, and the methods used to evaluate them, were reviewed prior to the research for this dissertation. They include: the Oregon Assessment of Knowledge and Skills (OAKS), work sample scoring guides adopted by the Oregon Department of Education, Programme for International Student Assessment (PISA), Middle Years Programme of the International Baccalaureate, National Environmental Literacy Assessment (NELA), My Environmental Education Evaluation Resource Assistant (MEERA), Assessments for Environmental Science Literacy-Michigan State University, and EUGENE The Next Generation Science Standards (NGSS) were reviewed as guiding document for the development of future assessment instruments. Raw data was added to a Pugh chart[1] shown in Appendix H scoring each of the existing instruments on three evaluation criteria: (1) construct validity, (2) reliability, and (3) equity. The OAKS was selected as the datum, or reference for comparison, since it has been used as a tool to measure Oregon middle school students' understanding of a number of science concepts and skills that resonate with the environmental literacy strands. The comparison served to identify confounding uncertainties associated with three aspects of the proposed research: constructing validity for environmental literacy assessment instruments, verifying the reliability of educators from formal and non-formal settings, and improving equity in assessment for middle school students across Oregon.

---

[1] A Pugh chart is a tool used by engineering designers to help select an appropriate design from multiple possibilities. It is used to decide which design to test before investing in building a prototype.

Oregon Assessment of Knowledge and Skills (OAKS). The OAKS for science is administered via computer to students in grades 5, 8, and 10. OAKS test specifications and blueprints are made available to the public through the Oregon Department of Education (ODE). The test was built from the 2009 science content standards. It is a multiple-choice test that includes interactive graphic manipulation of images for students to show understanding of science concepts. Accommodations are available for students who would like questions in both English and Spanish. Teachers can read the questions aloud to students as they take the test. Students may take the test twice in a school year. A score of 247 exceeds the science standard and 235 meets the standard. The items, or questions, on the test assess science content, vocabulary, and conceptual understanding. Scientific inquiry and engineering design questions measure students' skills in research and product development. The 2008–2009 Science Content and Assessment panel included 25 formal educators and 11 non-formal educators representing universities, community colleges, educational service districts, the Oregon Business Council, Northwest Regional Educational Labs, Portland State University Center for Science Education and the Oregon Forest Resources Institute (ODE, 2008). Formal educators use the same standards from which the OAKS is developed to create lessons and assessments for their students. Non-formal educators align the standards to their programs so formal educators recognize the skills and knowledge that students will experience. Table 1 identifies the Oregon science standards that resonate with the five strands of the OELP and are eligible for assessment on the OAKS. The test specifications require that "test items must be appropriate for students in terms of grade-level, difficulty, cognitive complexity, reading level, interests and experience; be free of age, gender, ethnic, religious, socioeconomic, or disability stereotypes or bias, and provide clear and complete instructions to students" (ODE, 2011, p. 53).

Table 1

Oregon Science Standards Adopted by Oregon Department of Education

Oregon Science Standard	Definition of Proficiency
6.2L.2	Explain how individual organisms and populations in an ecosystem interact and how changes in populations are related to resources.
6.2E.1	Explain the water cycle and the relationship to landforms and weather.
6.3S.1	Based on observations and science principles, propose questions or hypotheses that can be examined through scientific investigation. Design and conduct an investigation that uses appropriate tools and techniques to collect relevant data.
6.3S.2	Organize and display relevant data, construct an evidence- based explanation of the results of an investigation, and communicate the conclusions.
6.4D.1	Define a problem that addresses a need and identify science principles that may be related to possible solutions.
6.4D.2	Design, construct, and test a possible solution to a defined problem using appropriate tools and materials. Evaluate proposed engineering design solutions to the defined problem.
6.4D.3	Describe examples of how engineers have created inventions that address human needs and aspirations.
7.2L.2	Explain the processes by which plants and animals obtain energy and materials for growth and metabolism.
7.2E.1	Describe and evaluate the environmental and societal effects of obtaining, using, and managing waste of renewable and non-renewable resources.
7.2E.2	Describe the composition of Earth's atmosphere, how it has changed over time, and implications for the future.
7.2E.3	Evaluate natural processes and human activities that affect global environmental change and suggest and evaluate possible solutions to problems.
8.2E.4	Analyze evidence for geologic, climatic, environmental and life form changes over time.

Table 1.
Note: The 2009 Oregon learning standards for science were found by searching from *REAL: Standards by Design* available at http://www.ode.state.or.us/teachlearn/real/standards/sbd.aspx. They have since been replaced by the Next Generation Science Standards adopted in 2014. See Appendix G for correlations between the OELP and NGSS.

ODE uses a 7-year cycle to adopt new science content and generate assessments. Educators from Oregon helped author the Next Generation Science Standards (NGSS) completed in April 2013. In March 2014, the Oregon School Board adopted the NGSS as the framework to guide science instruction and assessment.

Next Generation Science Standards. The NGSS were based on the *Framework for K-12 Science Education* developed by the National Research Council (NGSS Lead States, 2013). The NGSS proposes standards that will assess students' ability to communicate their reasoning using scientific principles and crosscutting concepts. The term "systems" is used to describe the skills associated with interactions and modeling as one of the crosscutting concepts. The NGSS will guide educators' instruction and assessment in Oregon until they are reviewed as part of the 7-year cycle.

Art Paz, Jr., professor of architecture at the University of Oregon and former member of the Oregon State Board of Education questioned the completeness of the NGSS for curriculum design. He raised crucial questions regarding cultural framework, aesthetic sensibilities, and ecological intelligence. He also asked about the role of time and deeper thinking (Art Paz, Jr., personal communication, June 25, 2013). Ault (2015) also raised concerns about the nature of science as it is described in NGSS. He wrote:

> Often the sciences that address complex systems with many interacting variables bear fruits that are of high social value . . . The approaches taken by field science in the study of animal behavior suggest how to teach even the children in primary grades about complex phenomena. (p. 178)

The NGSS identifies knowledge, skills, and experiences that lead to environmental literacy, and the OELP deepens the ability of students and educators to explain the reasoning underneath their decisions as well as refine their actions.

PISA: Programme for International Student Assessment. By 1998, the assessment of science literacy had become well defined and was being assessed in 32 countries through the Organisation for Economic Co-operation and Development (OECD)/ PISA project (Harlen, 2001, p. 49). Harlen (2001) cites PISA's definition of science literacy as, "The capacity to use scientific knowledge to identify questions and to draw evidenced-based conclusions in order to understand and help make decisions about the natural world and the changes made to it through human activity" (p. 52). One key aspect of PISA is that science literacy is a "progression," which affirms students are naturally capable of observing and developing explanations for the natural phenomena they experience (Harlen, 2001, p. 52).

Similarly, research in the field of environmental literacy is an outgrowth of studies in environmental education. Educators identify its origins with the Tbilisi Act and Agenda 21 of the United Nations, published in 1977. The definition of environmental literacy developed by the OELP committee resonates with PISA's international agenda for science literacy: "An individual's understanding, skills and motivation to make responsible decisions that consider his or her relationships to natural systems, communities and future generations" (ODE, 2010, p. 4) Note that the definitions for environmental literacy and science literacy by the OELP committee and the PISA committee support assessments at different hierarchical levels ranging from microhabitats to global systems.

PISA (2009) published key findings in the document *Green at 15*, which showed the value of assessments that ask students to think in "multiple levels." The PISA assessment tasks are scored on an A to D competency scale. In 2006, four hundred thousand students from 57 countries were tested (OECD, 2009, p. 9). Results indicated an average of 84% of students having some proficiency at level D (OECD, 2009, p. 9). In addition to testing knowledge of environmental science, the PISA also surveys students on the following:

- Familiarity with environmental issues
- Sense of responsibility for environmental issues
- Optimism regarding environmental issues
- Awareness of complex environmental issues. (OECD, 2009, p. 8)

PISA data can be widely generalized. It can be used to make various correlations between demographics and survey questions as well as knowledge and skills. For example, "Students who report the greatest familiarity with complex environmental phenomena tend also to have high levels of proficiency. . . . Students with more disadvantages socio-economic status are no less likely to be committed to tackling environmental issues" (OECD, 2009, p. 10).

Another advantage of the PISA as a measure for environmental literacy is the inclusion of curricular "measures of school contexts, instruction and activities that promote learning about environmental issues, and parental perceptions of environmental issues" (OECD, 2009, p. 18). Of particular importance for the use of the PISA in Oregon assessing students with the indicator calculated for outdoor experiences. The 2009 report indicates:

> The most commonly used outside classroom learning activity for teaching about environmental science is outdoor education: almost eight out of ten students in

> OECD countries on average attend schools that use this approach. In Greece, Poland and the Slovak Republic, and in partner countries Azerbaijan, Colombia, Kyrgyzstan, Lithuania, Slovenia and Thailand, 90% or more of students attend schools that use outdoor education. (OECD, 2009, p. 72)

When the data is broken down by country, the scores of students from the United States fall slightly below the average of other OECD countries using outdoor education (OECD, 2009, p. 72). If PISA can disaggregate data by state, Oregon could use the PISA as an instrument for assessing environmental literacy.

Although the OAKS could be used to assess similar knowledge and skills using multiple-choice questions, it currently does not have PISA's robust survey component. OAKS also uses multiple choice questions for ease of scoring rather than providing opportunities for students to write their responses in their own words. The PISA could potentially be used to measure the environmental literacy of Oregon's middle school students.

Middle Years Programme (MYP). "These are exciting times," says Malcolm Nicolson, head of the redevelopment of the MYP, who is currently piloting program revisions with 200 schools in 40 countries (IBO, 2012). In 2014, MYP rolled out an external interdisciplinary E-assessment that will be optional for participating schools. Nicolson indicates that the E-assessment will include interdisciplinary concepts, which will allow students to demonstrate proficiency in unfamiliar contexts. Even though it is optional, Nicolson suggests that the instrument may be recognized and used by governments and universities to make decisions regarding a student's education:

> It provides a framework of learning, which encourages students to become creative, critical and reflective thinkers. The MYP emphasizes intellectual challenge, encouraging students to make connections between their studies in traditional subjects and to the real world. It fosters the development of skills for

> communication, intercultural understanding and global engagement, qualities that are essential for life in the 21st century. (IB Middle Years Programme, 2012, para. 1)

According to Nicolson, the MYP includes a focus on environment, which used to be one of the "areas of interaction" and has evolved into one of the "global contexts."

MYP is aware of the demands placed on schools to meet state and district requirements in addition to MYP criteria. MYP provides a curricular framework developing a "highly skilled global community." Of note are two of its five key objectives:

- Enables students to understand and manage the complexities of our world, and provides them with the skills and attitudes they need in order to take responsible action for the future
- Ensures breadth and depth of knowledge and understanding through the study of eight subject areas (IB Middle Years Programme, 2012, para. 3)

Nicolson (2013) explains that internal assessments are being redesigned so that teachers, students, and parents can more easily understand scores across subject areas. He says that all subject area scoring guides will have criterion-based scoring guides with four levels. In addition, the criteria will use similar descriptors for assessing students' work across subject areas, and each subject area will have only four different scoring guides. Educators from MYP member schools are permitted to include aspects of the OELP in units. MYP E-assessments and scoring guides could be used to measure environmental literacy. Its implementation process includes several steps that reflect similar aspects of implementing the OELP; however, a school must apply to become an MYP school, pay a fee, provide evidence of implementation, and successfully meet all requirements, including professional development for educators, direct interviews, and inspections.

The rigorous process of earning MYP certification assures face validity. According to Trochim (2006), face validity describes how well a construct, such as "managing the complexities of our world," is measured by a particular assessments instrument. He suggests asking experts to evaluate the assessment instrument. Educators are required to read technical guides such as *MYP: Principles to Practice*, and participate in ongoing professional development related to interdisciplinary curriculum development. Reliability is high because significant amounts of time are devoted to teaching the scoring guides that are used to assess students' work, and calibration of scoring between educators is used to improve curriculum, instructions, and assessment. Construct validity is improved because students are provided with multiple opportunities to be assessed. According to Trochim (2006), construct validity allows internal generalization between how a construct is actually measured, or operationally defined, and how the construct is understood, or its theoretical definition. The interdisciplinary E-assessment is clearly intended to generate evidence for convergent validity by showing the ability of students from 40 different countries to successfully complete the instrument. Trochim (2006) explains that convergent validity indicates a shared theoretical understanding of a concept across similar programs attempting to measure the same construct. With this kind of data, inferences can be made about the similar quality of the curriculum and instruction in member schools from different countries.

National Environmental Literacy Assessment (NELA). The National Environmental Literacy Assessment, which used an instrument called the *Middle School Environmental Literacy Survey* (MSELS), was completed in 2011 (McBeth et al., 2011, p. ix). The instrument was created for Phase I, which determined that the following "domains" are "critical to environmental literacy: knowledge, affect, cognitive skills, and behavior" (McBeth et al., 2011, p. xi). Phase II was a comparative research study that filled a gap in existing environmental research, which lacked content validity, only evaluated a few components of environmental literacy, and limited the ability to generalize only to individual programs (McBeth et al, 2011, p. 9). Interestingly, the researchers are self-depreciating in referring to "internal program reports" and "reports to funding bodies" as "fugitive literature" (p. 8). In fact, these kinds of documents are highly valued by stakeholders with resources to support environmental literacy. Their efforts can be documented with the use of instruments like the MSELS, which have high content validity. Content validity can be used as a kind of "checklist" to determine if one program shares the same characteristics of another program and has what it takes to be categorized as one that teaches the strands of the OELP (Trochim, 2006, para. 4). High levels of content validity depend on clear, detailed criteria and definitions of the phenomena that are measured. In Phase II, schools were nominated to participate if they have two classes participating in environmental education activities for two years (McBeth et al., 2011, p. 21). Researchers used web-based instruction to train 31 data collectors in how to survey 64 schools from the pool of 110 that responded to an invitation to participate in the study (McBeth et al., 2011, p. 22). The next instrument discussed focuses on the work of non-formal educators.

MEERA: Measuring the Evaluation Competency of Non-formal Educators. A recently published study by Zint, Dowd, and Covitt (2011) investigated the "evaluation competencies" of environmental educators who spent 10 to 100 hours of self study using an online evaluation instrument called *MEERA: My Environmental Education Evaluation*

Resource Assistant (p. 476). In its development, MEERA was designed for use by educators in colleges and universities, K–12, and non-formal institutions. One of those who participated in the study suggested more time and intention be given to working with formal educators:

> If I had time [in the future], I would try a participatory evaluation. Having gone through the process once now, I'm curious how it would have gone if the teachers had been involved from the outset. (Zint, Dowd & Covitt, 2011, p. 480)

MEERA provides links to an article by Zukoski and Luluquisen in the 2002 publication *Participatory Evaluation: What Is It? Why Do It? What Are the Challenges?* The article recommended shifting the responsibility for data collection away from "professional evaluators and outside experts" and toward "the evaluator and participating stakeholders" (Zukoski & Luluquisen, 2002, p. 3). Another educator who used MEERA struggled with developing an appropriate evaluation instrument, especially one that could "precisely" measure changes in their students' "self-confidence, systems thinking and long-term behavior change" (Zint, Dowd & Covitt, 2011, p. 486). In conclusion, they suggest that "self-directed learning resources about evaluation can play a role in enhancing EE's evaluation competencies and thus, [support] . . . evaluation efforts" (p. 493).

Two evaluation reports on the MEERA website characterize the strengths of formal and non-formal educators working together to measure the environmental literacy of the public school students they serve: *The IslandWood Evaluation Project* (Kearney, 2009) and *An Elementary School Environmental Education Field Trip: Long-Term Effects on Ecological and Environmental Knowledge and Attitude Development* (Farmer, Knapp & Benton, 2007). Kearney (2009) used clicker questionnaires and a cognitive mapping instrument to complete a two-phase study that measures environmental knowledge at three time intervals (before, one week after, and 6 to 8 weeks after) students participated in the overnight program.

> Increases were found with respect to both factual knowledge and to how students conceptualize "healthy" environments. (Kearney, 2009, p. 5)

Farmer, Knapp, and Benton (2007) used open-ended phone interviews one year after a class of 30 students participated in the Parks as Classrooms program (p. 1). Although only 50% of the students responded, the transcripts were coded, and it was discovered their memories included "parts of activities, plant and animal names, ecological terminology, environmental issues, and various other ideas encountered during the program" (p. 1). Although each study measured environmental literacy not only in terms of knowledge acquisition but also class dynamics and behavior change, each study demonstrates a mutual attempt on the part of formal and non-formal educators to quantify and validate not only their students' growth but the level of uncertainty in their findings as well.

Assessments for Environmental Science Literacy—Michigan State University. Researchers at Michigan State University are testing "learning progressions that lead toward environmental science literacy—the capacity to understand and participate in evidenced-based discussions of socio-ecological systems and to make informed decisions about appropriate actions and policies—for students from upper elementary through college" (Michigan State University, 2010, para. 1). Their work focused around four strands, or curricular objectives: carbon, water, biodiversity, and citizenship (Michigan State University, 2010, para. 2). The products of their work include an environmental science literacy assessment for a unit on the carbon cycles. All except a single page of the 10-page test include questions that pertain to the cycling of carbon through the ecosystem. Questions check for students' understanding of carbon's role in photosynthesis and plant respiration as well as carbon's role in human metabolism and global climate change. In addition to science knowledge, students use their understanding of science inquiry and data interpretation to explain their answers. Every answer requires an explanation from the student. The last page surveys students' opinions regarding climate change and global warming in relation to the impact they could make in reversing damage to the environment (Michigan State University, 2010, p. 10). Because the university provides the instructional materials for teaching the objectives associated with the assessment, the content validity is inferred to be fairly high.

Doherty, Draney, and Anderson (2012) presented their research on *Methodological Issues in Developing a Learning Progression-Based Assessment System* at the National Association for Research in Science Teaching (NARST) conference in Indianapolis (p. 1). They suggest that assessment instruments must be reviewed using "statistical criteria based on measurement theory and practice" and "conceptual criteria based on learning progression theory and practice" (p. 3). Essentially, the construct, or phenomena being measured, is defined using a "construct map," and a scoring guide is created to describe the qualities of a student's response, indicating their level of proficiency (p. 6). According to Doherty et al. (2012), highly proficient samples of student work indicate a student has learned "scientific discourse [and] [sees] how systems and processes are connected, applying principles and models across processes" (p. 7). Students whose work scores as proficient "will not see scientific connections between processes, but their accounts will have similarities because they draw on a common pool of linguistic and conceptual resources" (p. 7). In terms of designing specific items on the assessment, Doherty et al. (2012) recommends "measuring students understanding of principles and models with minimal effects from scaffolding and local knowledge" (p. 10). He explains that developing a scoring guide is an "iterative process" requiring "reliability checks" (p. 14). Doherty et al. (2012) clarifies that a student's response must have all the characteristics for a particular level of proficiency described in the scoring guide, and cannot be "partially correct" (p. 13). Doherty et al. (2012) uses the discrimination and weighted mean square to measure the limitations of specific items on an assessment instrument (p. 19). The environmental literacy assessments show a high level of correlation in items related to processes and items related to practice, which allows them to make a claim regarding students' underlying level of proficiency in environmental science literacy (p. 21). Two of the steps recommended by Doherty et al. (2012) were incorporated into the proposed process used for this research project: "(1) checking coding scoring guides for construct validity, and (2) recognizing the limitations of individual items" (p. 19).

The Environmental Literacy Project at Michigan State University has developed strong measures of reliability and validity for verbal and written assessment instruments over the past five years. They recognize the iterative process of examining the scoring guides and items on the assessment instrument for construct validity. The index of discrimination is used to correlate items on the assessment with overall proficiency. They have also developed a principles-based reasoning skills instructional process, which gives students the tools to explain their responses (Rice, Doherty & Anderson, 2013, p. 3).

Ecological Understanding as a Guideline for Evaluation of Non-formal Education (EUGENE). Ecological Understanding as a Guideline for Evaluation of Non-Formal Education (EUGENE) is an interactive item bank for creating online tests that was developed by the U.S. Forest Service, University of Georgia, Environmental Protection Agency (EPA), and University of Michigan (Andrews, 2013). It is designed to support environmental educators in creating evaluation instruments. It supports research designs that use a pre- and post-test approach to measuring students' understanding of seven ecological principles. According to Andrews (2013), Eugene Odum, the Father of Ecology, taught these principles to his students. The primary limitation of EUGENE is that it only measures students' knowledge of ecological principles: adaptation, behavior, growth and diversity, energy flow, limits, regulation and emergent properties (Andrews, 2013). The pervasive attention to students' understanding of key principles in environmental science demonstrated in EUGENE underscores the motivation of educators from formal and non-formal settings to work together in developing assessment instruments for environmental literacy. EUGENE ensures a high level of validity by limiting which questions an educator can use to create an assessment instrument to seven key ecological principles, but the tradeoff is that educators cannot adapt it to meet individual needs.

Recommendations

A comparison between the validity, reliability, and equity of the instruments described in this literature review indicated that the Assessments for Environmental Science Literacy from Michigan State University received the highest rating as a potential tool for measuring students' proficiency with the OELP strands (see Table 2).

Specifically, the evaluation process used by researchers at Michigan State showed the importance of identifying specific items in an instrument for revision by using multiple trials. Gotwals and Songer (2013) discovered that it was difficult to "place

Table 2

Comparison of Existing Assessment Instruments for Environmental Literacy

Evaluation Criteria	**OAKS (Datum)**	**PISA**	**MYP**	**ODE Work Sample**	**NELA**	**MEERA**	**AESL**	**EUGENE**
Validity	0	0	1	2	0	1	2	0
Reliability	0	0	1	1	1	1	1	1
Equity	0	2	3	3	0	3	2	0
Total	0	2	5	6	1	5	5	1

Note: PISA = Programme for International Student Assessment; MYP = Middle Years Programme; ODE = Oregon Department of Education Work Sample; NELA = National Environmental Literacy Assessment; MEERA = Measuring the Evaluation Competency of Non-Formal Educators; AESL = Assessments for Environmental Science Literacy—Michigan State University, and EUGENE = Ecological Understanding as a Guideline for Evaluation of Non-formal Education. Each instrument was compared for the criteria against the datum of the OAKS and score with a +1 if the instrument's characteristics appeared stronger than the datum, and -1 if they appeared weaker. Detailed explanations for each score are found in Appendix H).

students at a given level on [their] progressions" using their scoring guides (p. 597). They chose to work closely with the NGSS framework to show their treatments and assessment instruments were "teaching students to become scientifically literate citizens, who are able to make informed decisions about pressing scientific issues" (p. 597). They indicated that the way the construct is explained in the scoring guide needed revision before the instrument was used again. Reasoning, evidence, claim, and content were key to the scoring of students' decision-making skills based on Gotwals and Songer's evaluation design.

The definition of environmental literacy strands provided by the OELP represent the efforts of educators in formal and non-formal settings to develop common understanding, which may lead to a high level of reliability and validity for the instruments they create. Existing international instruments, such as the PISA, provide an important tool for understanding how students are doing relative to their peers. Since states use national standards to frame their assessment efforts, it seems efficient to continue to develop instruments like those currently used for science literacy and environmental literacy. However, assessment instruments need to be evaluated so the data gathered by their use with students is used appropriately and with integrity.

Research in the area of assessment for environmental literacy increased since Osborne (2007) identified increasing accountability for educators and the development of curriculum-based assessments as two reasons to change educational practice. In an effort to be accountable for successful implementation, the OELP identifies possible instruments for measuring environmental literacy: the OAKS, student work samples, and an adaption of the 2008 NELA, which was used by the North American Association of Environmental Education (NAAEE) to established a baseline measure of middle school students' environmental literacy. Assuming the purpose of educators in Oregon is to show evidence of each student's proficiency in environmental literacy as defined in each of the five OELP learning strands, two other types of assessment instruments appear to be valid, reliable, and equitable: work samples that are scored by teachers so scores can be passed through the school district to the Oregon Department of Education, and MEERA, where educators can find grant evaluations that use various instruments to measure goals specific to individual programs as well as requirements identified by funders. To avoid confounding uncertainties, and provide honest information to the public about students' proficiency with environmental literacy, these lessons from previous researchers can be applied:

- Consider a student's ability to show their reasoning and evidence in evaluating their decision-making skills.
- Verify the understanding of the construct of environmental literacy among those using the scoring guide.
- Recognize that the understanding of scientific principles and personal experience influence the development of environmental literacy over time.

In addition, the level of generalization that needs to be made from the data from a particular assessment instrument needs to be considered. The research design of this study limits the generalization of findings to a small, representative group of middle school students, and educators with skills in systems thinking and science concepts associated with natural phenomena, and experience interacting with community and natural systems from the diverse eco-regions of Oregon**.**

Summary

In this chapter, educational theories were discussed to understand the reasoning behind how assessments were created for middle school students. The literature was reviewed to determine how existing assessment instruments for environmental literacy ensured that the instrument actually measured the intended construct, and could be used reliably by formal, and non-formal educators as well as assure equity for each student. In Chapter 3, I will present a critical analysis of the methodologies that are relevant.

CHAPTER 3

Research Design and Methodology

In the earlier section, Chapter 2, I presented educational theories related to assessments for environmental literacy for middle school students, and evaluated a number of existing assessment instruments for validity, reliability, and equity. The type of instrument that was determined to be most appropriate for this investigation was a student's work sample evaluated using a scoring guide. In April of 2013, the Oregon No Child Left Inside Task Force published a revision to the Oregon Environmental Literacy Plan (OELP), which highlighted the application of systems thinking for environmental literacy. Their reasoning was that systems thinking provided a foundation for understanding the whole of a system as well as the interrelationships among its parts (p. 16). They argued: "systems thinking is not limited to any one subject and can be practiced in all curricular areas" (p. 16). Based on the importance assigned to systems thinking in the revision of the OELP, the purpose of Chapter 3 is threefold: (1) to create a systems thinking scoring guide that represented a group of middle school students, and formal and non-formal educators shared understanding of the construct of systems thinking, (2) measure the level of agreement between formal and non-formal educators scoring a sample of students' work for proficiency, and (3) look for antidotes to archetypical mental models associated with assessment, equity, and environmental literacy. Based on the review of existing assessment instruments for environmental literacy in Chapter 2, a scoring guide was selected as the instrument for measuring a middle school student's proficiency with systems thinking.

The research design investigated the following research question: "What does the level of consensus between non-formal, and formal educators reveal about designing an instrument to measure a middle school student's level of environmental literacy in Oregon?" Consensus was quantified using statistical measures for validity and reliability, including percent agreement and Cohen's kappa. I also used my lived experience as a student in Oregon, non-formal educator in Minnesota, Oregon and Washington, and formal educator in Oregon to confirm and question the research findings and make changes to my own practice.

In Stage 1 of this mixed methods study, a forum composed of middle school students, teachers, and environmental educators used the Delphi technique to identify and rank four sets of skills for a systems thinking scoring guide that demonstrates "an individual's understanding, skills and motivation to make responsible decisions that consider [one's] relationships to natural systems, communities and future generations" (OELP, 2013, p. 4). The Delphi technique provided a forum in which members could anonymously express their ideas and level of agreement so students' voices were equal to those of educators. In Stage 2, the scoring guide was field-tested online with a sample of students' work. Educator networks invited their members to score the students' work and provide feedback about the scoring guide using a link to a website created to collect data for the study. The educators identified themselves as either a formal educator, who assigns credit to student work, or as a non-formal educator, who does not. As Stage 1 and 2 were underway, I wrote frequent reflections about my lived experience as a student and educator for an autoethnographic narrative. The resulting narrative, titled "My Feet of Clay," provides social, cultural, and political context for this study, which revealed my deepened understanding of the constructs of assessment, equity, and environmental literacy. By using the Delphi technique, field-testing the scoring guide, and weaving together an autoethnographic narrative employing a constructivist approach, the voices of today's middle school students, and formal and non-formal educators in Oregon could clearly be heard above my own (see Figure 5).

Figure 5. Research Design for Revealing Commonalities between Educators

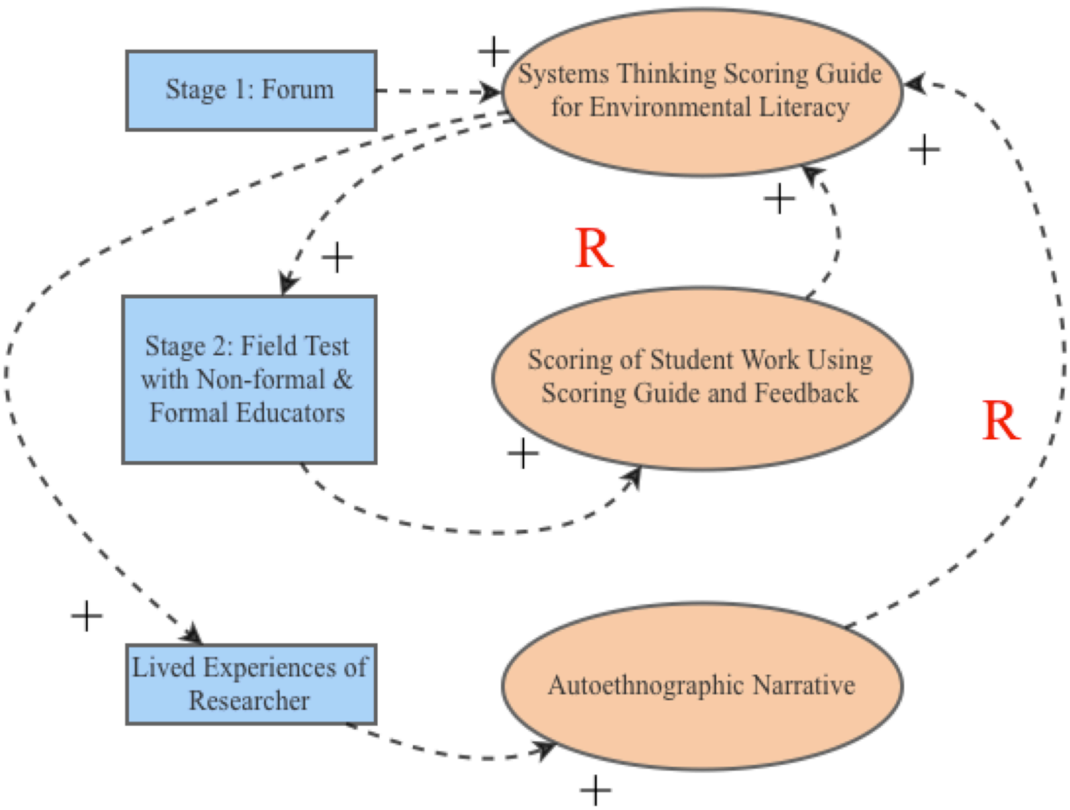


Figure 5. Research Design for Revealing the Commonalities between Educators involved in this investigation titled, Evaluating an Assessment Instrument for the Oregon Environmental Literacy Plan (OELP). The arrows indicate the flow of information between the forum composed of middle school students and educators, the formal and non-formal educators participating in the field test of the instrument, and the researcher's lived experience. The + symbol indicates increasing perspectives. The capital letter R identifies reinforcing loops where commonalities for assessing environmental literacy might be uncovered.

Mixed Methods Research Design

To design and evaluate an instrument to assess the environmental literacy of middle school students in Oregon, this mixed method study combined the quantitative analysis of the Delphi technique and field-testing with the qualitative aspects of autoethnography. Brewer and Hunter (2006) defined mixed methods research as "either single studies or more complex programs of continuing research, which systematically employ various combinations of field, survey, experimental, and nonreactive methods to address their research questions" (p. 14). Marcinkowski et al. (2012) documented a trend through content analysis of 248 dissertations published the *Environmental Education Research Journal* between 1991 and 2000. He found 49% percent used quantitative methods, 23% used qualitative methods, and 14% used mixed methods (p. 58). A mixed methods design was selected so quantitative methods could be used to help manage bias. Sosu, McWilliam, and Gray (2008) found that a mixed methods "approach can serve the dual role of confirming and elaborating findings" (p. 169). By using the qualitative methods of autoethnographic narrative, I invite those who read my stories to stand with me as I described what it was like to experience assessment, equity, and environmental literacy in formal and non-formal contexts. A mixed methods design supported the use of constructivist theory for this study, which measured validity and reliability using the commonalities between formal and non-formal educators. The scoring guide for systems thinking was analyzed from multiple perspectives—middle school students, formal and non-formal educators, and the researcher—who might otherwise have worked at cross-purposes due to the value their disciplines place on the deductive and inductive reasoning.

Based on Nuthalapaty's (2010) description of the distinguishing characteristics of qualitative and quantitative research designs, the study employed a mixture of methods (see Table 3). Commonalities between middle school students, formal educators, and non-formal educators creating the scoring guide were measured using an 80% level of agreement. Inter-rater reliability was measured in a field test of the scoring guide using Cohen's kappa. In addition, the major themes of the autoethnographic narrative revealed

Table 3

Overview of the Application of Nuthalapaty's Definitions of Mixed Methods

Qualitative	Quantitative
Explores the use of environmental literacy assessments in Oregon.	Comparative experiment between two groups of educators using a scoring guide for systems thinking as described in the 2010 OELP.
Gains insights into the construct, or phenomenon, of environmental literacy using the Delphi technique with a group of experts.	Driven by the null hypothesis: That there is no difference between the reliability of formal and non-formal educators using the scoring guide when compared to chance. H_1: Cohen's kappa = 0 or < 0 H_0: Cohen's kappa is between 0 and 1
Analyzes information from three groups: formal educators, non-formal educators, and students.	Measures the validity of scoring guide using percent agreement to identify potential difference in understanding of construct between formal and non-formal educators.

Note: The descriptors in Slide 15 from Nuthalapaty's (2010) slide presentation found at http://www.slideshare.net/fnuthalapaty/educational-research-102-selecting-the-best-study-design-for-your-research-question-3062530, were applied and specifically described in the context of this investigation.

potential unintended consequences of using a scoring guide to assess a middle school student's proficiency with systems thinking for environmental literacy. The mixed methods design ensured that the assessment instrument measured environmental literacy as its authors intended, with attention to the degree of reliability confounded by differences in the understanding of particular constructs.

Stage 1: Creating a Scoring Guide

Selecting participants for the Delphi technique. Since the intention of the scoring guide was to support the assessment practices of formal and non-formal educators, those who developed it needed to represent both educators and middle school students. The validity of the results depended on the selection and participation of these individuals (Keeney, Hasson & McKenna, 2011, p. 47). The selection criteria for participating in the forum was not intended to exclude potential participants from the panel. According to Keeney et al. (2011), "the more focused the criteria, the greater the limits are placed upon the study's findings" (p. 48). The selection criteria included on the Letter of Consent was:

- Share your ideas about the tools and skills used by systems thinkers
- Review examples of students' work using the links at http://goo.gl/sMLxM7.
- Access the Oregon Environmental Literacy Plan (OELP) at http://goo.gl/dAcGqk and read pages 16-24.
- Review a 1-page chart at http://goo.gl/D16ijT showing Next Generation Science Standards (NGSS) associated with the OELP.
- Optional: Learn more about systems thinking through self-study using these videos at http://goo.gl/CG7Ixg.

The selection criteria limited the generalization of the results from Stage 1 to educators and students with experience in systems thinking and/or a willingness to learn how to use its tools such as Connection Circles, Behavior Over Time Graphs (BTOGs), the Iceberg Model, and the Ladder of Inference.

The forum consisted of 11 people: two formal educators, three non-formal educators, and six students. The reason for using a small panel of experts was that they were representative of a population that could directly benefit from using the scoring guide. Inferences made from the OELP suggested: (1) formal educators might use the scoring guide to measure middle school students' proficiency using the crosscutting concept of systems described in the Next Generation Science Standards; (2) non-formal educators might use the scoring guide as an alternative measure of students' proficiency for meeting graduation requirements, or as a tool for reporting efforts to meet their program goals; and (3) students could use the scoring guide to improve their work with systems thinking. The purpose of uniting these voices in anonymous consensus was to assure that those who use the scoring guide in the future might understand its meaning because it was written by middle school students and educators. The scoring guide was not copyrighted, so others could use it to stimulate further discussion about measuring middle school students' proficiency with environmental literacy.

In order to initiate a discussion that would produce a systems thinking scoring guide for environmental literacy, superintendents (or their representatives) were contacted from urban and rural school districts in Oregon. A school district was approached based on whether middle school teachers had participated in workshops for systems thinking or were actively teaching systems thinking to middle school students. With approval from the school district, principals were asked to forward the invitation to participate to teachers, who then shared the invitation with their students. Non-formal educators were invited to participate through an announcement made at a conference of their professional association. Interested educators and students (with parent/guardian signature, if under age 18) signed Letters of Consent prior to participating in Round 1 of the Delphi, which involved responding to an online survey. Members were encouraged, but not required, to do self-study online using the *Systems Thinking in Schools* modules at WebEd, available through the Waters Foundation website. No one was turned away, and members had the freedom to participate as much or as little as they chose. To safeguard anonymity, I purposefully coded responses by color so I did not know if the member had identified as a student, formal educator, or non-formal educator. I intentionally used an electronic form where I had no way of knowing whose responses were whose. Members of the forum were encouraged to remain anonymous so that the ideas of each person on the panel would be equally valued. Each one had an equal voice and timeframe to respond to questions asked using a form distributed electronically. Any power differentials—often created by nonverbal responses, tone, controlling conversation, age, race, or gender—were masked by using electronic forms with the Delphi technique. The researcher's role was to catalogue and statistically rank ideas for review by the panel. The Delphi technique involved contacting experts, preparing and administering the forms, completing content analysis, and iterating the discussion twice more to identify skills to include in the scoring guide for systems thinking for Oregon's Environmental Literacy Strand 1.

Sequence of events for the Delphi . The Delphi technique was selected to address one of Gough's (2013) guiding principles for research in environmental education: the need to give each voice an equal opportunity to be heard (p. 10). The following series of questions guided the collection of data during the process of constructing a scoring guide for systems thinking with students and educators:

- What ideas and explanations do formal educators/non-formal educators/students use to describe decisions that are made with communities and natural systems in mind?
- Which systems thinking tools do formal educators/non-formal educators/students use, or reference, as they describe the skills middle school students need in order to explain their decision-making process?
- How do formal educators/non-formal educators/students prioritize the qualities associated with a proficient sample of a middle school student's work?
- How strong is the shared understanding of key ideas between members of the forum in terms of the level of consensus?
- Do any of the key ideas agreed upon by the forum resonate with the language used to describe achievement in the Next Generation Science Standards?

The Wiley publishing company granted permission to adapt, but not publish, the instructions and forms created by Keeney, Hasson, and McKenna (2011) to collect data. The role of the researcher was to be merely a conduit for anonymously passing ideas between members of the forum and to calculate their level of consensus. To avoid bias as the researcher, I wrote a scoring guide prior to the study so it could be used for comparison if necessary. The forum completed the scoring guide for systems thinking in three rounds over a three-month period.

Round 1: Open-ended questions. In Round 1 of the Delphi technique open-ended questions were used to discern the four most important steps and the associated skills that middle school students needed to demonstrate in order to be considered proficient in OELP Strand 1—Systems Thinking. The participants were asked to indicate, using a percentage, the level they would like the forum to reach for consensus. The median of all responses was used for the level of consensus. Content analysis in the Delphi technique used an approach similar to the qualitative approach used in phenomenology to collapse participant responses into themes. However, it was particularly important in the Delphi technique to be "as true to the wording as possible" (Keeney et al., 2011, p. 85). The statements that were “the same, or so similar that they mean the same thing” were identified first (Keeney et al., 2011, p. 85). These statements were collapsed and phrased to represent the meaning so the participants would recognize their original statement in the next round. However, to ensure the accuracy of each person’s position redundancy of forum statements was preferred to oversimplification. The skills were organized into four steps, or categories, named by the members of the forum. In this way, the information was organized in a format that looked like other scoring guides used to assess students’ work samples in Oregon.

Survey to avoid mono-operational bias. The scoring system was proficiency-based, and the scoring guide divided the qualities of the work in terms of meeting proficiency and then being able to teach others. Sturgis (2014) explained that proficiency-based learning systems are characterized by communicating expectations to students; scoring students’ work relative to those expectations rather than other students; and students can demonstrate their skills in “multiple ways” through “multiple opportunities” (para. 3).The scoring guide was used to reduce specific threats to validity identified by Trochim (2006): “preoperational explanation of the construct, and mono-operation bias” (para. 4). The forum in Stage 1, operationally defined the construct of systems thinking as a measure of environmental literacy. Mono-operational bias was avoided by involving formal and non-formal educators as well as students, but any claims made from the findings were limited to this one group’s operational definition of what proficiency in systems thinking for environmental literacy looked like in the work sample of middle school students.

Round 2: Rating skills on level of importance. The purpose of Round 2 was to reach consensus on the four categories and the skills that students must demonstrate to be proficient. The amount of information was unwieldy, as expected. In the second round of the Delphi, participants rated over 50 skills on a one to five Likert-type scale based on the level of importance. The forum used the following scale created by Keeney et al. (2001) to reach 80% consensus for the four steps and skills to include in the scoring guide:

1. Very unimportant
2. Quite unimportant
3. Neither important or unimportant
4. Quite important
5. Very important (p. 88)

The numerical ratings that participants assigned were disassociated from the steps and skills so they could be statistically analyzed without bias. The Statistical Package for Social Scientists (SPSS) was used to find the median, standard deviation, and mean for each set of ratings. Cohen's and Fleiss' kappa could not be calculated because responses were received anonymously, so they could not be coded to the same person from one round to the next. The following procedures outlined by Keeney et al. (2011) were used to analyze the data,

1. Each statement should be set up as a separate variable.
2. Responses from each expert [on the] panel should be inputted to the SPSS database alongside their master code.
3. Frequencies should be run on the entire dataset. This will provide output on the percentage of each overall statement. (p. 86)

The specific steps for inputting data and running the software can be found in Appendix I: Data Analysis Procedures for the Scoring Guide. The output showed the percent agreement for each statement; in this case, the importance of each skill to include in the scoring guide. The frequency output tables for each statement were used to prepare for Round 3 of the Delphi.

Survey to increase construct validity. Instruments from environmental education and science education literature were examined to identify the issues educators face when reporting students' level of proficiency. Guided by the Delphi technique, involving analysis of responses to survey questions and rating each member's level of consensus, an instrument was developed for measuring construct validity. One key variable to control in this part of the study was the understanding of different constructs, or phenomena, related to systems thinking and environmental literacy. Common understanding of proficiency and clear definitions for how to demonstrate systems thinking skills were indicators of strong construct validity. The issue of construct validity was given highest priority in this study because it measured the degree to which the assessment instrument actually measured the phenomena it is intended to measure. Brown (2000) indicated that "an accumulation of evidence…. using content analysis…demonstrating differences between differential groups" helps make the case for any claims regarding the strengths of a particular assessment instrument (p. 10). In this study, specific consideration was given to the statements in the scoring guide that described the phenomena, or the construct of environmental literacy in terms of systems thinking. This was identified as Strand 1 in Chapter 3 of the revised Oregon's Environmental Literacy Plan (2011).

Round 3: Reaching consensus. The purpose of Round 3 of the Delphi was to reach consensus regarding which skills to include in the scoring guide so it could then be field tested with a sample of students' work. Each participant received the forum's percent agreement on each skill, in terms of its perceived importance, along with the standard deviation and median. A smaller standard deviation represented less of a range of opinion for the reported level of consensus. For example, if the median was 2 and the standard deviation is 1, there was little variation from the mean regarding the importance of including the skill. Those statements that had already reached consensus were highlighted because they required no further action by the forum. In this final round, participants were asked to reconsider their ratings only on those statements that had not yet reached the consensus percentage.

The ratings from Round 3 were analyzed with SPSS using the same procedures as in previous rounds. When participants changed their ratings between Round 2 and Round 3, the change translated into a larger, or smaller score, and thus impacted the median and standard deviation. The change was used to determine if the group was moving toward, or away from, consensus. Those statements that gained a level of 80% consensus, based on their ratings as quite important and/or very important, were added to the scoring guide. The actual reasoning for including specific skills based on mathematical statistics was included in with Appendix N. At this point, the goal of the Delphi, to create a scoring guide for OELP Learning Strand 1–Systems Thinking, was complete.

Stage 2: Field Testing the Scoring Guide

Selecting an example of students' work with systems thinking. Ideally, students' work samples would have been selected from a representative sample of students of various ethnicities (see Figure 6). The actual example of students' work scored in the field test was created by a number of students who had participated as members of the forum. By agreement, the students' work was only used to test the

Figure 6. Ethnicities of Middle School Students in Oregon 2012-2013

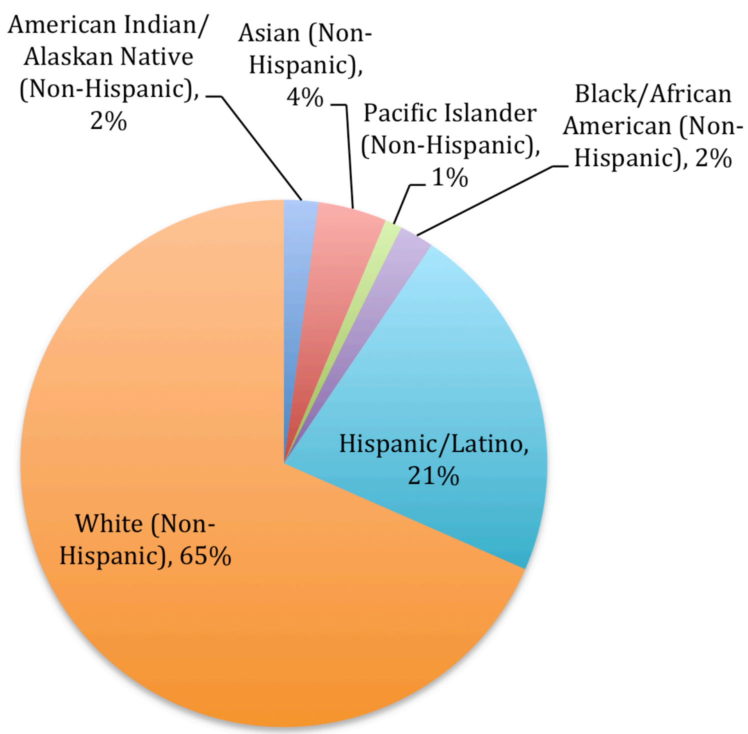


Figure 6. The percentage of students by ethnicity for selecting a representative sample. Calculated using Oregon Department of Education's *Fall Membership Report 2012–2013*, which included students enrolled in public schools in grades six through eight on October 1, 2012. The total number of public middle school students in Oregon during the 2012–2013 school year was 129,791 (ODE, 2013).

scoring guide, and not intended to serve as an exemplar. The students created the work outside of school, and it was not an assignment for credit. The students did not use the scoring guide to assist them in creating it. The scores that educators assigned to their work using the scoring guide were not shared with them, nor were they published, because of the expectation of, and commitment to protection of privacy between educators and students concerning assignments for grade. In this study, the appropriate use of the data from the field test was to measure the inter-rater reliability of educators using the scoring guide, not providing feedback to the students or making claims about the level of proficiency demonstrated in the students' work.

Selecting a representative sample of educators. The estimated ratio of formal educators to students in Oregon in 2013 is 1 to 20. This estimate was based on the number of public school teachers reported in the 2006 Oregon Teacher Standards and Practices publication, *Educator Supply and Demand: Implications for Staffing Oregon Schools*, and statistics in the Oregon Department of Education's *Fall Membership Report 2012–2013*. If 129,791 students were in middle school, the estimated number of formal educators working with them would be 6,379. Based on a model of middle school that placed students in three core classes each day, a third of these teachers might be asked to assess students' work for evidence of the Oregon Environmental Literacy Plan Learning Strands. So, the estimated number of formal educators is $N_{\text{formal}} = 2{,}126$.

The number of non-formal educators that work with middle school students is more difficult to estimate. However, Traci Price, Oregon Environmental Literacy Task Force chairperson, on July 1, 2013 estimated that "there are over 450 conservation education providers in the Metro region, and the Environmental Education Association of Oregon listserv reaches about 2,000 people" representing both formal and non-formal educators. For a fair test, the highest of the available estimates was $N_{\text{non-formal}} = 2{,}000$.

Because this study aimed to compare the scoring results of the two above groups of educators, the total population of educators represented by this study is approximately $N_{\text{educators}} = 4{,}000$. Using a sample size of n=94, the findings could be generalized to the population of educators working with middle school students to a confidence level of 95% and a confidence interval of + or – 10 educators. Since two groups were needed, an equal number of educators needed to self-select themselves into each group—formal and non-formal. $n_{formal} = 50$ and $n_{non\text{-}formal} = 50$.

Measuring inter-rater reliability. Inter-rater reliability was used to examine the agreement between the two groups. The raters consisted of 11 formal and 14 non-formal educators, who scored the students' work sample using the systems thinking scoring guide created by the forum. Educators were contacted through regional and statewide professional organizations for formal and non-formal educators with all levels of experience. Fifteen organizations were contacted for permission to send their membership the invitation (included in Appendix M), out of which, six organizations sent the announcement to their membership electronically. To prevent bias, and maintain my objective stance, my name did not appear on the invitation. The electronic form that was used to collect educators' ratings on the proficiency of the students' work made available for three months. Rather than provide a signature of consent, participants marked a box if they agreed to the following:

> I am an educator over the age of 18 years, who provides knowledge, skills and experiences that support education for environmental literacy. I have volunteered, or worked, with middle school aged children. By submitting this form, I consent to the use of my scores to compare the reliability with which formal and non-formal educators can use a scoring guide for systems thinking. If I chose to provide feedback regarding the scoring guide, my ideas can be used to improve the scoring guide and make recommendations concerning its use. I also agree not to copy, print, share or distribute this sample of students' work in any way.

Although the study failed to reach a statistically significant sample size to establish generalizability of the findings, the balanced response from formal and non-formal educators could be used to estimate the level of inter-rater reliability.

Collecting educators' scores for students' work online. Stage 2 of the study measured the inter-rater reliability between formal and non-formal educators from across Oregon testing the systems thinking scoring guide created by the forum. A quantitative analysis of the scores assigned by formal and non-formal educators was used to prove or disprove the null hypothesis that the differences in scores between the two groups is just as likely due to chance as it might be to any reliability in scoring. The educators who participated were treated as if their proficiency scores were used to assign academic credit. However, rather than using a ordinal scale for proficiency, the students work was scored simply based on whether it was proficient as described by the skills listed in the scoring guide. The rating scale was: 1 – No evidence; 2 – Not there, yet.; 3 – Proficient; and 4 – Highly proficient, appears able to teach another. This scale was used to prevent confusion between measuring the proficient use of systems thinking skills described in the scoring guide for making decisions versus potentially measuring whether the students' decision was a responsible one based on the opinion of the educator. Traditionally, formal educators receive scoring guides adopted by the Oregon Department of Education and are trained to use them until their scores calibrate; which is to say, they can score a sample of students' work with an accuracy of plus or minus one proficiency level. The field test was completed using the assumption that the scoring guide explained four sets of skills for systems thinking without the need for training and calibration. Percent agreement and Cohen's kappa was calculated to determine whether a group of formal and non-formal educators could score a sample of students' work more reliably than would be ascribed by chance.

Appropriate Use of Inter-Rater Reliability

Trochim (2006) stated that inter-rater reliability is "used to assess the degree to which different raters/observers give consistent estimates of the same phenomenon" (para. 2). For a general measure of agreement, a percentage was used to indicate the number of times raters gave students the same proficiency rating. Because a common scoring guide was used, raters showed where students fell along a continuum representing educational growth. The amount of correlation between the raters was estimated. The typical confounding variables identified for a lack of consistency, time and calibration, were accounted for by assuming each rater took all the time they needed to rate the students' work online, and by providing an opportunity to give feedback based on their understanding of the scoring guide. This feedback from this study could be used in future studies to improve the reliability and reaffirm the construct validity of the instrument. On the other hand, it would create an opportunity for a group of raters to skew the scale higher or lower. This study did not identify anchor papers, or particular case studies, to be archived for reference and future training purposes, which could have also supported the general understanding of systems thinking in environmental literacy over time. In this study, educators were given access to one work sample and did not have the opportunity to calibrate with one another.

Calculating Cohen's kappa. Cohen's kappa was selected to measure inter-rater reliability because the two groups of educators' scores could not only be correlated with one another, but also compared to a chance. This was important because the investigation did not include a control group. Cohen's kappa was suitable for analyzing data that can be put in ordinal and discrete categories. According to Zaiontz (2016), the two conditions were met to use Cohen's kappa: there were "exactly two raters (or groups of raters considered a unified group) and each rater [judged] all the subjects" (C. Zaiontz, email communication, March 24, 2016). The scores assigned to students' work samples were based on the following scale: (4) highly proficient (appears able to teach another person); (3) proficient; (2) not proficient, yet; (1) no evidence available in this work sample. This rating system was similar to those currently used in scoring guides adopted by the Oregon Department of Education and schools in the Middle Years Programme of the International Baccalaureate, but distinctly different in that proficiency was the only category described in the scoring guide.

Statistical measures of central tendency and reliability. When scores were more than one category apart, the score of a third person was not available to as has traditionally been the strategy used to determine the score. This strategy was used by Bartosh, Tudor, Ferguson and Taylor in their recent publication, *Impact of Environment-based Teaching on Student Achievement: A Study of Washington State Middle School*, and the Oregon Department of Education. Since the design of this study was to find the degree to which educators actually do give the same score without training and repeated practice, the third person technique was not used to generate a score for the students' work sample. The mode of scores assigned by the educators who score students' work samples was used.

Cohen's kappa is statistical measure that can be used to see how scoring systems works. Kappa controls for agreement against the value of agreement one would expect to find based on chance alone. A reliable scoring system for a systems thinking work sample would show a kappa value close to one. A negative number, or a number close to zero indicates that the agreement is more likely due to chance alone. To set up the table in SPSS, the number of scores representing each set of criteria was entered as nominal data. The first column was labeled, “sample,” in reference to the students’ work, and the second column was labeled, “RaterA” for the formal educators, and the third column was labeled, “RaterB,” for the non-formal. If there was a difference in score of more than one level of proficiency, the educators’ scores was were still included. The mean score of each groups was used to calculate kappa.

To run the kappa value for each part of the scoring guide, a new table was made before running the analysis. To run the analysis, *Analyze – Descriptive Statistics – Crosstabs* was used. *RaterA* was put in the box labeled, “Column,” and *RaterB* was put in the box labeled, “Row.” Then, *Statistics-Kappa* was clicked. To get comparative information regarding what the value of the scores would have been if due only to chance, *Cells* needs to be clicked, and both the *observed* and *expected* boxes needed to be marked. The difference between the observed and actual value indicated whether the scores between formal and non-formal educators were significantly different.

Interpreting Cohen’s kappa statistic. In the example shown in Table 4, the kappa statistic for the amount of agreement between the scores on students work samples from the formal and non-formal educators could be as low as .186, which indicates poor agreement. However, the value for the approximate significance might be less that one suggesting some statistical significance. The raters could agree that two of the work samples were WTP and the chance of both raters scoring two work samples, as WTP could be only 0.4. The agreement for HP may also be above the level of agreement expected due to chance. These values suggest that even though the level of agreement between raters may be poor overall, the raters do tend to agree on scores for work samples on either end of the continuum from WTP to HP. By using Cohen's kappa, the reliability of the scores was compared to chance rather than only to one another. So, these results can serve as a benchmark for future research.

Table 4

Sample Table of Kappa Statistic Comparing Two Groups of Raters

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
RaterB * RaterA	19	100.0%	0	.0%	19	100.0%

RaterB * RaterA Crosstabulation

			RaterA				Total
			HP	NP	P	WTP	
RaterB	HP	Count	2	4	4	0	10
		Expected Count	1.1	4.7	2.1	2.1	10.0
	NP	Count	0	3	0	2	5
		Expected Count	.5	2.4	1.1	1.1	5.0
	P	Count	0	2	0	0	2
		Expected Count	.2	.9	.4	.4	2.0
	WTP	Count	0	0	0	2	2
		Expected Count	.2	.9	.4	.4	2.0
Total		Count	2	9	4	4	19
		Expected Count	2.0	9.0	4.0	4.0	19.0

Symmetric Measures

		Value	Asymp. Std. Error[a]	Approx. T[b]	Approx. Sig.
Measure of Agreement	Kappa	.186	.123	1.716	.086
N of Valid Cases		19			

a. Not assuming the null hypothesis.
b. Using the asymptotic standard error assuming the null hypothesis.

Note: Table 4 is from the unpublished work of the author while learning to use SPSS to calculate Cohen's kappa between two groups of raters with samples of students' work. HP = Highly Proficient; NP = Nearly Proficient; P = Proficient; and WTP = Working Towards Proficiency.

Autoethnographic Procedures

Chang (2008) defined autoethnography as "cultural analysis and interpretation with narrative details" (p. 46). An autoethnographic narrative was written, using excerpts from my research journal, to attend to Gough's (2013) guiding principle for educators concerning the deconstruction of "stories of which we are a part" (p. 10). Haluza-DeLay (2013) wrote, "Children lack the resources to address problems arising on the sociopolitical level" (p. 399). However, I witnessed students choosing to act on their decisions by designing solutions to problems in their families and communities concerning access to clean water, clean air, and fatal interactions between wildlife and humans. As a science educator working in an Oregon school district, my professional life encompasses giving students opportunities to gain knowledge, skills, and experiences that they can use to make decisions and to act as members of the natural systems and communities in which they live now. In my opinion, designing an assessment instrument that measures the ability of each unique student, and meets the individual needs of each student, is possible. I believe equity depends on an educator's commitment to listening to each student define the context for decisions one chooses to make with the level of understanding, skills, and experiences they have at the time. Each student's decision is likely dependent on factors and parts of systems that I may never experience, cannot understand, and do not have the skills to address. In the back of my mind, I hold on to the possibility that a student might have just what it takes to make a big difference in their little corner of the planet. Entries were added to my research journal from the time of HSRRC approval until the first draft of the data analysis was composed.

Provocateurs. Peter Senge introduced the value of provocateurs to me at the systems thinking conference I attended in July 2015. The role of provocateurs is to challenge our ways of thinking and open our minds to one another. Senge's earlier work, *The Fifth Discipline: The Art and Practice of the Learning Organization* (1990) provided a formative foundation for my early efforts to leverage my skills for outreach as a non-formal educator. In order to prevent bias and challenge my mental models, I selected the authors of books—shared with me by neighbors, friends, and my advisor, on the topics of systems thinking, child development, indigenous wisdom, curriculum development, and well-being—to be my "provocateurs" (P. Senge, July, 12, 2015, *personal communication).* Donella Meadows' (2008) explanations in *Thinking in Systems: A Primer* were instrumental for crafting my analysis. Alison Gopnik's (2010) thoughts on child development, as outlined in *The Philosophical Baby: What Children's Minds Tell Us About Truth, Love and the Meaning of Life*, helped me imagine alternatives, or "counterfactuals," with middle school students who attend to both "exogenous and endogenous" factors as they make decisions (p. 110). Kimmerer's (2013) book, *Braiding Sweetgrass: Indigenous Wisdom, Scientific Knowledge and the Teachings of Plants,* provided a model for using research language in a way that transcends disciplines. For example, Kimmerer writes, "Doing science with awe and humility is a powerful act of reciprocity with the more than human world" (p. 252). In 2001, Jim Martin showed me how to use all of one's senses to find the "curriculum embedded in the environment," and teach others using the 5E model: "engagement, exploration, explanation, elaboration, and evaluation" (Bybee, 2006, p. 1). Munir Fasheh shared his experience of harmony by instructing me to "remove harmful words, ideas, convictions and perceptions from our minds due to manufactured and processed knowledge" (Fasheh, 2015, p. 33–34). By interacting with other's ideas, I better understood why I wrote a little sign for a few remaining plant starts that my science class transplanted to give away at Family Night. It read, "If you take care of this little plant, it will take care of you." A picture of three fully-grown green peppers popped up in my school email a few months later.

Demographics and sense of place. Autoethnography helped me synthesize various roles in which I had served the public. These included volunteer and work-related seasonal and temporary efforts, funded by governmental organizations, grants, nonprofit organizations, industry lobbies, and community partnerships. My teaching practices were adapted to serve families with the demographics shown in Figure 4, which were compiled using the U.S. Environmental Protection Agency's (2015) EJSCREEN: Environmental Justice Screening and Mapping Tool (p. 8). A number of areas in the neighborhood scored in the top 90% of environmental justice concerns as compared to other areas in Oregon. The EPA (2015) defines environmental justice as such:

> The fair treatment and meaningful involvement of all people regardless of race, color, national origin, or income with respect to the development, implementation, and enforcement of environmental laws, regulations, and policies (p. 6).

The U.S. Environmental Protection Agency's (2015) EJSCREEN identified a number of health risk factors including: the amount of particulates in the air greater in size than 2.5 microns, proximity to high volumes of traffic, ozone levels, distance from a toxic site on the National Priorities List, and proximity from effluents discharged to the water from industry. More than 50 locations were identified as hazardous due to the use of chemicals by industry. One area rated high on the environmental justice index for particulates.

Figure 7. Demographics of School Neighborhood

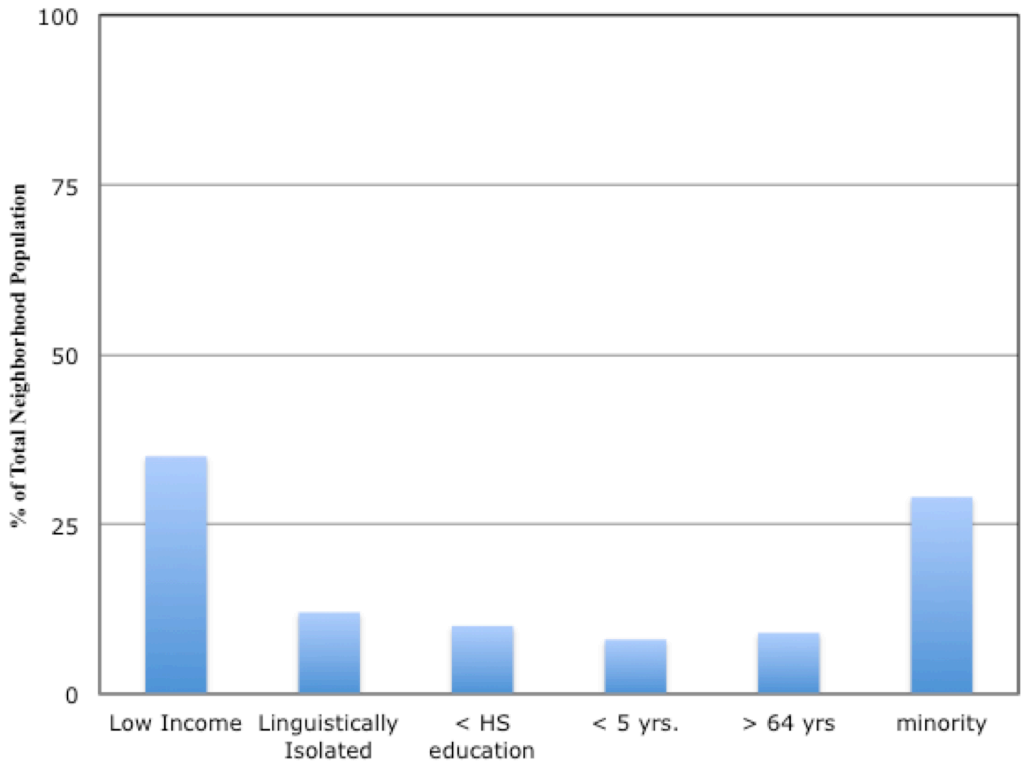


Figure 7. Demographics of School Neighborhood with an Estimated Population of 20,000 people. Compiled with EJSCREEN from U.S. Environmental Protection Agency data based on 2008–2012 American Community Survey.

About one third of the neighborhood was scored in close proximity to traffic since a highway passes through it. The school neighborhood was measured at about 100 square kilometers, and the longest possible walk from home to school was 5 kilometers, which takes an able-bodied person about an hour. Through the years, the school neighborhood's relatively large amount of undeveloped, natural areas were preserved. A barred owl made its home under the covered play area where we found regurgitated bones of small mammals. A peregrine falcon once scared away a mother squirrel in the school courtyard. We set up a small pool for the ducklings of a pair of mallards that decided to nest there, too. The resident flock of geese winter-over on the soccer field, which was once a wetland at the base of ancient lava carved away by an epic flood of water and rock released from melting glaciers.

Reflective journals. Reflective journals were kept to broaden my perspective as a formal and non-formal educator, as suggested by Haluza-DeLay (2013), who says, "environmental educators have reduced the scope of environmental sustainability and missed the opportunity to connect with more people and potential allies among a broader reach of civil society organizations and other educators" (p. 394). Certain questions guided my inquiry:

- What kind of language do educators use to describe how to measure environmental literacy?
- What do individuals believe about designing assessment instruments for environmental literacy?
- What kind of social structures exist to support equity in measuring environmental literacy?
- Where do commonalities exist for assessing environmental literacy between educators and students? Formal educators and non-formal educators? Educators and their stakeholders?

These questions served as reference points to validate the constructs of systems thinking and proficiency as defined by the forum that created the systems thinking scoring guide in Stage 1, and to situate the feedback from formal and non-formal educators who used it to score an example of student work in Stage 2.

Chang (2003) wrote, autoethnography "shares the storytelling feature with other genres of self-narrative, but transcends mere narration of self to engage in cultural analysis" (p. 43). My analysis involved analyzing what he called "border-crossing experiences" to reveal commonalities between the scoring guide, field test, and my lived experience of assessment and environmental literacy (p. 73). I heard five themes emerge from my lived experience as a student, environmental educator, and science teacher. Each spoke to my understanding of being fairly assessed and making "responsible decisions that consider [one's] relationship to natural systems, communities and future generations" (OELP, 2013, p. 4). The themes shown in Table 5 were summarized by topic and used to craft a narrative that addressed the guiding questions.

Table 5

Themes for Autoethnographic Narrative

Social Themes	Cultural Themes	Political Themes
Instructional relationships Fair assessment depends on instructional relationships that grow from attachments to natural areas as well as *understanding* observational and numeric data collected in the field.	**Deep culture decision-making** Decision-making is a deep culture *skill,* so honoring familial culture and involvement in local community is critical for constructing the meaning of *responsible*. **Educator as anthropologist** Systems thinking **motivates** educators and students to consider their *interrelationships* with natural systems. "A systems thinker steps back to examine the dynamics of a system and the interrelationships among its parts" (Waters Foundation, 2016, para. 1). It encourages stances beyond those taken by a curious anthropologist, who objectively simplifies unfamiliar, complex systems, or a missionary with a singular message that assumes sovereignty.	**Student sovereignty over defining the context** The context and relevance of the problem or opportunity, solution or possibility, and determination of one's interactions within a system's boundaries, purpose, elements, causal loops, archetypes, and antidotes belongs to each student due natural variations in *communities*. **Counterfactual futures** The exchange of credit for environmental literacy need not be limited to decisions based on the adoption of best management practices, capital investments in technology to extend the life of a renewable stock, or archetypical solutions to the Tragedy of the Commons, but rather helping middle school age students move into their *futures* valuing their sensory and academic skills in fairness and love with the "more than human world" (Kimmerer, 2013, p. 252).

Note: The words in italics resonate with the Oregon Environmental Literacy Plan's definition of environmental literacy, "An individual's understanding, skills and motivation to make responsible decisions that considers [one's] relationships to natural systems, communities and future generations" (Oregon Department of Education, 2010, p. 4). I listened for commonalities in language when I was reflecting on situations, or reading texts, describing experiences working with students in formal and non-formal settings.

In “My Feet of Clay,” I present my typical ways of thinking, or mental models, of assessment and environmental literacy, and I looked for antidotes so my thinking might value each person’s ability to make responsible decisions in accordance with their particular understanding of the systems in which they interact.

Rationale. I chose to include my own lived experience in the study because of the intentional choice I made 10 years ago to move forward in a career in non-formal education by teaching public school. At the time, I planned to serve for 10 years as a middle school science teacher and then return to my previous work providing environmental education support to teachers. I hoped a deeper understanding and awareness of school systems and their limitations would provide me with authentic credibility for encouraging educators to teach sustainability. I wrote stories in my familial language, using word pictures and diagrams, in order to connect with formal and non-formal educators as well as family and friends. I asked permission to include in my dissertation unsolicited stories that others shared with me in the course of our lives together. I situated my reflections within outdoor education, teaching practice, current legislation, and literature.

Role of the researcher. Mingé (2013) defined autoethnography as “a research practice that attends to sensory discourses, local concerns, and mindful action” (p. 427). Collecting data involved keeping reflective journals in response to my lived experience during the periods of April 2014 to May 2014 and January 2015 to August 2015. My roles included advocate, small group leader, facilitator, fourth-generation teacher, participant, citizen, volunteer, family member, and friend. I chose to interact with people and their ideas concerning child development, instructional time, Outdoor School, environmental literacy, intercultural communication, the Next Generation Science Standards (NGSS), and systems thinking. I actively participated and facilitated a number of events in which I was involved during the period of reflective data collection.

Artifacts. I gathered a variety of artifacts for the narrative, including documents, handwritten journal entries, testimony, legislation, photos, nametags, curriculum, and a podcast. As recommended by Mingé (2013), I categorized the data and participated in a variety of "sensory discourses, local concerns and mindful actions," including, but not limited to those shown in Table 6.

Table 6

Types of Experiences Described in Reflections

Sensory Discourses	Local Concerns	Mindful Actions
• Outdoor school • Environmental learning center • Interpretive center	• Legislative meetings • Rallies • Hearings	• Facilitator training to lead workshops for environmental education curriculum
• Field trips • Classroom visits	• Legislative meetings refining the definition of "instructional time"	• Cross-cultural communication class
• Volunteering • Teaching middle school	• No Child Left Inside legislation • Elementary and Secondary Education Act	• Systems thinking workshops • International conference about systems thinking
• Family and social events	• Oregon Environmental Literacy Plan	• Statewide science teachers conference • Statewide environmental education conference

Note: Mingé (2013) named these three categories as part of *Six Epistemological Lessons* that she learned from her family by doing her autoethnography titled, *Mindful Autoethnography, Local Knowledges: Lessons from Family.*

Weaving stories into narrative. Stories from my reflective journals were woven into an autoethnographic narrative titled "My Feet of Clay." I read through the journals five times for different purposes. In the first reading of my reflections, I labeled the excerpts from my journals into categories In the second reading, I uploaded my excerpts and coded them with a searchable database software program named ATLAS.ti. I searched by code, or code family, to gather excerpts and identify themes in the third reading. The purpose of the fourth reading was to identify relationships between topics, and compose short narratives. In the final reading of the excerpts from my journals, stories were woven into a comprehensive narrative about my lived experiences with assessment, equity, and environmental literacy in the fifth read.

First read: labeling, organization, and themes. In order to answer the guiding questions from my spiral notebooks, I converted them to electronic format and transcribed the handwritten text into tables. The tables categorized the data by date, and type of experience. I highlighted key quotes that would later serve to create codes. In order to make recommendations and conclusions, I specifically looked for missing feedback loops and antidotes to the archetypical mental models.

Primary and secondary labels. The purpose of the first read was to organize the data electronically in sequence, number each page, note the event, and associate key concepts. The artifacts were scanned into electronic format. The descriptions, key vocabulary, and quotes from handwritten notes were used to generate a list of primary labels to categorize written reflections about different kinds of events. Secondary labels were selected to categorize concepts related to the proposed guiding questions. The primary labels and secondary labels created from the first read are shown in Table 7. The primary labels played less of a role in organizing the data than the secondary labels that described key concepts from my experiences, reading, reflection, and interactions. I did not choose to weigh one type of data or experience over another, and intentionally valued my interactions with formal and non-formal educators equally. I found myself conforming to the norms of each group based on the type of nametags I wore.

Table 7

Primary and Secondary Labels from First Read of the Data

Primary Labels (Events, Artifacts, Locations)	Secondary Labels (Concepts)
Events: Workshop, Conference, Reading, Legislative Meeting, Political Rally, Classroom, Outdoor School **Type of Artifact:** Book, Video, Reflection **Location:** Inside, Outside, Field Study, Lab	**People Involved:** Formal, Non-Formal, Community, Family **Source:** Natural Patterns or Social Rules **Place:** Public, Private, Urban, Rural **Type of Assessment:** Gift, Individual, Group, Proficiency **Time:** Instructional, Paid, Volunteer, Citizen **Consensus:** Concord, Discord **Academics:** Engineering, Science Inquiry, Explore, Discover **Role:** Student, Teacher, Instructor, Expert, Scientist, Activist, Expediter, Beginner, Older, Younger, Staff, Leaders **Equity:** Justice, Mercy **Measurement:** Confidence, Resources, Knowledge, Skill **Context:** Student Choice, Personal Bias, Employer, Mission **Students Role:** Natural Systems, Community, Future

Thematic interpretation. As part of the first read, I did a first writing that focused on a description of cultural influences, while keenly aware of my mental models associated with settlers and land ownership. My intention was to understand the concept of equity by reviewing my family history. I learned how the roles that previous generations held in society influenced my choice of "the stable and able" plan, which meant working as a formal educator to inform my practice as a non-formal educator. I completed a "Culture-Gram" using Chang's (2008) worksheet to ascertain my interactions with social and cultural groups (p. 173). I purposefully sought out professional development opportunities to learn more about cross-cultural communication, equity, diversity, and race. Table 8 was used to summarize themes early in the data collection process.

Table 8

Early Themes

Language:	**Social Structure:**
Discovery, *exploration*, *claim*, *credit*, and *equity* are words that are often used in science research, which have cultural associations related to injustice, such as in situations where people settled and then used violence to protect their mental model of ownership.	Generational knowledge and epistemic trust can be found in natural systems and natural resource management based curriculum that often speaks to the Tragedy of the Commons. Science concepts play a role in defining the limits of natural systems and prescribe the limits for decisions. The definition of epistemic trust is: Trust in the authenticity and personal relevance of interpersonally transmitted knowledge. Epistemic trust enables social learning in an ever-changing social and cultural context and allows individuals to benefit from their (social) environment. (Fonagy & Allison, 2014; Fonagy, Luyten & Allison, 2014; Fonagy & Luyten, in press)
Economics:	**Assessment:**
The non-formal education field provides seasonal work with a message and bottom line determined by the funding organization. Formal education requires educators to refrain from indoctrination. Salary and benefits use hierarchy to enforce policy and law. The teacher serves as the *parent in loci* to act for the welfare of the child.	The violent language of assessment; e.g., "learning targets." Non-formal educators use group assessments. Formal educators use individual assessments.

Thematic relationships. While speaking at a systems thinking conference in 2015, Janice Jackson distributed a handout titled "The Cultural Iceberg." At the surface, culture involved traditional foods. Deep culture included "approaches to . . . raising children, decision-making and problem solving." I remembered a movie that I made with my father, titled *Living Along the Cowlitz,* which helped me understand how a small community of indigenous people, pioneers, families displaced by economic hardships and technological advances, loggers, farmers, migrant workers, and immigrants from Lebanon and the Netherlands accepted one another's gifts. I coded artifacts using ATLAS.ti so I could easily access the dates and notes needed to describe aspects of my deep culture. Codes included the following terms: *Diversity—Gender, Diversity—White, Equity—Self, Imminent Domain, Justice, Mental Model-Intercultural Sensitivity, Oppression, Platinum Rule, Political Structure Historical, Poverty, Social Self, Sovereignty, Stories, Straddling Two Worlds, Rural, Violence, Vision, Vocation.*

Additional coding. Additional codes were used to note concepts, or ideas, that could serve as antidotes my mental models. Concepts associated with constructing models were coded using the language of systems thinking (see Table 9) as described by Meadows (2008) in the book *Systems Thinking: A Primer.*

Table 9

Codes Related to Systems Thinking

Systems Thinking Codes
Experience: Event
Trend: Increasing, decreasing
Change
Structural Influence: policies, laws, physical structure
Relationships: Reinforcing, balancing
Mental Model: Assumption, belief, value
Cultural Meaning
Position Statement

Second read: coding with ATLAS.ti. The purpose of the second read was to preserve the integrity of my actual voice and perspective by transcribing the scanned journal entries from the table that I created in Microsoft Word to ATLAS.ti. ATLAS.ti was selected because of its power to organize quotes and comments by code. “In vivo” codes were taken directly from the text and added too (See Appendix J for complete list of in codes). Additional quotes and comments were added to provide more detail than was given in the first read. A total of 654 quotes were transcribed from 87 artifacts. At the end of the second read, 194 codes had been used. The codes were used to direct the third reading for themes. ATLAS.ti was used to analyze the words and metaphors associated with codes, particularly as they helped define and give voice to the meaning of assessment, environmental literacy, and instruction.

Third read: identification of central themes. During the third read, I identified an “exceptional occurrence” about a time when I had been assessed for environmental literacy outdoors (Chang, 2008, p. 131). I understood an exceptional occurrence to be a time where my mental model of how a particular social or ecological phenomena worked, changed. Since I was focusing my journaling around experience of assessment, environmental literacy, and equity, I expected to find an experience to jolt me from the familiar in an unfamiliar situation. As it turned out, it was the familiar experience of being assessed in an unusual way that became the seed for understanding my experience. I also found early assessment artifacts from that time in my life. By highlighting insights from the story, I identified five factors that related to the event. I described the major themes associated with each one, and listed the codes that I could use to sort corresponding quotes from my database in ATLAS.ti (see Table 10).

Table 10

Code Families by Theme

Theme	Description	Codes
Instructional Relationships	The purpose of the relationship between teacher and student, older and younger, instructor and peer mentor, peer mentor and middle school student, what Kimmerer (2013) names the "more than human world" and human, is what I lived as "kind regard" (p. 252). Educators are more than brokers of relationship between the human and "the more than human world." They disclose their own relationships with community and natural systems with a deep understanding that they are loved back without limit.	*Acceptance, Adaptation, Assigned, Belief, Challenges, Choice, Communication Skills, Concord, Cooperation, Cultural Competence, Cultural Identity, Cultural Influence—EE, Cultural Influence—School, Culture, Culture—Japan, Culture— Kalapuya, Culture Natural Resources, Dialogue, Environmental Identity, Experience, Explore, parent in loci, Instructors, Leaders, Mental Model—Learning, Sense of place, Sense of role, Social structures—Responsibility, Teacher, Teaching, Vision, Vocation, Youngers, Olders, Elders, Justice*
Deep Culture Decision-Making	Intergenerational knowledge and interaction transforms the language of crisis and panic into one of kind regard for curiosity and counterfactuals.	*Natural Systems, Systems Thinking—Archetypes (Tragedy of the Commons & Renewable Resource Stock and Flow Diagrams), Belief, Choice, Younger, Older, Elder, Parents, Cultural Identity, Cultural Competence, Cultural Influence, Culture, Curious, Dialogue, Experience, Family, Environmental identity, Sense of Place, Social, Social Structures, Mental Model—Culture, Mental Model—Sustainability, Instructors, Generational Knowledge, Stories, Straddling Two Worlds, Peer Mentors, Urban, Rural, Relationship*
Educator as Anthropologist	Educator as "anthropologist rather than missionary," which I learned as a student teacher from James Wallace, author of *Twins in a Two Room Schoolhouse,* published in 2012.	*Citizen Science, Outdoor School, Activist, Advocacy, Field Study, Instructional Time, Legislative, Missionary, Political Structure, State Board of Education, Structural Influence, Teaching*

Theme	Description	Codes
Student Sovereignty over Context	The context of the problem or opportunity, solution or possibility, and determination a system's boundaries, purpose, elements, interactions, causal loops, archetypes, and antidotes belongs to the student and is limited by their experience.	*Attention, Context, OELP, Outside, Student Voice, Problem Solving, Systems Thinking, Sense of Place, STEM*
Reciprocity	"Doing science with awe and humility is a powerful act of reciprocity with the more than human world" (Kimmerer, 2013, p. 252). Reciprocity between an educator and middle school aged students who do different kinds of assignments to receive different kinds of assessment.	*Acceptance, Adopt a Farmer, Assessment, Communication Skills, Commodification, Competition, Engineering, NGSS, Science Inquiry, Framework for Curriculum, Individual Success, Group Success, Knowledge—Group, Over Simplify, Peer Mentor, Reciprocity—Gift, Mental Model—Learning, Value, Students—Middle, Scientists*

Constructing the narrative. The short list of codes for each theme was used to "connect data fragments" and contextualize them into short narratives for analysis (Chang, 2008, p. 131). ATLAS.ti helped by organizing quotations with comments by code, and opening the source documents. Using ATLAS.ti, I grouped the codes into code families so I could associate groups and identify overlap. The quotation manager was helpful for reading the actual text in order to group short narratives together.

Fourth read: **Identifying short narratives and connections**. Ten strategies suggested by Chang (2008) were used to synthesize quotes and comments into short narratives (as shown in Table 11).

Table 11

Chang's Strategy and Matching Short Narratives

Strategy	**Title of Short Narrative**
Search for recurring topics, themes, and patterns	ongoing
Look for cultural themes	Cultural Iceberg
Identify exceptional occurrences	A Kind Regard for Not Knowing, Yet.
Analyze inclusion and omission	A Golden Man with an Ax
Connect the present with the past	Dung Beetles
Analyze relationships between self and others	Little Researchers
Compare yourself with other people's cases	Musk Ox Maneuvers
Contextualize broadly	Staybacks
Compare with social science constructs and ideas	Political Sandwich
Frame with theories	ongoing

Note: Chang's (2008) ten strategies listed in the left hand column, not only provided structure for the narrative, but also provided prompts and direction so I could keep writing (p. 131). The short narratives that are missing were noted as 'ongoing' since they will likely be written in the future reflections.

Connecting narratives. I used a linguistic approach called "Connectors," designed by Ryan and Bernard (1967), to make connections within themes and identify alleged causal relationships between themes for the final analysis. I organized words and phrases they recommended into a table and searched the quotes associated with each of them, using the "find" function in Word to find relationships within and between themes (see Table 12).

Table 12

Key Phrases Defining Relationships

Words that indicate relationships	Phrases describing kinds of relationships
because	attributes (e.g., X is Y)
since	contingencies (e.g., if X, then Y)
as a result	functions (e.g., X is a means of affecting Y)
if	spatial orientations (e.g., X is close to Y)
then	operational definitions (e.g., X is a tool for doing Y) or provenience (e.g., X is the source of Y)
rather than	examples (e.g., X is an instance of Y)
instead of	comparisons (e.g., X resembles Y)
is a	class inclusions (e.g., X is a member of class Y)
before	synonyms (e.g., X is equivalent to Y)
after	antonyms (e.g., X is the negation of Y)
next	circularity (e.g., X is defined as X)

Note: After using the 'Find' tool in the edit menu to locate these key words in several of the narratives, there were few matches. Ryan and Barnard (2003) cited Casagrande and Hale (1967) when describing the key words used to identify specific types of relationships in shown in this table.

At this point, I knew that the validity of any claims I made about correlation or causation using my experiences would be limited to my perspective. So, I identified quotations from authors who seemed to have stood near where I stood as I tried to make sense of environmental literacy and assessment. I used my findings to answer my guiding questions and propose antidotes to my archetypical mental model. My narrative represents only a snapshot of my decision-making process. Rather than recommending change, I changed and refined my mental models.

Fifth read: composing the narrative. I composed and reviewed the narrative, "My Feet of Clay," to settle my own internal disputes when trying to make decisions regarding the practices I use for teaching and assessment. In addition to deepening my understanding of environmental literacy, I learned to recommend small incremental changes repeatedly (iterating to optimize). But, I still sought to find antidotes to the limited perspectives used to inform my thinking and hope to continue a statewide conversation about the role of assessment for environmental literacy. Since I began my thinking using the archetype of Drifting Goals, my first analysis applied systems thinking to the problem. I re-evaluated my thinking by looking at the boundaries, elements, and interactions. I looked for causal loops to find missing balancing loops to the system. These missing balancing loops might support the assessment systems of formal and non-formal better than my original proposal of adding a systems thinking scoring guide to credit students for environmental literacy.

Formulating Claims

An assessment instrument itself must not be a barrier to students showing what they know and are able to do. It is assumed that the educators and students who participate in this study understand systems thinking and have English language skills that do not prevent an educator from understanding the student's work sample. Educators need to be able to provide both high-quality instruction and equal access to multiple assessment opportunities in a variety of formats to meet the needs of individual students. Hancock (2005) identifies threats related to construct validity when instruments are given to students whose first language is not that of the instrument. Careful consideration of the construct validity improves instruments that require narrative writing skills (Hancock, 2005, p. 608). By using student interviews in addition to written assessments, Doherty, Draney, and Anderson (2012) showed that written and verbal responses have similar means, although the range of verbal responses is much greater (presentation slide 23). Assessment instruments that can be used to guide an interview as well as write a narrative can share a reasonably similar level of validity.

Hancock (2005) suggested evaluating assessment instruments for students who are working on second-language development by asking how well the informative writing components are "situated with respect to their contexts and functions" (p. 607). Since the students are describing the context by using their understanding of interconnections between community, natural systems and the future, the educators learn the context from the students work sample. Since the scoring guide is developed by educators and students in Oregon, it is likely to serve the function of assessing each student for the OELP learning strand systems thinking than a nationally normed test used for the function of comparing students from one place with students from another.

Another limitation for making claims identified by Trochim (2006) was the "interaction of testing and treatment" (para. 6). He explained that the testing instrument is actually part of the treatment. Although it was not part of the design of this study, it was assumed that construct validity improves when students have multiple opportunities to show proficiency, or are asked to use the same instrument in different contexts. So, by focusing on increasing construct validity, the study created a type of instrument that was versatile enough to possibly be used in different settings, or to measure other strands of the OELP. Because percent agreement was used to measure "convergent validity," the inference was made that the two groups of educators shared a high degree of agreement regarding the "functioning and operating reality" of environmental literacy (Trochim, 2006, para. 1). In short, the evaluation criteria used to design a process for evaluating assessment instruments measured how well the Strand 1 of the 2013 OELP was translated for middle school students by educators working within different constraints.

Ethical Considerations

The use of human subjects required the use of informed consent forms for educators and students, which were distributed after receiving approval from the university and school district. Those who participated were given a letter that described their involvement in the research process and their ability to opt out at any point, along with contact information to make their wishes known (refer to Letters of Consent in Appendices K & L). Parents and students were asked to give consent for the use of their work samples. To protect students from unsolicited feedback about their work sample, and prevent the work sample from being used as a model of middle school students work outside of this context, the Letter of Consent for students read:

> You and your parents might be worried about posting your systems thinking project to a Google doc site where anyone who has the link can access it. Your work will only be available for the duration of the study and then removed. Although educators are asked to agree not to download, copy or share your project, I, as the researcher, cannot guarantee that educators will not violate their agreement. To safeguard against copyright infringement, you and your parents are encouraged to send a photo of the original work or an electronic copy in .pdf format. I will not publish a copy of your project as part of my report for the dissertation because my question is about whether educators can use the scoring guide, not how well students' projects score (for additional details refer to Appendix L).

Informed consent was not requested until after the Human Subject Research Review Committee and the school district had given approval.

To avoid potential risks to educators and students, and to safeguard their anonymity, their responses and work samples were coded. The scores assigned to students' work by educators were not traceable to an individual educator since they responded anonymously. Data was stored on an external drive and kept in the researcher's possession or stored in a locked cabinet. No subject was audio recorded or videotaped. The worksheets and responses of the forum, work samples, and scores will be stored until June 2017, when the external drive will be crushed and brought to the transfer station for burial in Eastern Oregon by a waste hauler.

Limitations and researcher bias. The study only partially responded to a call from Glenn (2000), who wrote the report titled *Environment-Based Education: Creating High Performance Schools and Students.* She writes, "More quantitative studies are needed, however, to convince doubtful teachers of the value of environmental-based learning" (p. 45). Her report, which was commissioned by the National Environmental Education and Training Foundation, iterated the methodology of seminal work by doing seven case studies in schools with environment-based education and focusing on academic scores. It has been 13 years since she wrote these words: "To make meaningful conclusions about the effects of environment-based learning on student achievement, we need many more studies, and quantitative ones that include analyses of the students and program characteristics associated with different types of learning outcomes" (p. 47). Demographic and educational program information was not added to the instrument to correlate students' level of environmental literacy with specific educational experiences. However, the experimental design of this investigation may support future research by increasing the sensitivity of a scoring guide for measuring systems thinking and improve the reliability of educators scoring students' work samples. To acknowledge my bias in moving the field in this direction, I created my own scoring based on my understanding of systems thinking, the Oregon Environmental Literacy Plan, and the Next Generation Science Standards prior to leading the Delphi with the middle school educators and students (see Appendix I).

Conclusion

The research design was structured to measure consensus. The level of importance assigned to each statement, amount of reliability from testing the scoring guide, and final comments for improvement revealed how to improve assessment instruments for environmental literacy. The evaluation process tested the notion that training educators to score work samples the same provided a more accurate method of assessment than providing a high-quality scoring guide to an educator that can be used to measure the proficiency of each student. By giving educators and students a voice in creating the assessment, formal and non-formal educators were given the opportunity to share the burden of measuring environmental literacy, and establish systems thinking as a distinct and measurable learning strand.

By using a common assessment instrument for environmental literacy, students were evaluated on their skills with environmental literacy as defined in the 2013 OELP rather than specific academic learning standards. Although the scoring guide included both adopted environmental literacy strands and adopted science standards, students explained their decisions to take action based on interactions between natural systems, communities, and the future. This study used a common assessment instrument which allowed formal and non-formal educators to report on how well they were able to use systems thinking skills as a measure of environmental literacy rather than whether students understood content specific to natural resource management or a particular academic subject. The scoring guide was as much about the educators' ability to use systems thinking skills as it was about the students'.

By describing interactions between one's self, community, natural systems, and the future, the students defined the context for their decision. They were not judged by the choices they made, but by how they had used systems thinking skills to describe how to leverage the system as they wished. They recognized how decisions they make now can influence the system in the future. The scoring guide directed educators' attention to specific characteristics of students' work. So, students benefited because educators had a common instrument with which to measure proficiency. Each student was scored in relation to the scoring guide rather than in relation to each other. Students also had the freedom to define the system as they understood and experienced it.

The scoring guide focused on the skills of systems thinking, so formal educators could directly assess for environmental literacy skills rather than academic content assigned to each discipline. It gave formal educators permission to make environmental literacy primary rather than secondary to academic learning. The scoring guide was used to evaluate students' work at one point in time rather than assigning the students the task of performing to the specifications of the scoring guide. It was created to improve construct validity so it could be used as a common measure for reporting back to the public if their organizations receive public funding. It was designed to support formal and non-formal educators who share a common understanding of how to measure environmental literacy in teaching the process skills of systems thinking to make decisions rather than teaching which decision is right or wrong for an environment or community. It did not require formal educators to find opposing perspectives and have them preapproved by principles or school boards, and non-formal educators were not suspect as environmentalists. A common assessment instrument for environmental literacy using systems thinking was worth investigating because it gave educators a common set of skills and tools for communication. The scoring guide identified skills that were important to middle school students, formal educators, and non-formal educators, and measured students' ability to make their own decisions in the community, natural systems, and future in which they live.

Summary

This chapter explained the design of this investigation to research the question: "What does the level of consensus between non-formal, and formal educators reveal about designing an instrument to measure a middle school student's level of environmental literacy in Oregon?" It explained how the voices of middle school students as well as formal and non-formal educators were heard, and how I was able to bring my own life experience to bear in understanding the social, cultural and political influences involved in assessing middle school students in Oregon for proficiency with systems thinking. In Chapter 4, a summary of the data will show: a high level of consensus for the construct of environmental literacy by the 11 middle school students, formal educators, and non-formal educators who created a scoring guide for systems thinking; a moderate level of inter-rater reliability between the 11 formal and 14 non-formal educators who field-tested the scoring guide using a sample of students' work; and what it is like to be responsible for teaching and assessing decision-making skills with middle school students.

CHAPTER 4

Analysis and Results

Based on the advantages of using mixed methods for studying the validity, reliability, and equity of assessment instruments described in the previous chapter, this section will provide a detailed description of the data. The purpose of this mixed methods study was to find commonalities in the evaluation of an assessment instrument for environmental literacy, using two groups of middle school educators—formal and non-formal. In Stage 1 of this study, middle school educators and students responded to an invitation to create a scoring guide that provided clarification of the construct of proficiency with systems thinking in environmental literacy. In Oregon, environmental literacy is defined as "an individual's understanding, skills and motivation to make responsible decisions that consider [one's] relationships to natural systems, communities and future generations" (OELP, 2013, p. 4). In Stage 2 of this study, the systems thinking scoring guide was field tested by a group of 25 people, 11 of whom self-identified as formal educators, and 14 non-formal educators. The educators were from 10 out of 36 different counties in Oregon. Using the scoring guide, each person independently rated a pre-selected sample of middle school students' work. Without any professional training, the group reached a moderate level of inter-rater reliability (k = 0.54) as measured by Cohen's kappa. An autoethnographic narrative titled "My Feet of Clay" provided social, cultural, and political context for the study. The stories supported two of what Trochim's (2006) Pattern Matching Theory would label "hunches": (1) non-formal educators are capable of creating valid and reliable assessments for environmental literacy that value middle school students' decision-making skills; and (2) decision-making skills are deep culture skills that non-formal educators are free to teach by giving voice to the hands and hearts that live with the consequences of inequity and environmental injustice (para. 2). Common understandings about using a scoring guide for systems thinking to measure a middle school student's environmental literacy were revealed by examining the following: (1) the level of consensus regarding the forum member's proposed criteria for a systems thinking scoring guide, (2) the level of inter-rater reliability between formal and non-formal educators field testing the scoring with an example of student work, and (3) personal observations and reflections that created "stories" from my journals concerning social, cultural, and political conditions related to assessment, equity, and environmental literacy in Oregon.

Stage 1: A Scoring Guide for Systems Thinking

The Delphi technique was used to build consensus between eleven people: two formal educators, three non-formal educators, and six students who had learned systems thinking skills from their teacher. The role of the researcher was to function as a conduit for communicating the ideas between members of the forum in such a way as to give each one an equal voice.

A high level of consensus. In the first round of the Delphi, each person responded to a survey which asked them to list the skills that a middle school student would need to demonstrate to proficiency with systems thinking. In the second round, the members of the forum of ranked the skills they had previously recommended. Through the survey, the forum established that consensus equaled 80% agreement. By the third round, the members of the environmental literacy forum agreed on four steps to guide students toward proficiency:

1. Modeling and Analysis
2. Systems Habits
3. Problem Solving
4. Refining and Proposing Changes

After rating their list of recommended skills twice, the forum agreed that 25 out of 52 systems thinking skills were *very important* or *quite important*. The forum used a median level of 80% consensus in order for a skill to be included. The skills were listed with each step, with the strongest level of consensus at the top of each column in the scoring guide, as shown in Appendix J. The first and only systems thinking skill to reach consensus in Round 2 was “identify short- and long-term consequences.” Each person’s voice, or expressed position, was represented in the final scoring guide. For example, each step included skills that had been suggested by a student, a formal educator, and a non-formal educator. The recommendations made by the members were distributed across the scoring guide. Rarely, similar statements were combined and reviewed by the forum for inclusion in the scoring guide together. The systems thinking scoring guide for Environmental Literacy Strand 1 of the Oregon Environmental Literacy Plan created by the forum was a compilation of the most important skills from the different scoring guides submitted by each member in Round 1 of the Delphi technique.

Students and educators prioritize decision-making skills. By using the Delphi technique to give middle school students equal voice as educators, a number of important decision-making skills were included in the scoring guide. The members of the forum agreed on three core skills that defined the construct of proficiency with systems thinking:

1. Making connections between the parts of the system and their outcomes, highlighting the interdependence of each part to make a whole.

2. Presenting the complex inner workings of a system in a simple and succinct way.
3. Creating solutions for systems that are not in balance or unsustainable.

The actions that the forum expected of middle school students transcended concerns tied to specific Oregon ecoregions. The Middle Years Programme named these types of actions “command terms” (International Baccalaureate Organization, 2016, p. 1). In the forum’s systems thinking scoring guide, proficiency meant demonstrating the ability to make connections, explain, identify, predict, suggest, create solutions, show, develop models, collaborate, present, display, track changes, and vocalize. The forum identified one particular skill that builds a bridge between students, formal educators, and non-formal educators, across academic disciplines, cultures, and ecoregions: “explore multiple solutions for the same problem.” The forum’s agreement matched the one key action Osborne’s (2007) recommended educators take to include all future citizens in decisions - give students the “opportunity to consider data which has no clear interpretation and to consider plural alternatives” (p. 179).

Systems thinking tools demonstrated in students’ work sample. The work sample provided by the students for field testing the scoring guide was actual evidence of their ability to use the tools of systems thinking with efficacy: Connection Circles, Causal Loop Diagrams, Behavior Over Time Graphs, and a Ladder of Inference. For example their Behavior Over Time Graphs (BOTGs) showed how particular elements in an ecosystem that supported salmon and big trees changed over time. They also used conventional notation such as dashed to extrapolate the data into the future. Arrows in the Connection Circle were draw from one element to another in the direction from cause to effect. An “S” was used to indicate that one element supported the other element in the system (direct relationship). An “O” was used to indicate that one had the opposite effect on the other (inverse relationship). The letter “B,” for balancing, was added in the middle of a causal loop that the students had identified in the Connection Circle. In total, they provided examples of two reinforcing loops and one balancing loop. They drew a Ladder of Inference that outlined the reasoning leading to their recommended action. At the base of the ladder, they identified an experience that gained their attention. Then, they listed which details they chose to notice, followed by the cultural and personal meaning they ascribed to what they noticed. Finally, they stated their beliefs in a way that made clear their reasons behind the action they would take, which was written at the top of the ladder. They summarized the system in two written paragraphs and identified two possible problems that could impact the system.

Systems thinking tools referenced in the scoring guide. Members of the forum were encouraged, but not required, to independently review the Environmental Literacy Learning Strands in the OELP, correlations between the OELP and Next Generation Science Standards shown in Appendix G, examples of students’ work with systems thinking, and systems thinking videos created by the Waters Foundation, available online at WebEd. The forum referenced four tools for modeling systems, which were explained by the Waters Foundation: Connection Circles, Iceberg Model, Behavior-Over-Time Graphs (BOTGs), and Stock and Flow Models. The scoring guide for systems thinking created by the forum included the following skill: “Use the Iceberg Model to show what people already know and the bigger picture of the system.” The Iceberg Model includes all the tools used by the students and all the tools identified by the forum (see Appendix A). The Waters Foundation’s Iceberg Model template takes students’ learning deeper than identifying patterns of behavior. It helps students "identify leverage points with greatest impact" for the purpose of "creating solutions for systems that are not in balance or unsustainable," which the forum agreed were important skills for proficiency. By choosing the Iceberg Model, the forum provided students with a critical tool they can use to explain "dynamic modeling with stocks and flows, and change variables until the desired outcome is achieved."

Inclusion of engineering design cycle. The forum's scoring guide for systems thinking asked student to use a pros and cons chart and the design cycle as tools for "exploring multiple solutions" and encouraging "creative thinking." The Oregon Department of Education's professional development for science teacher for engineering design was presented as a set of process skills to be assessed like science inquiry, where the context can vary. It was introduced to educators using projects that typically ask students to use science knowledge to optimize for a particular purpose based on criteria established by a client. They involved project-oriented, teacher-written scenarios, such as building wind turbines, trebuchets, and hot packs, that sometimes motivated students by including bonus points for low cost and efficiency. The Next Generation Science Standards for Middle School Engineering Design adopted by Oregon stated:

> Students who demonstrate understanding can:
>
> MS-ETS1-1. Define the criteria and constraints of a design problem with sufficient precision to ensure a successful solution, taking into account relevant scientific principles and potential impacts on people and the natural environment that may limit possible solutions.
>
> MS-ETS1-2. Evaluate competing design solutions using a systematic process to determine how well they meet the criteria and constraints of the problem.

> MS-ETS1-3. Analyze data from tests to determine similarities and differences among several design solutions to identify the best characteristics of each that can be combined into a new solution to better meet the criteria for success.
>
> MS-ETS1-4. Develop a model to generate data for iterative testing and modification of a proposed object, tool, or process such that an optimal design can be achieved. (NGSS Lead States, 2013, p. 2)

Comfort and familiarity with the design cycle as it is described in the NGSS was evident in the language used by members of the forum writing a scoring guide for systems thinking for environmental literacy.

Resonance with the Next Generation Science Standards for Modeling. The first step of systems thinking identified by the forum included modeling and analysis. The skills for modeling involved evaluating students' ability to "make a claim using evidence, and provide reasoning orally and in writing," as well as their ability to "construct an argument from analysis of data." A review of the Next Generation Science Standards (NGSS) shows that these phrases for "making a claim" appear in the standards for life science and earth science. In life science, the standards read, "Make and defend a claim based on evidence…" (NGSS Lead States, 2013, para. 1). In earth science, the standards read, "Make a claim about the merits of a design solution that reduces the impacts of…" (NGSS Lead States, 2013, para. 1). "Engaging in Argument from Evidence" is one of the Science and Engineering Practices in the NGSS. Another skill identified by the forum requires students to distinguish between "correlation and causation." Identifying "cause and effect relationships" is also one of the Crosscutting Concepts outlined in NGSS. "Developing and Using Models" and "Planning and Carrying Out Investigations" are key Science and Engineering Practices used to define performance expectations in the NGSS.

Roger Bybee (personal communication, 2012) taught me the importance of giving students ownership over the model rather than limiting them to replicating other people's models. Bybee also emphasized gauging students' understanding by providing them with multiple opportunities to improve upon previous models. The systems thinking tools explained by the Waters Foundation made it possible to field test the scoring guide with an example of students' work. Modeling with causal loop diagrams, and refining their models were identified as important skills by the forum that created the systems thinking scoring guide.

Dissonance. A major difference between the forum's systems thinking scoring guide and the Next Generations Science Standards was found in the performance expectations for MS-ESS3 Earth and Human Activity: "science does not make the decisions for the actions society takes" (NGSS Lead States, 2013, p.1). The systems thinking scoring guide for environmental literacy evaluated whether middle students can use their understanding of natural systems to suggest solutions "for systems that are not in balance or sustainable." To be scored as proficient in refining and proposing changes, students were asked to "use visual graphic skills to clearly present how changes affected the environment." They were asked to present sound models and suggest corrective actions. An inference was that students would demonstrate their systems thinking skills while attending to the precautionary principle of environmental science or the constraints that natural systems place on engineering design. Kriebel et al. (2001) noted two central aspect of the precautionary principle involved "taking preventative action in the face of uncertainty" and "exploring a range of alternatives to possible harmful actions" (p. 1). Since the authors of the scoring guide used the past tense, one can assume that the student acted on their proposed changes, "tracked change over time," and would continue to make "small, gradual changes over time." In terms of the scoring guide, environmental literacy differs from the Next Generation Science Standards because middle school students were considered part of natural systems and communities, which included the consequences of the decisions they make.

Stage 2: Field Test of Scoring Guide

Reliability of scores assigned by formal and non-formal educators. The purpose of the field test was to measure validity of the scoring guide. It also examined the

Figure 8. Field Test Responses by Oregon County

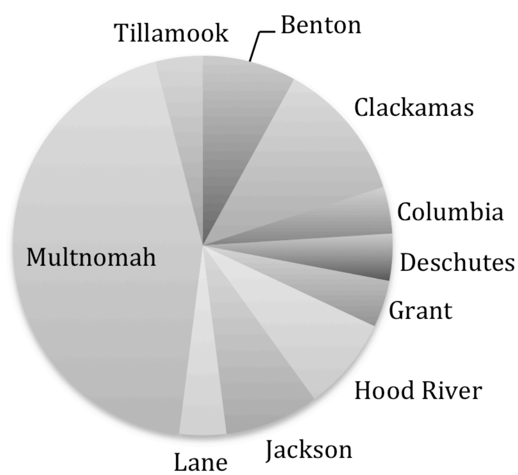


Figure 8. Graph showing responses from educators by Oregon county. The percentage of the total represented by the number of respondents is shown in Table 13.

amount of reliability between the scores of formal and non-formal educators reviewing the same example of students' work. Reliability between the formal and non-formal educators was measured using percent agreement and Cohen's kappa. Cohen kappa (k = 0.54) indicated a moderate level of agreement for the construct of proficiency with systems thinking as indicated in the scoring guide. Percent agreement was used to determine the validity of each statement. A higher level of agreement supported the premise that formal and non-formal educators shared an understanding of how middle school students demonstrated proficiency with systems thinking. Educators provided written feedback as well. The number of educators that participated in the field test included 11 formal educators and 14 non-formal educators from counties across Oregon (see Figure 8). When compared to the percent of population by county from the 2012 census, the proportion of responses was fairly representative (see Table 13).

Table 13

Field Test Responses by Oregon County

	% of Total Responses	% of Total Population (2012 Census)
Benton	8%	4%
Clackamas	12%	19%
Columbia	4%	2%
Deschutes	4%	8%
Grant	4%	0%
Hood River	8%	1%
Jackson	8%	10%
Lane	4%	17%
Multnomah	44%	37%
Tillamook	4%	1%

However, the limited size of the sample prevents any generalizations to the overall population, and is not reflective of the number of formal and non-formal educators in each county. Respondents received the invitation to participate in the field test through electronic correspondence from seven organizations supporting over 1,000 educators in science education, environmental education, and systems thinking in schools. The scores assigned by respondents were used to calculate statistics to test the null hypothesis that there was no significant difference in scores between formal and non-formal educators.

They are just as likely to agree on a particular score due to chance as it might be to any reliability in scoring. However, the actual sample size is not representative of the number of educators in Oregon and cannot be used to make generalizations.

Percent agreement. Percent agreement was used to determine the general level of agreement between formal and non-formal educators who rated the work with the following scores: 1 = No evidence; 2 = Not there yet; 3 = Proficient; and, 4 = Highly proficient (meaning, it appears they could teach the skill to another). Percent agreement does not account for chance, so it could not be used to prove or disprove the hypothesis. However, the percent agreement between the median of the scores assigned to the students' work by formal educators and the median of the scores assigned by non-formal educators was 76%. Educators rated the students' work on 25 skills, and the medians did not exactly agree for six skills. Disagreement did not exceed one level, but can be used to infer that the educators needed more information about the construct of proficiency with systems thinking in order to feel more confident with scoring students' work.

Further clarification of the construct. The six skills that formal and non-formal educators scored differently were shown in Table 14.

Table 14

Difference in Formal and non-formal Educator's Scores for Specific Skills

Systems Thinking Skills from Scoring Guide	**Difference in Scores**
The student work identifies long and short-term consequences.	1
The student work shows how a system's structure generates its behavior.	1
The student work identifies the problem of a situation.	0.5
The student work displays proposed changes and outcomes via easily understood diagrams. It uses visual graphic skills to clearly present how the changes effected the environment.	1

Table 14 (continued)

Difference in Formal and non-formal Educator's Scores for Specific Skills

Systems Thinking Skills from Scoring Guide	Difference in Scores
The student work analyzes data.	0.5
The student work shows evidence of collaboration. It uses the design cycle to explore multiple solutions for the same problem, and creates a +/- chart for each solution.	1

The overall score formal educators assigned to the students' work was .5 lower than the overall score assigned by non-formal educators. To determine if this difference was significant, and not due to chance, Cohen's kappa was calculated.

Cohen's kappa. According to C. Zaiontz (personal communication, Mar. 22, 2016), the kappa statistic can be used to remove the level of agreement between two raters due to chance if two conditions are met: (1) there are exactly two raters, or, as in this case, two groups of raters, and (2) each rater judged all the subjects. The number of times the group of formal and non-formal educators agreed and disagreed was used to

Table 15

Cohen's Kappa Comparing Reliability of Scores Between Formal and non-formal Educators

Cohen's kappa						
		Formal				
		No evidence	Not Yet	Proficient	Total	
Non-formal	No Evidence	0	1	0	1	4%
	Not Yet	0	9	3	12	48%
	Proficient	0	2	10	12	48%
	Total	0	12	13	25	
		0%	48%	52%		

Note: C. Zaiontz (personal communication, Mar. 22, 2016), explained that Cohen's kappa (*k)* is equal to the proportion of scores in agreement ($Pr_{(a)} = .76$) minus the proportion of scores due to chance ($Pr_{(e)} = .48$) divided by one minus the proportion of scores due to chance. In this case, Cohen's kappa for the scores assigned to the students' work by formal and non-formal educators was *k* = 0.54. A kappa value of 1 indicates perfect agreement, and a value of zero can be interpreted as no agreement. The level of agreement between formal and non-formal educators, accounting for chance, was interpreted as moderate, based on the following rating systems presented by Zaiontz: 0%–20% poor, 20%–40% fair, 40%–60% moderate, 60%–80% good, 80%–100% very good.

calculate the kappa statistic as shown in Table 15. The scoring guide was used by the respondents with a moderate level of reliability, according to Cohen's kappa, which is used to estimate the level of agreement that is not due to chance. The null hypothesis was that there is no difference between the reliability of formal and non-formal educators using the scoring guide, which is expressed as:

H_1: Cohen's kappa = 0 or < 0

H_0: Cohen's kappa is between 0 and 1

Cohen's kappa was calculated to be 0.54, which proves the null hypothesis. There was no significant difference in the reliability of scores between formal and non-formal educators.

Questions raised by educators' feedback. Specific feedback, from those who scored the sample of students' work showing the interrelationships between salmon, bear, insect larvae, and huge trees, identified ways to make the scoring guide for systems thinking easier to use (see Table 16). As part of the agreement with the students and parents, the students' work was not include in the results nor published. It was only to be used to field test the scoring guide. This was a purposeful choice so that no one piece of students' work would be used as an exemplar, and every student would be valued for their efforts. It might be helpful to note that the students who created the work simply created the work without being assigned to use the scoring guide. They presented their work to an audience of that included students and educators. The written paragraph that the students created for this study was an attempt to summarize the presentation. The feedback from the field test uncovered commonalities in the assumptions and questions educators have about assessment for environmental literacy.

Table 16

Feedback on Systems Thinking Scoring Guide from Field Test

Feedback on Systems Thinking Scoring Guide from Field Test with 25 Formal and Non-formal Educators
1. Are the students going outside and doing any field work?
2. Will requiring students to show their ideas using the Iceberg Model limit their creativity and how they express themselves?
3. Will those who struggle academically, or who are in the process of learning the English language, be able to perform at the level of the students who provided the work sample?
4. Is it possible to include the wide range of skills outlined in the scoring guide in one project?
5. Must the presentation be verbal, visual, or both?
6. Why are there no levels of proficiency?
7. How might my scoring change if I had been able to compare a number of students to each other?
8. How can this tool be made simple without over simplifying it?
9. How can the scoring guide be used if the students do not identify the problem?
10. How can I score without seeing them in action?

Introduction to Autoethnographic Narrative: *My Feet of Clay*

The function of my autoethnography is to be mindful of inequities and assumptions that affect my ability to provide a fair assessment of a middle school students' environmental literacy, specifically the use of the systems thinking skills outlined in the Oregon Environmental Literacy Plan. It is the compilation of my observations and reflections on serving as public middle school science teacher and non-formal environmental educator in Oregon, Washington, and Minnesota since 1988. It is important to recognize that my understanding of scientific principles, and my curiosity for the "living landscape" have influenced my working definitions of assessment, equity, and environmental literacy over time (Oregon Biodiversity Project, 1998, front cover).

The narrative begins to reveal antidotes to three mental models related to assessment: (1) Competition, (2) Completion, and (3) Correction. I describe a key experience where I was assessed with my field journal about a swath of woods. The quality of my fieldwork was measured against my own curiosity rather than the teachers'. My unyielding effort to protect the local woods was valued over the assignment requirements, and my teacher showed a kind regard for my not knowing all the names of all the plants. My stories take a long, deep dive into the influence my family's history has on my sense of equity. I follow my immigrant grandmother and pioneering grandfather to the edge of the Cowlitz river, which was dammed to power the City of Tacoma. My definitions of environmental literacy are provided through stories of my interactions with young people, where I actively participated in the process of giving cultural and personal meaning to experiences outdoors. I end by opening the doors of my classroom, and do all I can to help people experience counterfactual, the possibilities that I can imagine, with the middle school students I learn with each day.

The narrative is organized around five themes:

1. Instructional relationships
2. Deep culture decision-making
3. Educator as anthropologist
4. Student sovereignty over context
5. Counterfactual futures

These themes five themes emerged for my autoethnography after deep reflection and ongoing journaling beginning in 2012 through 2015 with two respites for cancer treatment. These respites provided opportunities to volunteer and heal with the empathy I experienced from family, friends, neighbors, and other caring individuals like those who donate blood. I am now grateful for the kind hearts, hands, and heads of generous medical scientists, donors, healers, and loved ones walking on two feet.

My Feet of Clay

Though I am an educator familiar with the working end of a wheelbarrow, parent-conferences, fixing food for 15 trail-weary teens, restoring streams for salmon, the Code of Conduct concerning controversial curriculum, and intercultural communication, I will always be learning the Platinum Rule: “Treat others as they wish to be treated.” I need to practice respecting the members of the many natural systems and communities that sustain me more than I know. One of my middle school teachers who was a strong advocate for integrating schools so students learned with people of different races, accused me of having clay feet. After years of reflection, I think I am stumbling toward a deeper understanding of what she meant. I used to use my interpretive skills to explain the chaotic dreams of those with power, like Daniel did for King Nebuchadnezzar. He explained the king’s dream of a statue made with a head of gold, chest of silver, and legs of bronze. Each different element represented another kingdom, all of which would fall when a boulder rolled down the mountain into the statue’s clay feet. The king rewarded Daniel for predicting the future. Understanding his dream gave King Nebuchadnezzar time to store food so people would ally with him when the impending crisis came, but he also threw people into a furnace for refusing to worship a golden statue of his design. The king had clay feet, too. Like Daniel, I warned future leaders that fluctuations in global climate would come due to increases in carbon dioxide and methane levels. I countered with the loss of vegetative growth due to global cooling caused by airborne particulates preventing sunlight from reaching earth's surface. The contradicting data felt like boulders, and it was hard to know what was real. But now I look back with chagrin realizing what The Oregon Biodiversity Project (1998) suggests so eloquently, "many of the conservation challenges the state faces can best be addressed through cooperative, non-adversarial efforts (p. 1).

I once met with an advertising executive sitting on a curb downtown that resulted in the donation of billboard space for a decade's worth of children's art. Children were celebrated for creating art with a message that would prevent waste. Now, I teach with a student who earned academic credit by performing in a government-funded production of the anti-consumerism play, *Barbie, Get Real!* Jennifer Gailus and Olivia Martin wrote the play when they were in high school in 1996 after visiting families with fewer resources half a world away from theirs. I remember the mixed feedback I received when we performed in a less-resourced, predominantly African American middle school. I was suddenly distressed to realize that I had asked those who did not have enough to want even less, and the play's message unintentionally crushed the hope that each person would have enough, but not too much. I learned that the solution to consumerism was systemic, not simply personal. Humbled, I remembered my uncle's admonishment to learn to carry my own shoes when all I wanted to do was run free through the surf. He told me, "I felt sorry for myself because I had no shoes, until I met a man who had no feet." Barbie and I both had feet of clay.

Instructional relationships. I have managed my mental model of success and failure only through the kindness of my classmates, parents, and teachers. My memories of those moments begin at age five in a school uniform with a nametag sewn into my sweater. When it was my turn to hold the flag and lead the pledge in front of the entire school, I forgot the words. Everyone said them anyway. When I left home to catch the school bus, unable to hold back the tears because I didn't know what would be on the test, my mother would tell me, "If you are going to fail, fail the best you can!" I failed well as a student, and I failed even more as an educator. My teachers' grades did not say so, and my students might have told you different, but I did. Student and teacher; instructor and mentor; interpretive naturalist and visitor; backcountry guide and novice mountaineer; mass and the angle of repose; older and younger sibling; hunger and huckleberries—I learned as one how to interact with the other. The function of my instructional relationships, in community and natural systems, stems from a disposition of curiosity for a planet, and a parent's love. One is loved back without limit. Anyone who has ever shared mangos or huckleberries has learned how much photosynthesis, the process that plants use to transform sunlight into food, loves us back.

In her story of a relationship with a teacher in *The Places That Scare You,* Pema Chödrön (2002) described her experience of "limitless love":

> This unconditional commitment to our selves, and to others, is what is meant by limitless love. The teacher's love for the student manifests as compassion. The student's love of the teacher is devotion. This mutual warmth, this heart connection, allows for a meeting of minds. It's this kind of love . . . [that] inspires us to step out fearlessly and start exploring the phenomenal world. (Track 21)

Devotion taught me the value of not knowing, and reigned in the imagined possibilities of my persistent curiosity. Greer, Mukhopadhyay, and Roth (2012) write, "Education is, fundamentally, about interpersonal relations between students and teachers" (p. 7). When being lost in the woods is my teacher, I am tested on whether what I hear is the wind, a waterfall, or the whoosh of cars, and I realize that the direction I choose to walk will have short and long term consequences. Chödrön reminded me of a teacher who taught me to cross boundaries created by political and scientific perspectives and encouraged me to notice seasonal patterns.

A kind regard for not knowing yet. My teacher rode his bike to school every day, even in the rain. He assigned me the tasks of written and oral response to sections of Commoner's (1979) *The Politics of Energy,* excerpts from Ehrlich's *The Population Bomb* (1968), Leopold's *Sand County Almanac* first published in 1949*,* and a plot of woods between my elementary and high school. He led daily Socratic seminars. The 15 of us sat in a circle made by tables in a chilly classroom with the rainy stench of our grass-covered tennis shoes. Our class was a locker room of young academics in the making. I wore jeans and a hooded sweatshirt kept warm by my internal radiator of panic. The panic came from not knowing, and rarely being able to form a coherent answer for any of his probing questions. Moss was my normal. My curiosity was secondary to safety, but my attachment to the birds near the creek and woods as *parents in loci* was not everyone's normal.

According to Gopnik (2010), who wrote *The Philosophical Baby: What Children's Minds Tell Us About Truth, Love and the Meaning of Life,* my brain was behind schedule. Her understanding of cognitive science suggests that, by 17, I most assuredly had an "unconscious causal map . . . of the way the world works," and I could use it to "act on possibilities after making predictions" (p. 38). Had I developed on time, I would have easily responded to a high school teacher's higher-level questions by imagining the counterfactuals, the "things that didn't happen" (Gopnik, 2010, p. 23). What she might have wanted me to realize back then is that I was just as capable of working on the world as it was of working on me:

> The evolutionary answer is that counterfactuals let us change the future. Because we can consider the alternative ways the world might be, we actually act on the world and intervene to turn it into one or the other of those possibilities. Whenever we act, even in a small way, we are changing the course of history. (Gopnik, 2010, p. 23)

For two years, my teacher's efforts earned my devotion, which became evident each time he communicated with my parents regarding my progress. It was remarkable to me that my attitudes were evaluated along with my skills and knowledge:

> In discussion Susan was sometimes shy, but her interest was so great that when she had a good idea she would blurt it out almost despite herself. There were moments when her synthesizing powers were strong, and she was an attentive student whose comments contributed to the discussion. . . . She perceived that the problems we studied in early American History had their counterparts today. After writing a paper on the violation of minority rights and how well minorities resisted these violations, she attended public hearings . . . where she saw the dilemma of those few citizens whose houses would have to be moved (or so it was claimed) in order to provide for the convenience of the majority.

A year later, he wrote to my parents:

> Her final paper achieved a rare blend of rational thought with real emotion. When discussing her argument for managing the school woods for the educational and aesthetic values. She urged that one should consider the value of the feelings created by the life on my plot as an asset to learning. After all, the school symbol is a tree within a circle. If they decide to cut down the trees, they are weakening their values. If they disturb the natural cycles of the plants and animals, their symbolic circle is broken. When Leopold's land ethic is applied to my plot, 'to preserve the integrity, stability and beauty of the biotic community' it allows the individual plants to grow together to create a biotic community in a school community, which is also dependent on the growth of the individual.

What he failed to capture in his written report was what he observed about my not knowing. When it was my turn to show him my field notebook and interpret my plot of woods, I did not know the genus and species of every plant. I had struggled for weeks with a vine that had two shapes of leaves that kept changing. When he asked me to explain, I simply shared my observations over time. That is when my teacher saw me, and applauded my not knowing, and my curiosity, as important. It was the kind of curiosity evident of a "free-range, latchkey, couch-surfing, runaway kid" before teachers used those labels to describe students like me. The relationship that I developed with the forest was serving me well as I responded to my high school teacher's questions. No doubt he had walked that path, too. The forest was obviously teaching me along with the text, and he valued my ability to speak for it when others imagined a school without it. The forest was a tool for my education and provided a sense of community for me. Because of his wise assessment, I never felt alone in my not knowing.

I wonder if Kimmerer (2013), who wrote *Braiding Sweetgrass: Indigenous Wisdom, Scientific Knowledge and the Teachings of Plants,* would have rated me any differently than Gopnik on my reading of the woods:

> Science and traditional knowledge may ask different questions and speak different languages, but they may converge when both truly listen to plants. . . . Sustainable harvesting can be the way we treat a plant with respect, by respectfully receiving its gift. (p. 165)

The vine that grew different shaped leaves at differ times with the seasons taught me that science requires more than categorizing and naming in order to have a correct answer. My teacher respectfully received my gifts: written papers and verbal attempts at reasoning. I not only showed up to class on time, but listened as well as a saturated student can. I wrote my final paper in the sun in the field by those woods without a sense of crisis, even as the volcano in my grandmother's backyard was covering the landscape with ash. The field notebook he requested was filled with more illustrations and questions than Latin names and answers, which testified to my 12-year attachment to a temperate forest I knew as the woods between school and home. Had it not been for his compassion the persistence of my curiosity would have been my undoing.

The kind regard my teacher and I exchanged with natural systems and the local community is the counterfactual, the possibility from my middle years, that I now imagine we could indemnify. Imagine a tool that could measure a student's decision-making skills as a memorable conversation with an educator and a temperate forest, coastal estuary, or oak savannah. An assessment supported with illustrations showing change over time in a middle school-aged academic's soggy field notebook. Extrapolating from Gopnik's (2010) premise—"Counterfactuals let us change the future"—middle school age students, who can imagine how the context in which they live could be different, can act on the world to make those possibilities true (p. 23). As a soil instructor for Outdoor School, I noticed the impact that sleeping and feeding people from different schools together had on day-to-day decisions. By the end of a week, individuals were more than their parents' hearts walking on two feet, and naturally chose to reach out to one another with compassion.

Deep culture decision-making. I have a history of making destabilizing choices based on three mental models: settler, immigrant, and the Law of Eminent Domain, where the public good outweighs individual property rights. My vocational choices are driven by economic gains associated with credits earned for general education and professional development, combined with lived experience and common sense that my mother refers to as "the stable and able plan." I am not an expert, just curious and persistent. While my mother celebrates her Dutch ancestry, my father enjoyed his

Scottish heritage. I identify with two cultures—settler and immigrant—yet remain grateful for a third culture: indigenous.

An intergenerational indigenous basket. Starting in childhood, my father and grandmother made me the keeper of the family stories. My dad gave me a basket that he had treasured. It filled him with memories of his grandmother, Nellie. Her husband was the postmaster who ran the store. I know this because Dad kept a little milk bottle that held a receipt for bread, signed by my great-grandfather as a promise of payment, or credit, to my other great-grandfather. The basket reminds me of the afternoons my grandmother and father spent driving me across the county, showing me places like the storage garage that was once Nellie's one-room schoolhouse. The story of how the basket came to the family is one of relationship whose details I do not know. However, it is old enough to speak for itself, the person who made it, and the one who gave it to my family. I know I've changed since the basket came to me. I am less comfortable wearing the official nametags and logos that symbolize the mental models of governments and organizations. I am more eager than ever to spend time by the river with my back against a cedar. Maybe it would serve others for me to try to swim upstream, but I am resting now, here with this basket, listening to the past in the present.

My father's culture is celebrative of settlers who came to the Pacific Northwest, after Lewis and Clark's Voyage of Discovery, hell-bent on discovering gold and claiming land. Decisions were based on paternalistic, Christian assumptions that focused on sacrificial giving rather. Sacrificial giving was the kind of giving that risked a human life to obtain natural resources. The number of people who died from sawing cedars and firs into lumber seemed to be of less concern than profits. The white Bible, in the cedar box branded with the logo of the forest product company, did not bring an uncle, a father, or his lost income, back to my hungry, bear-killing cousin and his sisters. Balance would have respected the forest's regeneration rate over an ever-increasing population of human labor for harvest. I met one company that planted cedars two and a half human generations ahead, reflecting the investment the ancient cedars had made in their family's company. Another cousin confessed that changes in Forest Practices Laws were giving value to what he was taught to call trash trees. The salmon streams had made fine skid roads, and immigrants were often hired to place the chokers used to pull the trees, turn logs, and float them in rafts to the mill. I was surprised when he apologized to me for the logging company's behavior rather than making amends by planting a forest in his backyard like uncle.

My mother's culture is one of immigrants, characterized by a handful of traditional Fryslân values. For example, it is preferable to avoid too much "drukte," which is the word my mother taught me to use for "too much trouble to make sense of, not knowing what to do, or overwhelming." Advice included, "Go where the work is." Temporarily believing one's self to be free of a family's limitations, my great-grandmother carried her baby daughter, and the "drukte" that comes from borrowing too much without the means to pay it back, below the deck of a ship that left for America just weeks after the *Titanic* sank. She could not have imagined a future where her little girl's home, and the homes of people from the Indian Nation called the Cowlitz, would be flooded by the damming of a river to generate power for the City of Tacoma.

I grew up unaware that I used phrases from Chinook jargon and Dutch, as a natural part of my thinking. I used words like "skookum" to indicate "everything was in its place just right," or asked for a "skootleduke" to wash the dishes. On school field trips to the longhouse, elders from Chief Lalooska's family told us stories of how Raven posed as the beloved child of his grandfather so he could release the sun, moon, and stars from the cedar boxes in which they were kept. I ate salmon and picked berries, mostly local farmers' strawberries, grandmothers' raspberries, or the blackberries that invaded our backyards. I was not born the race of the indigenous people, whose government built the Princess Kapiolani maternity hospital in Hawaii where I was born. I will always be grateful that my dad who was a Navy doctor returned with other would-be fathers from a mission to Cuba, which was resolved without using nuclear weapons. My birth certificate has a picture of the princess on it, along with the motto "To build and increase the race." Mom did that by using hypnosis to manage the pain I caused at birth, while my father took advantage of the opportunity to learn more gynecological science.

The "stable and able" plan. "Nurse, teacher, or secretary?" my mom pleaded with me. "I don't take orders from people very well, so secretary is out," I said as I turned to see her face break from motherly concern to uproarious laughter. What I really needed was a vocation that allowed me to interact with others in the outdoors. "Teacher," I finally responded. I almost chose nursing, but I remembered the havoc my tears created for teachers the day we lined up to get vaccinations. Dad and I had to rehearse for weeks before I could do my part to prevent the spread of disease. In the plan, she paid my tuition for a teaching license from her own work as a mental health nurse just as her mother had paid for her nursing license by washing floors.

In the first pages of their book *A People's Curriculum for the Earth*, Bigelow and Swinehart (2014) get right to the heart of the matter: "Our curriculum must confront the false dichotomy between the environment and people" (p. xi). "Yes, finally," I think. "Two seasoned teachers from schools in my little corner of the planet have announced and published what I know to be true. I am not separate from the natural systems that sustain us, and these natural systems respond to our choices." But what about the moment my feet hit the floor, when I turned myself over to what my mother referred to as the "stable and able" plan. In this plan, I walked into a classroom with 36 middle school students, and spent 12,250 instructional minutes loving kids and science trapped by four walls. So, the stability of teaching would have to have my back while through the window the seasons changed.

A golden man with an ax. What I call a mental model of controversy and crisis is characterized by phrases like "two Oregons," which is made in reference to urban and rural communities that actually depend on the flow of resources between one another. I wonder what the man with the golden ax on top of the state capitol building was thinking. For example, why did he choose cutting trees and replanting fast growing trees, rather than tending to the systems dynamics of cedars. I wonder if recent legislation will bring the same kind of unintended consequences that national funds for vocational education brought in the early 1900s. Will it create an inequitable system that hires general educators to train citizens to test scientific theories, and vocational educators to train tree farmers? Systems thinking and the emphasis on modeling in the Next Generation Science Standards have commonalities that might find antidotes to archetypically mental models for environmental literacy. I hear phrases that begin with possibility, like "Yes, and . . ." and "Not there, yet."

In the past, federal legislation provided national funds for vocational education that brought unintended consequences for people oppressed by Euro-Americans. In the early 1900s, Booker T. Washington and W. E. B. Du Bois debated progress in terms of people supporting the economies of states. Du Bois debated the best way to train students in schools, whether urban or rural. Du Bois (1903) wrote:

> We need to-day more than ever, the training of deft hands, quick eyes and ears, and above all the broader, deeper, higher culture of gifted minds and pure hearts (p. 30).

Washington (1901) described the power of vocational talent when he said:

> The product of field, of forest, of mine, of factory, letter, and art, much good will come, yet far above and beyond all material benefits will be that higher good. . . . In blotting out sectional differences and racial animosities and suspicions, in a determination to administer absolute justice, in willing obedience among all. (p. 134)

Using technologies for more than extending our senses and capabilities, and copyrighting information that is necessary for the public good, can give some parts of a systems an advantage, but other elements may no longer be able to function. For example, Meadows (2008) uses the example of extending the reach of fishing industry by making it easier for a few lucky boats to catch the few fish that are left with state of the art nets, but the flow of resources from a renewable stock may not be able to recover (p. 67). However, Du Bois justly spoke to the continued oppression of people by others making decisions based on the belief that technological advancement is itself the goal rather than a means for reaching the higher good.

“Blaming, disciplining, firing, twisting policy levers harder, hoping for a more favorable sequence of events, tinkering at the margins,” Meadows (2008) suggests, “does not work” (p. 112). The criteria for building technologies is time bound by natural systems that create equity for those who return to those natural systems for their needs. The undersides of newly melting glaciers were once the fusion of ancient snowflakes. I teach science because I know the conceptual understanding of natural phenomena is put together anew in each child's mind. It is each one's unique combining of ideas that creates new possibilities while sustaining the old. Meadows would not have subscribed to my Velcro theory of learning, which states: "Do enough science labs and activities that keep kids engaged, and some concepts are likely to stick with them for when they need them to make decisions in the future." Meadows (2008) says of bounded rationality: "It's possible that you could retain your memory of how things look from another angle, that you burst forth with innovations that transform the system, but it is distinctly unlikely" (p. 108). She's right. I have been accused of not seeing the forest for the trees, too, much like a golden man with an ax chopping.

I notice the violence of the vocabulary for assessment—e.g., "learning target"—used to describe an arsenal of scientific knowledge and skills used to attack problems. Only three generations back, the U.S. government sent guns to help my great-grandfather protect Boistfort School District #1 of the Washington Territory in a place by a river that sustained the First Foods of one of the First Nation called the Cowlitz. In the next generation, these Cowlitz would lose their first lawsuit against the U.S. government that also inundated my mother's parents' land for public power, I mean, electricity. When the Cowlitz sued again, after the dam was built, the win returned 50 cents per acre, which is not really winning. The dam is still there. Like returning the faces of four founding fathers in the Black Hills, returning the land to the indigenous people would have be justice (Dunbar-Ortiz, Book Talk, 2015). In spite of Meadows sense that it is unlikely for people, including middle school students, to burst forth could burst with innovations to transform these kind of unjust systems, I believe that collectively they might *have a shot at it.*

During my mini lectures, students beg me to go slower so they can *capture* all the words on paper, or take pictures of the screen with their phones so they *have* knowledge. I take a breath and ask them just to stick with me until they can see the bigger picture. What are typically referred to as “bullet point” to denote subtopics, I call “peace dots.” I feel *threatened* by an administrator entering the classroom to ask any student to recite which “target” is the focus of the current moment. Each moment is a small but significant portion of the 11,100 hours of instructional time a child in Oregon’s K–12 schooling is *assigned* to me. I realize how important it is to reteach the 1% missed on a test where the student scored 99%. The purpose of my teaching practice is to help middle school students realize they know science concepts so they can use them to both solve problems and create possibilities. Teaching gives hope for the future, but I realize I need to quickly give students a chance to influence our now. I think like a Marine, *leaving no child behind!* The 1% matters. Arun Gandhi shared his story with me about the day he asked his grandfather for a new pencil. He had thrown the small stub he had left into the woods on the way home from school. His grandfather gave him a flashlight and told him to go find it. Our grandparents hold us accountable for a future they will never see, and they seem to have mastered ways of thinking that free us. They have lived enough to prevent problems in natural systems and communities so we don’t need to go back and look for our resources out of lack of respect, in a time of crisis, with a flashlight.

Educator as anthropologist. As a student teacher, my professor taught me to be an "anthropologist rather than a missionary." He made room for questions in class seminars so I could make sense of the social structures of schooling that science curriculum guides rarely addressed. I think it was his laughter in my unpredictability that still makes me appreciate the summers of student teaching classes when I run into him now at the gym. I needed to hear his anthropologic voice. I grew into my twenties witnessing a number of "mountaintop conversions" while working for Christian camping ministries. Sure, I could lead a prayer and a Bible study, but what my young charges really needed to know was how to (1) self-arrest when they found themselves careening backward down an icy cliff, (2) empty the canoe they had just sunk in a waterfall, and (3) jump off high rocks into an ocean filled with seals. If we could develop decision-making skills while tending to the needs of one another, certainly we could decide to love and be loved.

We needed the metaphors of wilderness to carry us through our urban ordeals, as much as we needed to experience the natural consequences of our actions. I worked hard to retain the outdoor life to which I had become accustom by guiding people into the backcountry and brokering life-changing experiences with mosquitos, marmots, lava tubes, and snow. All it took was discipline, courage, and the ability to read: the topography lines on the trail map, the shape of cumulonimbus clouds, the eddy behind a rock, the moment on a summit when one's eyes gently close to embrace the wind, the sunlight reflected in dripping sweat, faces with berry-stained smiles, bird songs, the pain inflicted by hornets' nests, and badger holes. It was common in those days to just "bag" peaks as if they were trophies, but "guiding" meant teaching others like your life depended on it, because it did. For the most part, I had been an anthropologist filling backpacks with only the things trail-weary teens would need to carry. I knew their heads and hearts could find the rest of what it would take for reaching out to one another, along the way, and a short time alone on solo would teach them they were not alone. As a classroom teacher, the closest I came to guiding in the backcountry was the Saturday a 6th-grade student was on a science club field trip with me to the airport. The flight instructor had brought her plane from across the river and was passionate about teaching others to fly. I took a backseat and held on. Afterward, the student confidently walked with her dad toward their car, carrying more than a good report card under her wings.

Before guiding others, I spent years with volunteer instructors who warmly provided instruction in rock climbing, paddling, cooking, cleaning, building, pest control, ground maintenance, vehicle repair, plumbing, and counseling. I attempted to create a safe environment in which my young charges could fail: walking into snake nests, pulling leeches from legs after an afternoon in the pond, warming waxy frostbite back to stinging red, finding friendly travelers with medicine to treat waterborne illness, as well as pulling porcupine quills and washing skunk musk from our half-wolf companion. So, when I read Cajete's pedagogy of indigenous education, I realized I needed to keep quiet, sit tight, and listen as I had that day the student flew the plane:

> One can only learn and understand to the extent that one can establish a direct and participatory relationship with the natural, cultural, and historical reality in which one lives. (Cajete, 2012, p. 47)

Quietly listening, I hear Cajete say, "Foster authentic dialogue" (p. 48). The kind of dialogue that keeps the plane on course while you as an educator are actually in the plane witnessing a middle school-aged child piloting for the first time, or a high school track star talking herself through tying the bowline knot at 3,000 feet: "The rabbit comes out of the hole, around the tree, back down the hole. . . pull tight!"

I begin to wonder why the backcountry, Outdoor School, and visits to environmental learning centers have such a long half-life in the hearts and minds of people I have met around the country. These moments resist decay: Sandy Soils eats dirt to prove it's full of nutrients; the minerals in our gold pans reflect the sunlight; we build condor caves; and the challenge course is socially structured so I have help securing the harness before swinging out of a 100-foot cedar tree over the Sandy River. Then, I realized the mercy of not being able to go to Outdoor School with all the rest of the students and doing school outdoors in the school neighborhood. I shift my perspective when given the resources as a science teacher to "establish a direct and participatory relationship with the natural, cultural, and historical reality in which one lives" (Cajete, 2012, p. 47–48).

The staybacks and the maybes. I added the title "Ms." to my green-ribboned, wood-cookie nametag and wore it over my school ID. Each morning, I put over my head a key and an identification tag with a picture of a tired, brown-haired lady, half-dressed for success. My outdoor school students arrived without sleeping bags, while their classmates prepared to spend four nights away from home. I promptly distributed wood cookies and taught the functions of the layers of a tree, as I had to hundreds of children:

"See that dark area in the middle? That is the heartwood filled with lignin that makes the center strong. When I say 'heartwood,' you say 'tall and strong.' Heartwood!"

"Tall and strong," they say.

I threw my shoulders back and lifted my arms, knowing we would need much more practice. I asked why the heartwood was not exactly in the middle of the tree cookie. They responded as I expected, with the "maybes":

"Maybe it was too shady."

"Maybe it was too dry so they had to grow more toward the sun on that side."

"Maybe there was too much competition."

We finished the story as kids across the country had done with me hundreds of times:

"Sapwood—We pump! We pump!"

"Cambium—We make new cells."

"Phloem—Yum, yum, yum, yum."

"Bark—We protect."

Two dozen or so young people wrote their names on their tree cookies as I took roll. I'd learned to pronounce the alphabet so I got more name right than usual. I showed them how to tie an overhand knot. Today mattered. We mattered. I showed them how to work the binoculars donated by the Audubon Society: "Strap over your head, point your nose toward the bird, and then lift and focus." We made two lines. Each one taught the next one until their math teacher came to get us—the same math teacher who would keep the whole week on an even keel. The rock climber and ice climber math teacher knew what "stupid arm" felt like after hours of climbing when the muscle that move one up, begin to fail. He didn't worry that I had given out the compasses with no expectation of ever getting them back. I was all for each one teaching one a skill they had just learned.

The next morning as students were entering the halls, I heard a voice asking, "Do you want to learn how to walk to the beach?"

"Put 'W' (270 degrees) under Fred."

"Stick your thumb along the compass out from your belly button."

"Put Red in the Shed and follow Fred."

"See!" he said.

I said, "You earned a bead today." Quizzically he put the bead on a paperclip strung on his wood cookie nametag, and wore it in school all week, and the next. The previous year, when I had gone to Outdoor School, a bead intended for a student fell in the duff. After a gallant search, I told the sad fellow, "Don't worry. Whether you have that bead or not, you know that you know what was taught today. No one ever loses that, and no one, not even gravity, can take it from you." Now, I think, "We aren't the Staybacks, who for whatever reason cannot go to Outdoor School. We are the Maybes."

Little researchers. "Go forth and make little researchers," my soon to be partner teacher said as I walked into the shoes of a middle school science teacher. I had just come out of the woods from counting tree species in a riparian area in the Coast Range. I had been working for an interagency-funded learning center named after one of the first flowers to bloom in spring (Think: house by a creek, with a lease to the school district, transformed into a field station near the Columbia River). My partner had a Region 10 Environmental Protection Agency grant, and a van full of shiny new Vernier probes, life jackets, and vision for salmon restoration that linked classrooms and resource professionals. We could learn in the field by measuring changes in suspended sediments from the removal of a dam. Yes, outdoors! I was responsible for the interactions between fast-moving water and middle school students in boots and life jackets in January. We resurrected our senses and the science needed to calculate the river's Water Quality Index: discovering tracks, shooting photo points, and discussing data over lunch with hot chocolate at the Grange Hall. Everything leaked out: excitement, peanut butter and jelly, dried bones, fear, ice-covered puddles, songs, discoveries, and multiplication. For a 3-year grant period, we were another generation of field researchers retrieving the river's feedback on its response to the effects of collective actions. A decade later, I received an email:

> It is important for children to know that there is an adult community that values intellectual pursuits and that these pursuits can be creative and rewarding both when you are a child and as an adult. Science Fairs are one of the few times when we have such a cross-generational interaction, with adults mentoring, supporting, and encouraging young people as they pursue questions and knowledge that is of interest to them. These are the things that students will remember.

When I was a 12-year-old researcher, I published a description of my grandfather, whom I had never met, for the Lewis County Historical Society: "He had so much love, understanding and compassion for people. He was always ready to reach out a hand to anyone who needed a lift." Now, on route to a teaching career and the removal of a dam to return the salmon, I had met an educator like my grandfather, with a vision for science education that not only closely connected science research and sustaining natural resources, but students and adults as well. I learned that science research included teaching the ethics of experimentation and revisiting past assumptions through change over time. Witnessing the dam's removal brought meaning to the Water Quality Index (WQI). Measuring turbidity gave students a chance to hear what recent generations of fish thought about the WQI. Cajete asserts that students know they are "agents of transformation in their own social and cultural contexts" (Cajete, 2012, p. 49). So are fish.

Teachers without instructional borders. I stood by my classroom door, as I had stood in the mirror image of other classroom doors, and watched as clothing passed by containing different kids wearing backpacks half their size. The same sizes, the same hurried gait, the same heavy binders are carried like Velveteen Rabbits with their fur rubbed off for comfort. Our computers are behind the technology curve even though the new downsized chips with their upsized processing speeds are created just miles away. Media messages move through my mind, suggesting a child is fortunate to attend kindergarten and be in the 50% of students who learn to read well enough by 4th grade to complete all of high school.

My mother's prayer for my vocation was to plant trees to restore the devastated forests of Haiti. The example set by my parents serving families in Tijuana as medical missionaries made me an amiable after-school volunteer. Because the program's leadership genuinely loved young people, and knew how to enlist the support of family, it was easy to support a 40-year-old after-school program with the mission to help students develop their full potential as global thinkers through science, math and engineering. The professional development that we, as advisors, received in engineering design taught us a mental model for a different kind of normal. Science and technology were not superpowers and secret weapons, but rather the means to meet the needs of one another. The baby warmer project, the prosthetic limb project, and the water-carrying project were all antidotes to the competition of building relics of war called trebuchets. Regardless of proficiency, we shared a concern for our mutual mortality that prompted us to learn more science in order to optimize the possible design of our models.

My memory flashes on eager hands and bright hearts categorized as "underrepresented" and "represented" students in the fields of science, technology, engineering, and math (STEM): the two older students that taught two classes of younger students to build and race solar cars; the student at the regional science fair whose worm gear earned recognition from the Army Corp of Engineers; the student whose videos for programming Arduinos taught an older class of students to make light shows and rechargeable solar lights for doing homework at night; and the mini donut–loving middle school graduates who returned after school every week to teach pliers and wires to others following in their footsteps. One student was "inspired by a transformed world," and is quoted as saying, "I know it will take a lot of work by a lot of people and I want to help with that." My memories and mission yield to appreciation for the knowledge of science, and the skills of engineering design that allow students to cross borders established by the inequitable distribution of resources. My stint as an after-school advisor with a mission simply involved providing opportunities after school so student could interact as the "community of global thinkers" that they already are.

Tragedy of the commons. Edward T. Hall was quoted by Forester (C. Forster, *personal communication,* Jan. 31, 2015) in the intercultural communication class I attended as saying, "Culture is communication, and communication is culture." Creating common assessments requires crossing familiar boundaries in educational systems. Non-formal educators have excelled in teaching the social, economic, and political lessons associated with the Tragedy of the Commons. They have consistently represented stock and flow models using scientific simulations for an economy dependent on renewable resources by transforming students into various species and encouraging them to race for poker chips in predator prey games. But, those are not the skills that restore species; species restore species based on their reproduction rate. An anthropologist's, or educator's, indirect and emotionally restrained stance discourages the kind of communication necessary for fair assessment of environmental literacy. Cajete warns:

> This form of ultra-objectification denies the reality of interrelationships and reduces participation and learning to only an intellectual exercise of applying a preconceived objective method or model. (Cajete, 2012, p. 47–48)

The Oregon state song still voices a culture described as a "Land of the empire builders . . . Conquered and held by free men, Fairest and the best" (Buchanan, 1927). Cajete writes:

> The result is a perpetuation of dependence on an "outside" authority and the maintenance of the political power brokers behind such authority. Indigenous People who are "administered" education, extension services, and economic development in these terms remain oppressed and gradually become dependent on the authority. (Cajete, 2012, p. 47–48)

Students need to be able to sing their own songs, standing in the context of now, and reason through potential consequences. They are one of many species born to experience the limits of higher authorities—the natural laws written in the number of heartbeats that govern our interactions. The instruments we choose to use to measure proficiency need to give students internal authority, or simply, a sense of being skilled.

Assessments for environmental literacy can be more than just another administrative dragon to slay. It might be a memory that decays more slowly if it resonates with the context in which both educator and student thrive, not a problem or crisis laid on the shoulders of a student to solve in order to prevent a perilous future. In her lecture on *Intercultural Conflict Resolution*, Forster (2015) indicated equity means, "doing what needs to be done to make things fair" (C. Forster, *personal communication,* Jan. 31, 2015). She suggested that language tends to be the number one stressor in cross-cultural communication. She explained, “emotions are universal but how they are portrayed varies.” She described how cultures that express emotions and ideas directly place responsibility for communication on the speaker. She used the metaphor of a funnel to show two how the speaker traditionally begins with open-ended questions, followed by probing statements that lead to specific facts that support a proposed solution; (2) On the other hand, cultures that express emotions and ideas indirectly place the responsibility for communication on the receiver and open up possibilities by making closed statements followed by probing questions leading to open-ended questions. It is clear that, either way, both the communicator and receiver are in agreement when they are in the middle of a discussion: they are both asking probing questions to either funnel down toward a solution or open up to new ways of thinking. We need an hourglass model, where information and possibilities flow both ways to communicate inter-culturally. Systems thinking models use a graphic language to show elements, interactions, and functions. It leaves a great deal of empty space to add missing feedback loops, which people who participate in the places they live can see better than anthropologists.

Student sovereignty over context. From my perspective, sovereignty is the means to access what one needs when one perceives to need it, and ability to respond in a meaningful way to the needs of others. Rarely, have I lived without provision. February was known as “out of provisions” month for the people of the Kalapuya (Juntunen, Dasch & Roger, 2005, p. 24). Before people indigenous to Europe came to claim acres of Oregon, many generations of people who share a common language lived and provided for one another from the oak savanna and the rich soils brought there from ancient glacial floods. August was “out of provisions” month for me, and it only happened three times.

The first time was not far from a ridge where blueberries the size of a quarter were growing. We were out of money after hiking 600 miles up the Appalachian Trail in New England, and the nice woman at the only store for miles would not take a check from a bank in another state. I still had energy to climb the ridge, and so I picked the berries I needed for a couple of days until my friend's brother drove out to Maine with his girlfriend to pick us up. The second time was just 5 rope lengths from a summit, without water. I was dehydrated and hypothermic, so my friends slept me in between them and we climbed down together 26 hours later. The third time, upon discharge from the research hospital, I had no cash for lifesaving medication, but a merciful 3rd-year resident made it happen. Unsettling and reorganization was necessary for bringing me into balance. I barely understood systems function, but remained open to new and different kind of information in order to get my needs met. Thankfully, I had help. I experienced the self-organizing principles of natural and social systems well enough to know that I needed to improve my decision-making skills.

Maybe, my birth on an island nation made me wonder what it meant to be born sovereign. Sovereignty was taught to me differently than what I came to understand later as white privilege: my family genetics provided me with opportunities stolen from other families by decisions I perpetually embodied. My self worth came out in little ways.
I remember standing in the recess line in 5th grade, wearing a button my dad had put on my jacket. It read, "BKTM."

"What do the letters on your button mean?" my teacher asked in front of the whole class.

I responded with confidence: "Be kind to me. God hasn't given up on me yet."

The parent at the home where I couch surfed had a sign on her desk. It read: “Work is love made visible.”

After college, sovereignty had more to do with living out the answer to the questions I asked myself. My co-worker put a sign on the entry table, with a question that baffled me: “Do you do what you do because you are what you are? Or are you are what you are because you do what you do?”

When the answers were not obvious, I tried to remember what someone who had my back would say: “No matter what you feel, or what people may say, or even try to lead you to believe, you are loved, and you matter.”

Another mantra came from a person who took an interest in my recovery after the kind of extensions that only surgeons can give: “I am alive, here, now, in this moment, and everything is okay.”

When I used to call grandma, she would answer, “I’m still here,” and we’d laugh.

We would retell each other the story of helping the woodpecker find its way to fly out of the barn. The wonderful bird was a determined flutter of brown, black, tan, and red feathers and had been trying all day to fly out of a dusty, cobwebbed, closed window that appeared to be the only way out. All we could do was open the barn door and gently guide the bird with the broom. There was always something important about the part of the story where the woodpecker had gotten itself out. Then, she’d hang up and walk to the garden where she’d grown all the vegetables I would ever eat as a kid of privilege who had a choice.

Autonomy, self-determination, or what I have understood as agency, were not lost on me as a teacher. I have told students overwhelmed with anxiety: “Be like the mountain. They make their own weather." As an educator, climbing mountains of updated standards, and articulating the missions of the agencies that paid me, I tried to model what I meant about educational sovereignty in the context to which I was assigned—giving a cave tour, testing a pH lab, taking an inventory of *Thelotrema* (Pierced Nipple) lichen, or making a trail map with high school students, in February, where people without homes lived in the invasive English Hawthorne on top of a butte in the city where the mayor and his staff would have bunkered if a nuclear war ensued. More importantly, I have stories of time and again where a middle school student knows the obvious course of action in the present for a non-solution from the past to which my lazy synapses cling.

Muskox. A couple of years after reaching the summit of Ten Peak Mountain with a dozen high school students and newly wed, I was offered a job with the Conservations Corps by the Minnesota Department of Nongame Wildlife. It was a welcome respite from my job at a care facility, where I washed windows, cooked, cleaned, sang, and protected my elderly charges when it was 20 below zero outside. I was assigned to educate teachers about balance, especially balance between people, game, and nongame wildlife. I used a curriculum written by a national group of teachers, game hunters, and conservationists, called Project WILD. One of the simulations that became part of our repertoire was Muskox Maneuvers. We would transform ourselves into a heard of muskox on the tundra, which wasn't hard to imagine since we were in below-freezing temperatures outside in the snow, wearing our insulated boots. I assigned a few teachers to play the part of wolves, and the rest worked together to protect the group. Of course, we soon learned how to organize with the weakest, youngest, and eldest ones in the center of the circle, protected by the older ones, who faced outward, our tailbones gently reminding them everything would be okay no matter what. Once the teachers got the gist of it, it was pretty obvious that muskox herds had internal hierarchy and self-organizing principles that prevented the wolf pack from taking too much.

My supervisor—who also took time to harvest trumpeter swan eggs from Alaska with his son in order to restore the flocks that migrate to Minnesota, and write books about attracting birds to your backyard—saw fit to send me to Washington, DC to meet my brothers and sisters in the corps from around the nation. They asked me to testify in a room full of their representatives and senators to support a bill for what would become AmeriCorps. The program would provide small wages and education stipends for interns, which might lead to a living-wage job. It would be a welcome improvement to the vision of the Civilian Conservation Corps (CCC) that put people to work building roads and walls when we, and our grandfathers, had no work. We all needed CCC opportunities that would lead to stable income and protect us from the wolves at our doors. Painting over the graffiti in our neighborhoods, only to see it replaced each night by others in the neighborhood, was not building the relationships needed to secure everyone an economically safe place in the community.

The next month, I started a job as one of many directors of environmental learning centers (ELC), in a state park along an 11-mile lake in the maple-basswood forest. It was not the coveted ELC of the Northwoods, where they track wolf packs and make methane candles from marsh biomass, masking tape, and soup and tuna cans. But we ice fished with kindergartners, made maple syrup with 4th graders, and put snowshoes on the older kids who built snow shelters and tracked the animals. Before getting on the bus, we would explore the restored tall grass prairie, looking for nests, and tie a red bandana on a student who wanted to be the lost horse so the rest of us could all run to find it. High school football players came for a week with their classmates and wrote poetry about the Yellow Lady's Slipper and masses of frog eggs. It was their teacher who taught me to never accept less than a salary commensurate with a teacher's salary for working with teachers and parents to transport, feed, house, instruct, sing, and tell stories with hundreds of loved ones in our care. At the request of the administrators, who often came overnight with their students, I would visit the schools in the evening to answer questions that brothers and sisters had not already answered about spending the night in the cabins. These farming families found a way to pay the not-for-profit costs. When I returned home to Oregon to search for similar work, I found no positions for Directors of ELCs, or salaries commensurate with teachers. The land ethic of Leopold seemed buried in traditions from a culture uninterested in songs about mosquitoes that could carry you away. The maple sap did not run here like it did on sunny days when the nights froze. But, I served as an instructor for Outdoor School at Trout Creek anyway, and gradually learned to teach about the Columbia River basalt flows embedded in waterfalls.

Twenty-three years later, a fellow Outdoor School instructor named Mouse told me a story while a family of three generations of Oregon Outdoor School graduates and my citizen-self waited to talk with our senator. We were hoping he would vote for Senate Bill 439, which would establish an Outdoor Education Account at the Oregon State University Extension office for the purpose of providing outdoor school programs, like the one my older brother, younger sister, and I got to go to as 6th graders in the 1970s.

The familiar lump of homesickness rose in my throat, and tears filled my eyes, when she talked of counseling a cabin with a girl who could just not stop crying on the first day of Outdoor School, *like me*. She had tried all the strategies recommended to high school counselors, but the tears continued into evening after campfire, teeth brushed and back from the bathhouse. So, she left the girls alone in the cabin to seek some more support. When she came back, she saw that the girl had taken her mattress off the bunk bed and laid it on the floor. She was resting quietly in a circle of mattresses all around her, where 11 girls were sleeping like muskox. These younger girls understood that community was more than cause and effect. Deep in their spirits, self-organizing principles generated a feedback loop to replace the missing dopamine of familiar comforts that the proteins of individual tears could not. Not to mention, they had witnessed Mouse use every bit of wisdom and love she had to include the homesick girl in all the day's activities together with them. It took a while for cultures, and traditions, and exhaustion, and the community to act. Too much had been asked of the little girl that morning without yet knowing the strength of a group to reorganize the mattresses into a model of healing relationships.

When I was in middle and high school, the adults in my world had strange ideas about me that I had never considered. The health teachers tried to explain that there would be times where "Your hormones have taken your decision-making powers from you." Another kept reminding me, "Communities of the world are facing crises that you must solve, or we will all perish." A third idea was right around the corner: "Your genetics will perpetuate disability in future generations." The dynamic scientific forces of nature were pitted against my family and community. These educators seemed to be asking me to live with a foot in each world: my family's traditional understanding of the world, and counterfactuals that exceeded natural limits. Pondering Mouse's story, I realized that the basics of creating feedback loops are already in us. We just needed to make the "limitless love" of a muskox explicit. Unfortunately, the two worlds are pitted against each other even though science and community are one in the same. The world has turned out to be even more difficult than I imagined:

> Ecological economists argue for reforms that would ground economics in ecological principles and the constraints of thermodynamics. . . .We continue to embrace economic systems that prescribe infinite growth on a finite planet, as if somehow the universe had repealed the laws of thermodynamics on our behalf. (Kimmerer, 2013, p.)

As my friend finished her story about the homesick little girl, I remembered going to Outdoor School at the age of 11 and leaving the familiarity of a classroom with a teacher who fed us candy bars from another county made of chocolate and bees, and bowling with friends on Thursdays. My friends from the bowling league were in other cabins. I knew the world had limits. I knew I had limits. Too many counterfactuals had made it hard for me to stop crying. In my social story, chocolate was not supposed to have bees in it, and I always had friends on Thursdays because we bowled. Eventually, I met other children who could imagine a world in balance that flowed with common sense. The kind of common sense I also learned from teaching teachers to be muskox.

Dung beetles. Great laughter erupted when I shared my anxieties around submitting grades as a first-year teacher in 1994. "What the hell difference does a middle school grade make anyway?" The teacher I asked chuckled. "Didn't they matter to the students and their families?" I thought. "Wouldn't there be some consequence for poor performers, and some reward for those who were stellar?" Feedback from teachers never robbed me of the provision of my parents. But now, working in a place the teachers referred to as "Felony Flats," I worried for the kids' lives. Surely, the daughter of a general from Vietnam, who could learn English well enough to explain science, has a gift worthy of an "A"!

So, when asked, "Is this for a grade?" or "Do we have to do it?" I responded with a wild, unforgiving stare that communicated the question, "Whoever could make anyone do anything?" I had learned in kindergarten how to follow the rules and stay out of Brother Dowd's office where he religiously hit us with a ruler according to the number of marks our teacher chalked on the board for laughing during naptime or sharpening our pencils out of turn. I learned I had not been alone in the world when I heard Fasheh (2015) read his paper, *Over 68 Years with Mathematics: My Story of Healing from Modern Superstitions and Reclaiming My Sense of Being and Well-Being*, "Young people in Arab countries, but also for youth in general: they have been victims of control, mainly through being constantly measured" (p. 33). As I pined over all the unknown attachments parents would make to the grades in order to motivate and cajole their children into success, I despaired. "Justice and mercy," my dad told me. I felt the commodification of knowing as opposed to learning. Testing and grades capitalized upon the due dates of progress reports rather than helping a learner confirm they know measured by the widening of their eyes and arms raised in celebration for balancing a challenging chemical equation. I wanted to build communities that gave students a place to share their gifts in an environment where they feel appreciated for sharing them.

Rosemary William Wray (1993) taught me how to be a teacher who gives grades well. She said, "Every child's schoolwork is a gift." The teacher who gave us chocolate covered bees expected an organized, sequential story with characters that wandered through a beginning, middle, and an end, but I wrote a detailed and quite graphic tale of my nightmare. I needed my mother to talk with the teacher in order to understand how I had misunderstood the meaning of *story*. That was the first time I realized that grades were not gifts of reciprocity that acknowledged an intimate discourse with a teacher. I could write in cursive and count by 5s before kindergarten. I could sing and read with confidence, appraising an author's attempts to have me believe in a world with "Red fish, blue fish, one fish, two fish" and that I would grow up to have a dog named Spot. I could do all these things before teachers told me I could do them by assigning grades according to their practice. My schoolwork was a gift, and grades simply a snapshot in time.

So, why did I feel like a cash register churning out receipts by making corrections and marking scores on students' beautifully written work? I took roll, asked students to do a warm up, had them share, gave a lesson, called on students, organized labs, changed seats, collected papers, told stories, showed videos, offered ultimatums in exchange for silence, and gave a wrap-up. I stepped over giant binders that students rolled up academic hills like dung beetles. Worksheets and lab reports overcame them, and rolled them back to new starting points every time they seemed to reach the learning goal. I simply wanted them to know when they were able to see parallax, and that I valued their knowing that they knew. If they knew they did not know, we rolled forward to the next demonstration. I asked them to look for parallax by holding a finger up to a reference point, closing one eye and then the other, to notice when the reference point leaps from one side of one's field of vision to the next. And, there was the answer to all of my grade vending, my field of vision. Was it because the four walls of a classroom limited me to measuring their knowledge and skills on space science using standards rather than taking them outside to actually look at the stars?

Before going outside, Meadows (2008) might have asked me to reconsider, from a systems thinking perspective, the level and pace at which I used instructional time. She might have determined that learning, like a renewable resource, was "flow limited" (p. 71). She would have told me:

> They can support extraction or harvest indefinitely, but only at a finite flow rate equal to their regeneration rate. If they are extracted faster than they can regenerate, they may eventually be driven below a critical threshold and become, for all practical purposes, nonrenewable. (p. 71)

Had I developed the stock and flow diagram to consider my choices, I might have seen what Meadows (2008) generalized for all systems with renewable stocks, including collecting assignments and using them to produce grades. I had two choices for preventing oscillation or collapse: "(1) Notice that the critical threshold beyond which the [students'] ability to regenerate had been damaged; and (2) Rapidly slow the [pace of instruction] as the [students] become depleted" (p. 72). I became curious about how to structure the class to aid regeneration. Maybe we needed to tell our stories, laugh, and cry more often? After all, I had learned that crying was a body's last autonomic mechanism for re-establishing relationships. A strategy Meadows might have referred to as one of the latent behaviors of a system that assigns grades to gifts.

> The trick, as with all the behavioral possibilities of complex systems, is to recognize what structures contain which latent behaviors, and what conditions release those behaviors—and, where possible, to arrange the structures and conditions to reduce the probability of destructive behaviors and to encourage the possibility of beneficial ones. (Meadows, 2008, p. 72)

Meadows would have made a great middle school administrator. As I am asked to teach and assess with higher order cognitive thinking skills, Fasheh's (2015) lecture lifts me up:

> A million plastic flowers put together for a hundred years can't produce a single plastic flower, whereas one seed of a real flower, after it withers and dies, can generate a million flowers. It is the spirit of regeneration that makes the difference between what is real and alive on the one hand, and what is make-believe on the other. (p. 41)

I had to consistently remind myself that the young people in my midst were not their grades. I had forgotten to ask how people across the planet still use their knowledge of the stars to make decisions about when to plant or harvest, as well as to figure out which way leads home. None of the test questions measured how well they could prevent parallax from causing them to wander off course.

In a discussion about educational testing policies, a parent said to me, "What we count impacts what teachers teach and what students learn." I remembered this as I was sitting with educators by the Sandy River. We discussed what to assess and count, and what mattered most was quality of life, meaning, relationship, feeling safe, and feeling loved. I told the group,

> I am not speaking for all teachers, but my experience as a teacher has felt like being a customer and a source of little people to train, who go home to train their parents with whatever message the educator is obliged to teach because of the organization that pays the salary of the person with the message.

After our conversation, I wanted to measure how well my work brought students closer to discovering that they are capable of making change, self-advocating, and leveraging natural systems. Did they know that natural systems consider them an essential part of the community? Could they "speak for the trees," like the Lorax in Dr. Seuss' stories, as a demonstration of their gratitude for their function within the context of their familiar world? Could they hear the unfamiliar human, and weather communities talking in patterns, or use data and observations gathered by species beyond our own who can be trusted to extend our senses?

Two phrases guided my indoor and outdoor educational efforts: (1) “learning by doing” through strategies like labs, simulations, and practice; and (2) “each one teach one,” where students teach each other, using strategies like partner talks, videos, class discussions, presentations, or posters. Because I was trained as an elementary teacher, I tended to teach ecology using familiar graphics like food webs and energy pyramids, and though we did not have the computer skills to model flow mathematically, we gained a reasonable sense of how energy flows and matter cycles. The scientific method was an easy complement to data collection, but I generally fit the content to the context. The result was that I often gave young people responsibility over the context, which was more concrete if we were outside. For example, an inventory of stream invertebrates indicated fast, cold, clear water, or slow, warm, turbid water, which reduced the probability of a salmon growing to adulthood. But for older children bound to a classroom, assigning purpose seemed more abstract and complicated. Gopnik (2009) writes,

> For older children, attention gradually becomes more controlled by their internal agenda rather than by the intrinsic interest of external events. So it becomes more difficult to use their attention as a reliable indicator of what they see. And for adults, of course, if we decide to attend to the ball even the wildly unexpected gorilla won’t distract us. (p. 118)

I saw my teaching as an experiment within a larger experiment that Gopnik (2009) referred to as schooling. The experiment that my grandfather had begun by consolidating one-room school houses in order to hire more teachers and bring large groups of students together for band and sports.

According to Yong Zhao, one of the provocateurs speaking at the systems thinking conference called Camp Snowball 2015, the function of the school system was "not to fix deficits according to external standards, but to help each person become successful in their own way." Zhao said that education should come with "side effect warning labels" that say things like: "This reading program may boost test scores, but may make students hate reading forever." He claimed, "Christmas Eve is here, and every student is Rudolph"—in other words, the confident one who will lead Santa's sleigh through the mythical snowstorm of education to become a confident entrepreneur in a community of diverse wants. "Every individual has talent, and every individual has passion." He confirmed my initial hunch—that a child's schoolwork is a gift—when he said, "Direct instruction is very good at generating knowledge, but the worst in supporting curiosity." Ask anyone who's been exploring with me, and they'll tell you that my favorite phrase is, "I wonder if . . . ," coupled with "I don't know, but let's figure it out." This was not the kind of not knowing my principal wanted to hear in my interview, but it was the kind of curiosity and passion that helped me repair the radiator hose of my car with a foam sleeping pad and duct tape in order to get to the interview on time.

Zhao continued, "PISA is one of the most destructive forces of the 21st century" because it reinforces the idea of competition. Fasheh (2015) seems to agree with Zhao, since, as he writes:

> Sitting on our behinds for 12 years and looking at meaningless words (on boards, papers, and screens), with no action and no context, and calling that learning, has caused much harm. Myths existed in other civilizations, but the modern one is the first to measure intelligence, one's worthiness, and a country's development using numbers—and to claim that such measures reflect reality. (p. 44)

Zhao encouraged, “Don’t fix the past. Provoke a new paradigm that is driven by students’ passion.”

When I was taught the language of systems thinking—systems’ boundaries, functions, elements, interactions, causal loops, archetypes, and antidotes—I was told its purpose was not like other curriculum designed to be the solution to problems in schools. Connection circles were to be used tacitly to model systems behaviors. When I tried using the tools, I heard students’ voices that had been sitting silent a long time. Because the students were considering multiple variables, they participated in a critical discussion about a myriad of consequences. Like conversations with cougars, I could hear more from the students’ eyes and the angle of their shoulders. I backed away, knowing who was in charge of the present moment, who was sovereign, and who would need to remain so in order to provide for one’s self and others well into the future.

Responsible science. Scientists are the people in my culture who are closest to the patterns and cycles of the natural world. Learning science gave me the ability to fend for, and care for, myself, and others. It connects me to the natural systems that sustain me. David Sobel, in his 2008 publication *Childhood and Nature: Design Principles for Educators*, concludes:

> The pathway to responsible environmental behavior is a bit trickier than knowledge leads to attitudes lead to behavior. It’s more like a sense of agency and control leads to knowledge of issues and action strategies, which lead to an intention to act, which under the right precipitating conditions leads to environmental behavior. (p. 145)

Golley (1998) writes, “It is obvious that both analysis and synthesis are necessary for a full understanding of phenomena. In the United States, however, the emphasis is on problem solving, and therefore the analytical approach is most commonly used” (p. 16).

National Geographic (2015) writer Joel Achenbach describes why science struggles to elicit environmental behaviors:

> It's their very detachment, what you might call cold-bloodedness of science that makes science the killer app. It's the way science tells us the truth rather than what we would like the truth to be. (p. 47)

Sobel's solution fits with Achenbach's research: "Science appeals to our rational brain, but our beliefs are motivated largely by emotion, and the biggest motivation is remaining tight with our peers. . . . That need to fit in is so strong that local values and local opinions are always trumping science" (p. 47). The knowledge and understanding of science, especially environmental literacy, is sometimes hidden from us by those "experts" who would protect us, those "profiteers" who benefit from our not knowing, or its complexity and our willingness to attend to it.

Velcro theory. I stand at the door of a classroom and fight to keep it open wide to the science community: local engineers, nurses from the neonatal care unit at the hospital, 4th generation farmers, satellite launching team leaders, and software designers. I use my Velcro theory, which is based on the hope that students will remain engaged in an activity or lesson long enough to generate their own questions. They would be the kind of questions that will keep them searching for deeper understanding for the rest of their lives. With every guest speaker and field trip, I'm hoping those who apply conceptual science to their daily activities will help students find industrious vocations. I toss students toward the metaphorical walls of the solar industry and wastewater treatment, the microscope industry and wind turbines, hydrogen fuel cells, and the limits of the human body. I put probes and binoculars in students' hands to extend their senses, but make sure they can approximate without them. Time is given to their questions. They speak to one another, listen to one another, question one another, and rest with one another in the knowledge of patterns they can predict as well as those they cannot. All because my dad taught me to fish by looking in their stomachs to see what they had been eating so we would know which insect to put on the line next time. Dad studied the stream to predict where the fish would get his next meal. Science is not a commodity. When I focus too much on the simple cause-and-effect principles of science, it is an injustice to those interdependent systems that function from mutual causalities.

In my opinion, the commodification of science is worse than not yet understanding science because one does not have access to the tools that provide the appropriate kind of information to make particular decisions. Participating in the processes that sustain us is exactly the science to which we need to remain present. Science has been commoditized with Latin-derived medical codes and easily confused prescription names that require an industry of people to decode their secrets. By teaching both photosynthesis and genetics, food could simply be food without secret genetics. Eliminating rewards for the commodification of science, would end the need to separate the seeds in the greenhouse and code them for confidentiality because we would have the science we really needed to make choices about our food in the first place.

We are only three generations into an experiment called Outdoor School in Oregon. What if we really lived like the calving of ancient glaciers, caused from carbon dioxide increases in the atmosphere, could be arrested, like we did before when we worked to prevent the extinction of the condor and to repair the hole in the ozone. Science is not for building technologies; it is the constraint that informs us of our boundaries and our responsibilities to natural systems. Outdoor School began because natural resource specialists saw that students were losing their reference points and science knowledge of natural systems. Both the natural systems and communities in which students live are at the root of a place called Oregon's economy and well-being. Kimmerer (2013) writes, "Doing science with awe and humility is a powerful act of reciprocity with the more than human world" (p. 252). Our technology is not just what we do with science knowledge to extend the life of community and natural systems, but whether technologic innovation chooses to ensure that products, processes, and systems function to restore regeneration rates to the levels communities and natural systems can sustain.

Counterfactual futures. Intergenerational knowledge and interaction transform the language of crisis and panic into one of curiosity and counterfactuals. I thought maybe by purchasing the book *The World of the Kalapuya: A Native People of Western Oregon*, I could read the refuge where I walked. I wanted to see with "hunter and gatherer" eyes, but I ended up seeing through the eyes of people from the year 1880 (Juntunen, Dasch & Roger, 2005, p. 110). The people were forced into the Indian Manual Labor Training School by Euro-Americans, like me.

> At first some teachers taught reading and writing in English, but later, vocational skills were considered more important. Boys learned skills such as blacksmithing and carpentry, and girls learned how to cook, sew, clean and do non-Indian crafts such as embroidery and sewing beads on to leather. . . . Children were punished for speaking their native language or practicing native ways. (Juntunen, Dasch & Roger, 2005, p. 111)

I thought if I looked harder at the landscape, plants, and animal communities, they might interpret the story for me. The dry grasses and recently planted oak saplings told of people whose great-grandparents burned the grasslands to encourage the growth of the trees. Reading the experts' interpretations did not fill the generational gap. I was walking in the midst of natural systems that could help me, but the authority of the text was once removed from the voices of the people, plants, and animals with which I wanted relationship.

An exciting project. My mindset for environmental justice and community has changed through the years, but once I was not unlike a man featured in Courtney Martin's (2010) *Do It Anyway: A New Generation of Activists.* His name is Tyrone, and he told her:

> Having class privilege means I get to see living on a small budget as an exciting project rather than a stressful necessity. Truly being poor is expensive, and having had good healthcare my whole life, never having to go into debt, not having to take financial care of my family, and a million other things make it easy for me to live cheaply. (p. 121)

Tyrone summed up my old mindset pretty well:

> All you have to do is read an article about climate change to get totally freaked out about the future. But that's the psychology of capitalism, right? Make everyone feel so insecure that we hoard all the resources we can and forget how to share and take care of each other. (p. 130)

I simply acted "green" in order to prove that I was not part of the problem, which ended up negating my ability to reason. Choosing a self-powered, popular, low-impact lifestyle shortened the time delay between waste and sustainability, but it did not meet my need to be able to sustain myself, and others.

In summary, it was not my privilege to live directly from the forests I knew in my youth. My great-grandfather consolidated the school children that would be displaced by hydroelectric power for the industry of the port city on the ocean side of the mountains. My grandfather brokered the knowledge of agriculture and home economics by hiring vocational teachers for my mother and father. My grandmother brokered the vegetables from her garden into canning jars that we brought home to store in the hall closet. My father showed us how to keep bees and spin honey. I knew every cow I ate and the cousins who gave us salmon and elk to eat, but I only knew how to shoot tin cans, pick berries, and pick up apples—overcome by gravity—for sauce. Decisions guided by my grandparents' copies of the *Farmer's Almanac* were beyond the scope of my understanding. By age 12, my schooling had more to do with listening and writing than cooking and sewing for myself. But, it has been my privilege to remain close to the forest (the red cedars, pack of coyotes, the barred owl, the hawk) and provide the public service of teaching science concepts and thinking processes to middle school age children. I believed that the ability to use science knowledge to engineer solutions to everyday problems would provide individuals with training that would allow them to live and help others. By using science in love, they would be capable of putting the "v" in their vocation. After all, "Work is love made visible."

Sitting outdoors by the recently wood-chipped trail near the lodge, I found myself with a diverse, yet remarkable, united-in-heart quartet, trying to make sense of what to assess for environmental education. They were not asking me to apologize for telling the truth about how much I feel like a customer, rather than a public servant, when I sit with people who are trying to sell me environmental programs and messages. I have little fiscal authority, only the responsibility for teaching science with 100 middle school students, nine months of the year, in a classroom, using our ingenuity and resourcefulness. It feels as if the resources are right there, but I do not have the skills or knowledge to unlock access to them. My reflection reads:

> I would like to measure how well my work brings students closer to discovering that they have the skills and knowledge to leverage the systems of the world to help one another.

Assessing a middle school student's hopes for a counterfactual future gives educators the same kind of responsibility as parents, or what my teaching partner and I call *parent in loci*. Formal and non-formal educators share responsibility for preventing harm to young people by ensuring they have the skills and knowledge to access the resources they need now. Environmental literacy scores that show proficiency for systems thinking measure a student's ability to show where a system might leverage itself to regain balance. In the event that one chooses to implement a model of environmental literacy, the community will be glad that siblings, peers, grandparents, parents, and educators offered feedback and support to the student.

Each of my assignments was a gift, and a grade of "A" was like a thank-you card. I attached myself to those teachers who gave me the feedback and attention that I craved, to the point that I tried to adopt them as parents, completely taking for granted the parents at home who were trying to help me survive with sewing projects and haunted houses and go-carts and cooking and weeding and macramé and wood carving and silk screening, and a thousand other creative opportunities that turned life with my family into art. My attachment to caring adults was displaced by grades, gold stars, percentages, and comments. The games I played as Teacher's Pet quickly escalated to an addiction to red pen; I would scour a teacher's written feedback for any kindness and any twinkle of being understood.

Teaching people to use a scoring guide with fidelity is usually a process of moderation and calibration. Calibration involves looking at multiple works and establishing a point of proficiency. A student's work is used as an exemplar to adjust the rater to recognize a specific level of achievement. The educator works like a trustworthy lab probe, providing precise and accurate measurements each time. An assumption of meeting the goal, and a perceived level of quality, is at work in calibration, so the measure can be used to make decisions about the ramifications of students and educators of reaching, or not reaching standards. It is congratulatory at best, stifling at worst, and results are used to reinvigorate a mental model of crisis and controversy. To prevent Drifting Goals, systems thinkers would never settle. Each effort would be a step for the next best effort.

In moderation, a collection of students' work is used as evidence of proficiency in many different contexts. Moderation helps raters maintain a significant level of accuracy. For environmental literacy in Oregon, moderation is essential because of the diverse ecoregions. Each population has developed specific relationships with different natural systems over time, so the mental models associated with indigenous species in one region may be entirely unfamiliar to another. This is where the student is sovereign and educators must learn to cross assessment and cultural boundaries by listening closely to students. Simply said, be ready to advocate and protect the middle school student's voice. Middle school age children have a crucial role to play in testing existing systems because they are so clearly beginning to make decisions for themselves that impact others and their future. A model that is counterfactual to historically trusted models will need as much supporting evidence from experience as any. Make no assumptions about the limitations of 12 to 15 years of experience on a planet. My father taught me how to catch a raccoon. He said that all I had to do was put something shiny in a log with a hole just a little smaller that a raccoon's fist. The raccoon would see it and become curious. When it reached in and took hold of it, it would be trapped until it let go. The counterfactual models of environmentally literate people of any age have helped more than one raccoon act differently so they could be freed from the trap.

Combined Results of the Delphi, Field Test, and Autoethnography

This mixed methods study used three techniques to determine the level of consensus shared by formal and non-formal educators: (1) the Delphi technique engaged 11 middle school students, teachers, and environmental educators in defining the construct of proficiency with systems thinking as described in the 2010 OELP; (2) the scoring guide created in Stage 1 was field tested with 11 formal and 14 non-formal educators by measuring inter-rater reliability with Cohen's kappa; and (3) an autoethnographic narrative was written describing how the researcher came to make meaning of assessment, equity, and environmental literacy through her experience as a student, her work as a teacher in public schools, and multiple opportunities to serve as a non-formal educator.

The commonalities between formal and non-formal educators reveal a high level of agreement (80%) for the importance of including specific skills as a measure of proficiency with systems thinking. The reliability between the two groups of educators field testing the scoring guide for systems thinking with a sample of students' work indicated a moderate level of agreement accounting for chance by using Cohen' kappa (k = 0.54). The difference between the median of the two groups ratings on the proficiency of the students' work sample did not exceed 1 level supporting my "hunch" that formal and non-formal educators could use the scoring guide with equal validity. Feedback from the two groups however raised questions about: the instruction students received prior to completing the work, whether the students had spent time outside, and whether they had spent time outdoors as part of the investigation. Variation in the median of the ratings assign by the two groups indicated that six skills in the scoring guide needed further clarification:

1. The student work identifies long and short-term consequences.
2. The student work shows how a system's structure generates its behavior.
3. The student work identifies the problem of a situation.
4. The student work displays proposed changes and outcomes via easily understood diagrams. It uses visual graphic skills to clearly present how the changes effected the environment.
5. The student work analyzes data.
6. The student work shows evidence of collaboration. It uses the design cycle to explore multiple solutions for the same problem, and creates a +/- chart for each solution.

The autoethnographic narrative identified three issues concerning equity in assessment for environmental literacy: (1) student sovereignty over the context, (2) the historical administration of education to students from the nation of the Kalapuya, which dismissed indigenous knowledge and skills by choosing to teach vocational skills, and (3) the commodification of science that prevents the flow of information that citizens need to support their decision making processes. The narrative found that assessment in formal and non-formal education served a similar function, which was to support students by acknowledging their gifts or skills with attention, tokens, written notes to parents, and witnessing personal expression of failure and success. The narrative described the influence of intergenerational interactions on decision-making, and the type of dispositions between educators and students required to manage conflicting data. The combined results of the study confirm what Osborne (2007) identifies as one key action educators can take to include all future citizens in decisions - give students the "opportunity to consider data which has no clear interpretation and to consider plural alternatives" (p. 179). In the words of the middle school students and educators who decided to participate in this study evaluating an instrument for Oregon's environmental literacy plan: create solutions for systems that are not in balance, present the complex inner workings of a system in a simple and succinct way, collaborate, explore multiple solutions, and share ideas in a way that people will understand you.

Summary

In this section, I presented an analysis of the data for answering the research question, "What does the level of consensus between non-formal, and formal educators reveal about designing an instrument to measure a middle school student's level of environmental literacy in Oregon?" The construct of proficiency with systems thinking for environmental literacy was defined by a forum made up of 11 individuals including: formal, and non-formal educators, and middle school students. The level of inter-rater reliability was calculated as moderate (*Cohen's kappa* = 0.54). The five themes of my autoethnography provided social, cultural, and political context for this investigation. In the final chapter, I will present recommendations for policies and practices in teacher and administrator education to support the assessment of proficiency with systems thinking for environmental literacy.

Chapter 5

Discussion and Conclusions

In this final chapter of my investigation regarding what the level of consensus between non-formal, and formal educators reveals about designing an instrument to measure a middle school student's level of environmental literacy in Oregon, I will review the systems thinking scoring guide designed to measure the proficiency of middle school students in relation to the social, cultural, and political context of my lived experience. In addition, I make recommendations for changing my own practice of assessment as a formal and non-formal educator along with the preparation of educators and adminstrators. I also suggest specific rule changes in Oregon for requiring the devotion of instructional time to the learning strands of the Oregon Environmental Literacy Plan, and advocate for parity between formal and non-formal educators.

This mixed method study was based on a constructivist theory of learning, where validity and reliability were "derived from community consensus regarding what is 'real,' and what has meaning, especially for future action" (Guba & Lincoln, 2005, p. 197). An evaluation of the validity, reliability, and equity of eight existing assessment instruments informed the design of this study (see Table 17):

Table 17

Existing Evaluation and Assessment Instruments

Instrument	Score
Oregon Assessment of Knowledge and Skills (OAKS)	datum
Oregon Department of Education Work Sample	6
PISA: Programme for International Student Assessment	2
Middle Years Programme (MYP)	5

Table 17 (continued)

Existing Evaluation and Assessment Instruments

Instrument	Score
My Environmental Education Evaluation Resource Assistant (MEERA)	5
Assessments for Environmental Science Literacy—Michigan State University (AESL)	5
Ecological Understanding as a Guideline for Evaluation of Non-formal Education (EUGENE)	3
National Environmental Literacy Assessment (NELA)	1

Note: These eight assessment instruments were reviewed for construct validity, reliability, and equity. For an explanation of the reasoning behind the scores see Appendix H. Each instrument was compared for the criteria against the datum of the OAKS and score with a +1 if the instrument's characteristics appeared stronger than the datum, and -1 if they appeared weaker.

A comparison of these instruments using the evaluation criteria shown in Table 18 indicated that scoring guides were valid and reliable instruments for assessing constructs outlined by the learning strands in the 2010 Oregon Environmental Literacy Plan. A forum made up of middle school students, formal educators, and non-formal educators

Table 18

Evaluation Criteria Definitions

Evaluation Criteria Definitions
Validity: • Operational construct is framed before use • Operational construct includes strands identified in Oregon Environmental Literacy Plan • Interaction between testing and treatment includes outdoor experience
Reliability: Provides measures of proficiency in environmental literacy that are meaningful to educators from formal and non-formal settings
Equity: Allows educator to provide specific supports to individual students. The initial definition of equity was based on Singleton & Linton's (2006) description: "an operational principle that enables educators to provide whatever level of support is needed to whichever students require it" (p. 47).

Note: A Pugh Chart was used to compare the eight different assessment instruments listed in Table 17 using these criteria. Each one was compared to the Oregon Assessment of Knowledge and Skills (OAKS) and scored based relative to the OAKS. For an explanation of the reasoning behind the scores see Appendix H.

used the Delphi technique to create a scoring guide for systems thinking (see Appendix J). It was then field tested with another group of formal and non-formal educators using a sample of middle school students' work before examining it through my social, cultural, and political understanding of equity as a formal and non-formal educator. Even though the findings of this study begin to describe consensus between middle school students, formal educators, and non-formal educators regarding proficiency with systems thinking in environmental literacy, the construct is "still open to new interpretation as information and sophistication improve" (Guba & Lincoln, 1998, p. 211).

Synthesis of Findings

In the quantitative part of this mixed method study, Stages 1 and 2 were designed to build consensus around the construct of proficiency in systems thinking. Bias was managed by using measures of central tendency to describe the level of agreement, between middle school students, formal, and non-formal educators who participated in the Delphi, and field test. Reliability was measure with Cohen's kappa ($k = 0.54$) in order to determine level of agreement between the two groups of educators, and to find which definitions of the construct needed further clarification. The autoethnographic component, served as the qualitative aspect of the study, digging into my own preconceived, and continually changing notions of assessment, equity, and environmental literacy. In some cases, the stories served as potential antidotes to assessment traps identified by educators who field-tested the scoring guide.

Construct validity for proficiency with systems thinking skills. Findings from this study indicated that middle school students, formal educators, and non-formal educators were able to reach a level of 80% consensus on 25 skills that were important, or very important, for demonstrating proficiency with systems thinking (see Appendix E). The formal and non-formal educators who field-tested the systems thinking scoring guide shared a level of 76% percent agreement for operationally understanding the systems thinking skills being assessed by the scoring guide. The statistics revealed six skills that need further clarification:

1. Identifying long and short-term consequences.
2. Showing how a system's structure generates its behavior.
3. Identifying the problem of a situation.
4. Displaying proposed changes and outcomes via easily understood diagrams, and using visual graphic skills to clearly present how the changes affected the environment.
5. Analyzing data.
   Showing evidence of collaboration, and using the design cycle to explore multiple solutions for the same problem by creating a +/- chart for each solution.

Reliability of formal and non-formal educators' scores of students' work. The inter-rater reliability measured using Cohen's kappa was 0.54, indicating a moderate level of agreement that was not simply due to chance. Educators voiced their concern that outdoor experiences, such as field study, were not required as a prerequisite for proficiency, and that they would need to see students "in action" in order to give an accurate rating. The skills assessed by the systems thinking scoring guide resonated with what I learned from Janice Jackson's *Cultural Iceberg* about deep culture decision-making skills that cross cultures and generations. The scoring guide included systems thinking tools that I was introduced to via Meadows (2008) in her book, *Thinking in Systems: A Primer*, which gave meaning to the golden man with the ax on top of Oregon's capitol building. Natural systems that depend on renewable resources continued to speak to us through their regeneration rates even while technologies gave us a competitive advantage for securing the remaining resources. Kimmerer reminded me of how "reciprocity" heals, or what the scoring guide refers to as the skill of creating solutions for systems that are not in balance or unsustainable (p. 189). She writes,

> One of our responsibilities as human people is to find ways to enter reciprocity with the more-than- human world . . . through gratitude, through ceremony, through land stewardship, science, art, and in everyday acts of practical reverence. (p. 190)

Gopnik (2010) referred to the kind of acts that middle school students are capable of bringing into existence as counterfactuals (p. 110). Pipher (2013) described those who take responsibility as "the most practical among us who come out of denial first" (p. 6). As a science teacher and environmental educator, the study encouraged me to change my mental models of equity and consider potential unintended consequences of using a scoring guide to assess environmental literacy.

Instructional relationships that acknowledge growth. In my narrative, I expressed appreciation for an experience when a teacher assessed my knowledge of a swath of forest at school. He received my inability to identify and categorize a particular species of vine with kind regard for my persistent observations. He honored my effort to speak for conserving the forested land by acknowledging my reasoning skills in a progress report to my parents. He modeled the kind of teaching practice that acknowledges another of Gough's (2013) principles: "to recognize that knowledge is partial, multiple and contradictory" (p. 10). Because the systems thinking scoring guide created in this study encouraged students to refine and propose changes, it acknowledged that learning and demonstrating environmental literacy was not a static process measured by the demonstration of skills upon demand. I became what Chödrön (2002) describes as a “devoted” student, because my teacher showed “compassion” in his assessment of what turned out to be my deep curiosity, which was rooted in my 12-year relationship with a temperate rainforest between home and school. I believed Gopnik’s (2010) premise that “counterfactuals let us change the future,” and have learned to acknowledge the strength that middle school students have for imagining how the context in which they live could be different (p. 23). After reflecting on how I had been assessed as a student, I realized that it did not require proficiency with every decision-making skill to act on the world in a way that makes possibilities true.

Standing midstream with Native American youth in an urban creek, I quickly realized I never wanted to be an expert again. Cajete (2012) was absolutely right about my mental model of problem solving, “where schooled ‘experts’ observe a reality or situation at a distance, then develop a solution or dictate an action or policy . . . [which] decontextualizes the problem from the totality of human experience and leads to a distorted perspective of the problem as an event that has relationship only to itself and to nothing else” (p. 47–48). As I traced the vision my grandfather had for schooling back to the Smith-Hughes Act, I realized that the federal funding for vocational education in my parents’ rural Washington community had actually oppressed students of the Kalapuya in Oregon. Vocational skills were taught rather than general education and English, and students were punished for using indigenous language and skills critical to making decisions. I learned that leveraging resources to support one element in a system had opposite effects on other parts of systems. My cultural definition of problem solving led to the ethical conundrum of whether the rights of the individual were more important than the needs of the group, and vice versa. The story of the inundation of my mother’s childhood home, and the valley where indigenous families live, in favor of generating power for the city was not a single event simply solved by moving my grandparents and their cows into town. Paying for a home underwater did not bring a displaced community into balance; it simply created a new kind of normal that threatened access to intergenerational knowledge. Gough (2013) was right; my experience showed racism in environmental education (p. 10). Decision-making is to be accepted as a deeply held cultural skill, and the scoring guide showed agreement for the construct of proficiency with systems thinking. Educators demonstrated that it could be used reliably, but my experience indicated that a number of mental models for assessment, equity, and environmental literacy would best be made explicit as part of the conditions for its use.

The scoring guide for systems thinking created in this study identified a set of skills that values a middle school student for assuming the responsibility to make decisions, and act by sharing one's ideas and observations with others. (See Appendix J for the full scoring guide, and the Tables 19-22, which present each set of systems thinking skills step by step.) Integrating environmental literacy with academic standards that specifically shift the burden for decision-making on "insiders" or "outsiders" creates discord in educators who practice in formal and non-formal settings. The more fruitful feedback would come from a student in conversation with educators from both perspectives. Stories from my lived experience as a formal and non-formal educator indicate that I have been searching for antidotes to the common assessment traps. Educators, who participated in the field-test identified competition, completion, and correction (the "need" to have the right answer), too. Even still, the systems thinking scoring guide created by middle school students, formal, and non-formal educators, which is considered here in more detail, provides a trustworthy map.

Modeling and analysis with systems thinking. In the Next Generation Science Standards, I learned that it was essential to give students multiple opportunities to refine and reconstruct models for their ideas of natural phenomena. In systems thinking, I learned that students show how changes in behavior over time could impact many different elements in a system at once. Using a Connection Circle. In teaching engineering design, I found that the application of one's models for natural phenomena also defined the limits of students' creativity and became a driver moving them toward new, and more diverse perspectives. The systems thinking scoring guide created by the forum began with a best practice in science: making a claim using evidence (see Table 19). It echoes traditional scientific methods of making predictions and proposing solutions. By focusing on creating solutions for systems that are not in balance, the scoring guide directed middle school students' attention to the kind of science that works with others to transform the world for a future that is not magical for all its wonder, but a future that is wonderful for its ability to function.

Instructional relationships. In terms of the instructional relationships needed to use the systems thinking scoring guide, my experience suggested it takes time and attentiveness to become attached to the "inner workings" of a temperate forest. The relationship between the forest, instructor, and student suggested that any natural system, when seen using the "big picture," will have many purposes, not all of which support only humans (see Table 20). I agree that the term *function*—rather than *purpose*, as

Table 19

Modeling and Analysis Skills in Step One of Systems Thinking Scoring Guide

Modeling and Analysis
Make a claim using evidence, and provide your reasoning orally and in writing. Construct an argument from analysis of data.
Predict how changes in one part of a system could affect the rest of the system. Identify variables and differing outcomes with changes to variables.
Create solutions for systems that are not in balance or unsustainable.
Explain if relationships are “correlation” (a mutual relationship between two things) or “causation” (one action causes another).

Note: For the complete table showing all four steps in *A Scoring Guide for Systems Thinking an Oregon Environmental Literacy Strand*, see Appendix J.

recommended by Meadows (2008)—be used to encourage mental models that move beyond using natural systems for anthropocentric purposes (p. 15). Expect failure and respond with compassion for students. “Presenting the complex inner workings of a system,” like figuring out which macroinvertebrates fish eat, takes as much persistent curiosity as science. Science serves as a gift to limit us from hurting ourselves, since our communities interact with natural systems. Technology was a tool to extend our senses and provide the information, when needed, to make decisions. The inner workings of systems and their self-organizing principles, draw attention to short term and long-term consequences.

Counterfactual futures. Situating a lab-based science classroom to reflect mental models for education held by families in community, and their middle school aged loved ones, was a more daunting a task than I ever imagined. I began to notice the dis-service I was doing to students when I resorted to canned labs that focused on known relationships

Table 20

Systems Habits Skills in Step Two of Systems Thinking Scoring Guide

Systems Habits
Identify long and short-term consequences.
Identify the purpose of the system and why it is important.
Identify long and short-term consequences.
Identify leverage points with greatest impact. Suggest how to use leverage to affect the system.
Develop models. Use an Iceberg model to show what people already know, and the bigger picture of the system.
Show how a system's structure generates its behavior.
Present the complex inner workings of a system in a simple and succinct way.
Make connections between the parts of the system and their outcomes, highlighting the interdependence of each part to make a whole.
Show how elements in the systems change over time. Track changes over time

Note: For the complete table showing all four steps in *A Scoring Guide for Systems Thinking an Oregon Environmental Literacy Strand*, see Appendix J.

of cause and effect. A strategy that turns science into a vending machine for getting predetermined results. Teaching optimization rather than bigger, faster, and better—and ensuring the student always defined "better" in terms that could be measured—were two other science practices I emphasized. Other practices emerged as I tried to leave students with an understanding of science as a dynamic balance between knowing and not knowing. The interactions of our body systems with the environment and one another was the greatest gift of science, and imagining the possible outcomes of our actions gave everyone a curious sense of purpose.

I used to work at a camp where the cook never learned to use the dishwasher. I suspected that she could have easily learned to use the dishwasher, but her function was to cook and feed, not to clean and set the tables. It was one of those systems where there was plenty of work to be done, and people could choose not to learn so they would not have to take responsibility. I was eager to move, eager to interact with people, and basically eager to have an integral function in the system. I learned how the dishwasher worked. I kept many dishwashers, and myself, working through many thousands of plates in many communities, even communities where people accidently left their false teeth in the coffee cups on their trays. I loved them anyway. One makes decisions. Even deciding not to understand, or deciding not to interact, is a decision. In practice, simply acknowledging that middle school students were capable of making decisions acknowledged the crucial function they have in community and natural systems. It does not take an expert to wash the dishes.

Educator as anthropologist. As an educator, I learned to switch perspectives from the role of outside, objective observer—which is usually associated with evaluation—and instead become a participant, by resourcing counterfactual models for stocks and flows proposed by the students. However, the limits of a system as expressed through intergenerational knowledge and science alone can guide the students' efforts when made explicit. The difference between causation and correlation is worth noting, but the student may quickly learn this on their own by making multiple iterations of their models.

Student sovereignty over context. The teacher, who assessed my knowledge of the natural history of a swatch of woods, created an assignment that matched what I already knew I knew about temperate forests. It mattered greatly to me that he valued the curiosity that led me to that forest, and that he showed that value by assigning academic credit to my field notebook and interpretive conversation between the biotic and abiotic interrelationships. It was the first time my skills and interests seemed to be of value to the community, and it was at a time when I was beginning to learn to care for those who cared for me. It was just one hour in the woods, shared with an adult who seemed to use his observations to make decisions of consequence. In the scoring guide for systems thinking, the responsibility for "communicating one's ideas, and remaining available for questions" were placed on the student. Each student has a unique perspective from which they can create antidotes to unsustainable mental models, even in situations where others presume an antidote does not exist. Students were asked to design "creative solutions" that reflect a range of communities with which they identify and connect (see Table 21).

Table 21:

Problem Solving Skills in Step Three of Systems Thinking Scoring Guide

Problem Solving
Identify the problem of a situation.
Create solutions that could mitigate the problem, and predict how changes to the system could emulate those solutions. Make inferences from experience.
Use creative thinking. Use the design cycle.
Use dynamic modeling with stocks and flows, and change variables until the desired outcome is achieved.

Note: For the complete table showing all four steps in *A Scoring Guide for Systems Thinking an Oregon Environmental Literacy Strand*, see Appendix J.

Deep culture decision-making. In my experience as a formal educator, I have found that decision-making skills improve not only from presenting multiple sides of an argument, as required in professional practice, but also in conversations that cross culturally familiar boundaries, languages, and mental models. The language of crisis and controversy tends to send people to safe, familiar places in their minds and forces them to take sides. The scoring guide placed the responsibility for communication on the one proposing changes and imagining outcomes (see Table 22). To be proficient, one was asked to share ideas "in a way that people will understand you." At middle school age, students can identify a culture's hypocrisy and injustice. For example, in Oregon, there are as many generations of farmworkers as farmers. Generations of indigenous people live reciprocally with indigenous plants and animals in Oregon, too. The scoring guide

Table 22

Refining Skills in Step Four of Systems Thinking Scoring Guide

Refining and Proposing Changes
Display proposed changes and outcomes via easily understood diagrams. Use visual graphic skills to clearly present how the changes effected the environment.
Analyze data.
Develop a sound model. Suggest corrective actions by finding leverage points and making slow gradual changes.
Collaborate. Use the design cycle to explore multiple solutions for the same problem and create a +/- chart for each solution.
Use strong presentation skills to share ideas in a way that people will understand you. Vocalize the proposed changes and answer questions pertaining to the ideas shared.

encouraged collaboration and careful exploration of multiple solutions, which was at the core of my experience of teaching. I applauded the forum that developed the systems thinking scoring guide for identifying deep culture decision-making skills. My experience in outdoor education affirmed that group success and individual success are one in the same. A sound model was defined as one that suggests corrective actions by finding leverage points and making slow gradual changes. As Gough (2013) indicated, environmentally literate decisions are not "culturally-blind" (p. 10). My experience suggests that decisions require intercultural communication skills so people can shift from historical mental models that do not consider the long-term consequences of championing the golden man with the ax over the woman with the basket (p. 10).

Recommendations for Teacher and Administrator Preparation

Teachers and administrators need to be taught systems thinking skills so they can identify antidotes to assessment traps. The feedback from educators who field-tested the scoring guide resonated with themes related to assessment, equity, and environmental literacy revealed in my autoethnographic narrative:

- Competition: one student must be better than another, and others must be worse.
- Completion: to be proficient, the students' work must include every skill.
- Correction: overemphasis on solutions to problems, which often leads to an assumption there is only one answer.

My recommendation for policy and practice would be to continue the discussion of what it means for educators to be environmentally literate in a system that provides parity to formal and non-formal educators. Each day formal and non-formal educators are working together to kick denial to the curb and doing the job of interacting with middle school students in community with natural systems. Studies that give communities the opportunity to witness middle school students in the act of making decisions, refining their ideas, and acting on their beliefs in the places they live now, can serve as antidotes to those traps identified by educators who participated in this study. Educators can use systems thinking to leverage competitive forces into balance so their combined efforts demonstrate Gough's (2013) environmental principles for educators,:

> To recognize that knowledge is partial, multiple, and contradictory; to draw attention to racism and gender blindness in environmental education; to develop a willingness to listen to silenced voices and to provide opportunities for them to be heard; and to develop understandings of the stories of which we are a part and our abilities to deconstruct them. (p. 10)

One provocateur in particular, from Oregon, spoke directly to the actions that educators can take to deconstruct mental models for assessment that are fueled by a level of competition that exceeds the amount of resources available.

Competitive force. Zhao (2015) helped me to understand that competition was one of the "most destructive forces in the universe" and challenged me "not to fix deficits according to external standards, but to help each person become successful in their own way" (Y. Zhao, personal communication, July 2015). It came to my attention that feeling like a vending machine, dispensing scores on assignments, was a symptom of a mental model that still assumes a group of students must be rated in comparison to others using the variation shown in a bell-shaped curve. Zhao provided me with the antidote when he said, "Every individual has talent, and every individual has passion." My antidote was to treat each student's schoolwork as a gift, and to treat each assessment as a snapshot in time, whose purpose was to give each student an opportunity to reflect on the fact that *they knew that they knew* a particular concept or could demonstrate a particular skill. Gough's (2013) four guiding principles for educators encouraged me to consider whether my use of a systems thinking scoring guide for environmental literacy demonstrated "a willingness to listen to silenced voices and to provide opportunities for them to be heard" (p. 10).

Implications for Environmental Literacy Policy

The timeframe for my dissertation paralleled the political cycle for the revision of national and state educational policy. In 2008, the U.S. House of Representatives passed the No Child Left Inside Act (H.R. 3036) to restore environmental education to its nation's test-heavy accountability system outlined in the Elementary and Secondary Education Act named No Child Left Behind. The final artifact that I uploaded for data analysis came from the No Child Left Inside Coalition website in July 2015, indicating that the U.S. House and Senate passed the Elementary and Secondary Education Act (ESSA) or Every Child Achieves Act, with the following provisions:

- Environmental science education would be an allowable subject included in the STEM (science, technology, engineering, and math) grants.
- Environmental literacy would be part of the after-school programming sections of the bill.

Their December 2015 update indicates that President Obama signed the Every Child Achieves Act into law, and Title IV indicates the following:

- Environmental education is called out as eligible for funding under a $1.6B "well-rounded education" grants program.
- Environmental literacy programs are eligible for funding as part of the $1B 21st Century Community Learning Centers program.
- The prioritization of STEM activities, including "hands-on learning" and "field-based or service learning" to enhance understanding of STEM subjects, may provide additional opportunities for environmental science education programs.

It is my understanding that because the state of Oregon's legislature passed the Oregon No Child Left Inside bill (H.B. 2544) in 2009, and adopted the Oregon Environmental Literacy Plan in 2010, that the state can apply for federal funds provided in the ESSA.

From a policy perspective, the language of the 2015 Every Child Achieves Act categorizes national funds for education in a way that perpetuates artificial boundaries between formal and non-formal educators. The consequences of this policy decision are unfortunate. Funding for environmental education, or environmental science education efforts that include STEM subjects, field-based and service learning opportunities is greater. Environmental literacy programs are associated with 21st Century Community Learning Centers described as after school programming.

Recommendations for Rule Change. In 2015, the Oregon Board of Education removed "Outdoor School" from a list of examples of appropriate uses of instructional time from OAR 581-022-0102(30) in the process of doing rule work. Based on findings from evaluating an assessment instrument for the Oregon Environmental Literacy Plan, which involved listening to middle school students, formal and non-formal educators equally, my recommendation would be to propose a rule for instructional time to teach and assess the environmental literacy strands adopted in the OELP as an integral component of science at the middle school level. In practice, this would involve school districts ensuring 100% of all students are scheduled to receive a minimum of 40 hours of instruction per year in environmental literacy, which must occur outdoors near the school, or at other field sites under the direction of educators with environmental literacy certification. Students could also receive credit for proficiency by presenting evidence of skills with systems thinking. Such a rule would accredit environmental literacy based on proficiency, and create an environmental literacy career path for the licensing and salary of educators working for environmental literacy commensurate with the level of responsibility they assume in the relationship of *loco parentis* at schools, field study sites, or learning centers. Such a ruling would value the work of those in public schools as well as those who work with educators in schools to thinking systematically about the consequences of decisions on the natural systems and communities in which we live. So, although I hesitate to suggest policy based solely on the findings of this study to evaluate an assessment instrument for the OELP, I do feel strongly based on my experience that leveraging political, economic, cultural and social systems would prevent educators from drifting to low performance, and continue to move educators towards a mutual goal of helping students recognize they are capable of "using understanding, skills and motivation to make responsible decisions that consider [one's] relationships to natural systems, communities and future generations" (OELP, 2010, p. 4).

The role that teaching and assessing environmental literacy with parity between educators in school buildings and educators in the field, or learning centers, would be to acknowledge the strengths of each, or in the words of W.E.B. Du Bois, "Give each to each those characteristics both so sadly lack (Du Bois, 1903, p. 130). Middle school students who learn the limits of natural systems with environmental scientists might be more likely to employ the Precautionary Principle that limits the designs they engineer. Those students, who work with scientists and electron microscopes to understand the smallest systems, may be able to prevent the overshoot of scaling systems up because they can also see the bigger picture using systems thinking. Students who spend time in the field and serving in community would recognize the role that time delays play in systems and how each generation's incremental changes influence the next generation's decisions. Middle school student would understand the responsibilities associated with the function they have in natural systems that thrive on epistemic trust, the kind of trust that our grandparents knew with the natural systems in which they also lived.

I now understood why the unfunded Outdoor School for All bill (S.B. 439) unanimously passed with bipartisan support: trust and hope. The Oregon Legislature assigned responsibility to Oregon State University Extension Service (OSU Extension) for assisting school districts and educational service districts in providing outdoor school programs. OSU Extension was assigned the responsibilities of "administering a grant program, providing program leadership and providing program maintenance" (S.B. 439, p. 1). My experience with the politics of educational funding for outdoor school started to follow the funding stream paddled by my grandfather, who graduated from OSU as an agricultural teacher, his education having been paid for with federal funds from the Smith-Hughes Vocational Education Act of 1917, which allocated funds for building industrial arts classrooms and paying vocational teacher's' salaries. When my grandfather was selected to the Vocational Agricultural Department for a rural school district, his duties included: high school principal, basketball coach, science, manual training, and vocational agriculture. These were the commonalities valued between generations in my family: vision, competition, knowledge of how the natural world works, building, and growing food. I wondered what the men who argued Smith-Hughes into federal funding over a hundred years ago would have thought of my grandfather's efforts to build a school system that embodied their vision. I worry today for the unintended consequences created by the artificial boundaries between environmental science education and environmental literacy in the national Elementary and Secondary Education Act in Oregon at a time when educators are crossing boundaries to implement the Oregon Environmental Literacy Plan.

Charles Prosser, executive secretary of the National Society for the Promotion of Industrial Education, was an author of the Smith-Hughes Act. The Act was proposed to "meet a compelling need of the new America—the need to provide American industry with the complicated work skills required in a technological society" (Wirth, 1972, p. 2). Prosser was a physics, chemistry, and literature teacher from a steelworker family in New Albany, Indiana, and "he was a proponent of the social efficiency philosophy" (Wirth, 1972, p. 1). According to Wirth (1972), Prosser's explained:

> The purpose of vocational education is to help a person secure a job, train him so he can hold it after he gets it, and assist him in advancing to a better job. . . . Training for useful employment. (p. 3)

Prosser's definition of vocational education was intentionally in sharp contrast to what Wirth (1972) termed "traditional scholastic education" though it surely sounded like a plan to keep a person stable and able (p. 3). Additionally, Prosser understood the mission of vocational education was to "establish habits of correct thinking and doing" so the minds of the students were treated as a "habit forming machine" (Wirth, 1972, p. 3). Students would be able to learn these habits best in a learning situation that was as close to the real situation as possible. He suggested that tasks related to a trade should "be taught by the craftsman-teacher skilled in the task, rather than by general mathematics or science teachers" (Wirth, 1972, p. 3). Prosser's narrowing of the curriculum did not resonate with the "broader, deeper, higher culture of gifted minds and pure hearts" called for by Du Bois (1903, p. 130). It made me wonder what the curriculum would be like if it were taught by nature itself—the trees skilled in the task of sequestering carbon, and ancient soils skilled in the task of collaborating with sun and insects to transform nitrogen into a form that can be absorbed by roots. I hoped the systems thinking scoring guide would free a student to be sovereign over, and explicit with, their decision-making skills, rather than becoming a "habit forming machine" (Wirth, 1972, p. 3).

The 2016 initiative petition to propose the use of unassigned lottery dollars to fund a full week of Outdoor School for every 5th or 6th grade student in Oregon is a resonates with the voices of Washington, Du Bois, and my grandfather:

> Outdoor School is a proven answer. Kids come alive—curious and engaged—when they get outdoors. Outdoor School is a smart, time-tested, hands-on science-based week of solid effective education. Breathing fresh air, surrounded by wonder, collaborating with other kids builds confidence and self-sufficiency as kids learn to value and make responsible choices about our incredible natural resources. . . . The benefits of Outdoor School are clear: strengthening today's economy, creating tomorrow's leaders, and preserving for all time the natural resources and natural legacy that make us who we are—Oregonians. (Outdoor School for All, 2016)

According to the campaign brochure, the measurable benefit 10 years from approval would be seen in the form of a "$270 million economic impact" (Outdoor School for All, 2016). The campaign defined Oregon as a "stunning tapestry of natural wonders and natural riches" (Outdoor School for All). The political implication of using a systems thinking scoring guide for assessing environmental literacy reminded me of the potential for it to be misunderstood as simply a tool for retelling the story of the golden man with the ax. On the other hand, I found solace knowing that middle school students might be more concerned with the population of slugs they could attach to their jacket sleeves and safely return to the humus once the honeymoon was over. Since 1957, Outdoor School has given Oregon's middle school students a week in which their lives are structured intentionally around not only making friends with people from across town by sharing a cabin, but soil, water, forest, and wildlife as well. This vision left a deep enough impact such that complete strangers form a strong bond at the mention outdoor school when asked to sign a petition. Adding an assessment instrument to measure the impact of outdoor school, and other equally valuable programs that bring natural systems and communities to mind, could make environmental literacy and science both integral aspects of decisions that affect middle school students now and in the future.

Implications for Practice

I wondered if my mental models for decision-making were any different than my grandfather's, which was equally possessed of the poetry of Walt Whitman and the pragmatism of feeding a family. I worried whether school districts would work with educational service districts to assess environmental literacy without perpetuating a mental model that oppressed people like the Kalapuya by teaching vocational skills over English, and punishing students for using indigenous language and skills critical to making decisions which have sustained cultures and natural systems for centuries. Will educators find the language systems thinking too difficult to teach, and instead repeat the history of vocational educators in the past that oppressed others by lowering their goals? I stumbled between the language of quantitative and qualitative mental models, and the function of environmental science and environmental literacy. I wondered if my colleagues voluntarily adopt grandmothers and grandfathers, neighbors and friends, who can teach them to know that the Earth loves you back when you give and receive one another's gifts with respect for one another's regenerative limits. As I reviewed the systems thinking scoring guide created by the forum of middle school students, formal educators, and non-formal educators for the Oregon Environmental Literacy Plan, I found the fact that they agreed upon the importance of multiple perspectives, collaboration, and speaking so people can understand you critical to making decisions. Given the hundreds of decisions I must make as an educator each week, the importance of working together towards meeting standards is important. I live for the days outside, too, where people can hear the owls respond with gratitude for reaching our goal of "aligning human behaviours, actions, practices and social conditions towards a sustainable future" (*Tbilisi Communiqué*, 2012, p. 1)

Conclusion

Nagel (1996) wrote of "the importance of acknowledging how people learn and how people use their knowledge in life" (p. 150). She called this model of integrated teaching the "real-world problem-solving process," which "promotes the active integrative sense of learning through student-directed work on issues of concern to both the students and the larger community" (p. 150). She explained that authentic assessment should bring students and teachers "together to develop criteria for determining satisfactory performance" (p. 107). Stevens and Levi (2005) concurred with involving students in constructing rubrics (or scoring guides):

> Surprise rubrics happen when we grade an assignment with a rubric that students have never seen before, and then hand back the assignment with the rubric attached. When this occurs, students are justifiably miffed. . . . Involving students lets us share the "burden of explanation" with them and we are no longer alone in explaining how to complete an assignment. (p. 50)

In order to fairly assess environmental literacy, my autoethnographic narrative indicated that we not only need to add the voices of students, but also the voices of what the Oregon Biodiversity Project refers to as Oregon's "living landscape" (1998). Where cedar trees assess the skills of middle school students to not simply "develop and dictate an action or policy" but to "establish a direct participatory relationship with the natural, cultural and historical reality in which one lives" (Cajete, 2012, p. 47–48). I realized that the scientists in my life were those who defined the steps and skills of the natural systems that sustained me: the vocational teacher who took me to a fallen cedar that had been transformed to a nurse log; the farmer who taught us how to dibble the soil for planting genetically modified and non-genetically modified seeds; shelling walnuts and baking bread with my grandmother; shooting a bow with my cousin; and being fully present to the forces of life and death in the wilderness and the city. When I was with these intergenerational teachers, they seemed to have an internal scoring guide that rubbed off on me.

Ever since I could use language, the word "skookum" was used to describe the moments when everything was balanced and as peaceful as highly functioning systems can be when their elements have limitations. With this writing, I learned that the word did not travel to the Pacific Northwest with my great-grandmother in the belly of a ship from the Netherlands. The use of common jargon to create shared meaning was a gift from the First Nations who lived with the salmon and helped my pioneer ancestors understand unfamiliar systems. According to Native Languages of the Americas, the word "skookum" was used to describe the powerful, man-eating forest creature also referred to as Sasquatch. This example of my poor cross-cultural understanding of the construct of balance revealed that what seemed to be just right for me was actually a powerful, man-eating creature that put others' lives at risk. Jackson (2015) showed me that decision-making is at the bottom of the Cultural Iceberg, along with our understanding of fairness and justice, where learning concepts naturally leads us deeper into a place of choice. Meadows (2008) cautioned that the leverage points where we choose to intervene in systems to solve problems find us "pushing with all our might in the wrong direction" (p. 146). She writes:

> Insistence on a single culture shuts down learning and cuts back resilience. Any system, biological, economic, or social, that gets so encrusted that it cannot self-evolve, a system that systematically scorns experimentation and wipes out the raw material of innovation, is doomed over the long term on this highly variable planet. (p.160)

I wonder if the golden man on top of Oregon's capitol building laid down his ax, and took stock of the future he helped to create, what skills he would use to make decisions with natural systems and community in mind. Would he encourage the variability and diversity that Meadows (2008) suggested keeps systems in balance, or remain unaware of the unintended consequences of acting without first climbing his Ladder of Inference (p. 160)? Would he stop and ask himself about the meaning he attaches to his experience, and how he chooses to act based on those beliefs? Middle school students, formal educators, and non-formal educators share a high level of consensus for the construct of proficiency with systems thinking skills. Educators can use a scoring guide to recognize these skills in a written sample of students' work that uses systems thinking tools (Connection Circles, Behavior-Over-Time Graphs, Causal Loops, and Ladders of Inference), with a moderate level of reliability, even though they have had no training. Based on my lived experience, the function of an insatiable curiosity for understanding natural phenomena, including our interactions with one another, is a precautionary, intergenerational story of acting in dynamic balance. As an educator, assessing students in the context of their lives means acknowledging those moments that they know that they know and can do what they envision. Being there in that memory matters. Being a vending machine, where students add assignments for credit, can quench their curiosity for the moment. But know that the thirst that comes with what Senge (2014) described as caring will return, and students will need to make decisions with natural phenomena and communities in mind—systems that include middle school age youth and their voices. My grandmother summed it up for students pretty well: “You fall down. You can’t get up. Someone helps you up. You are still a good person, no matter what, and don’t let anyone ever tell you any different!”

References

Absolute Astronomy. (2013). Re: An Inconvenient Truth [Online forum]. Retrieved from http://www.absoluteastronomy.com/topics/An_Inconvenient_Truth

Andrews, M. C. (2013). *EUGENE: Ecological understanding as a guideline for evaluation of non-formal education* [Web-based assessment tool]. Retrieved from https://projecteugene.org/cgi-bin/eugene

Bartosh, O. (2003). *Environmental education improving student achievement* (Master's thesis). Retrieved from http://wa.audubon.org/PDFs /Education-research-Oksana-1-3.pdf

Bartosh, O., Tudor, M., Ferguson, L., & Taylor, C. (2009). Impact of environment-based teaching on student achievement: A study of Washington State middle schools. *Middle Grades Research Journal*, *4*(4), 1-16.

Brewer, J. (2006). The Multimethod Approach and Its Promise. In J. Brewer & A. Hunter (Eds.), *Foundations of Multimethod Research.* (pp. 2-16). Thousand Oaks, CA: Sage. Publications, Inc., doi:10.4135/9781412984294.n1

Bybee, R. W., Taylor, J. A., Garder, A., Van Scotter, P., Powell, J. C., Westbrook, A., Landes, N. (2006). The BSCS 5E instructional model: Origins and effectiveness (A Report prepared for Office of Science Instruction, National Institutes of Health). Colorado Springs, CO: Biological Sciences Curriculum Study. Retrieved from http://www.bscs.org/bscs-5e-instructional-model

Cajete, G. (2012). Contemporary indigenous education: Thoughts for American Indian education in the 21st century. In S. Mukhopadhyay & W. Roth (Eds.), *Alternative* *forms of knowing (in) mathematics: Celebrations of mathematical practices* (pp. 33–52). Rotterdam, The Netherlands: Sense Publishers.

Chang, H. (2008). *Autoethnography as method.* Walnut Creek, CA: Left Coast Press, Inc.

Chödrön, P. (2002). The spiritual friend. On *The places the scare you: A guide to fearlessness in difficult times* [CD]. Boulder, CO: Shambala Audio.

Cloud, J. P. (2012, November 5). Educating for sustainability with the brain in mind [Web log post]. Retrieved from http://cloudinstitute.org/blog/category/jaimie-cloud

Cutter-Mackenzie, A. & Smith, R. (2003). Ecological literacy: The 'missing paradigm' in environmental education (part 1). *Environmental Education Research*, *9*(4), 497-524. doi:10.1080/1350462032000126131

Doherty, J., Draney, K., & Anderson, C. K. (2012). Methodological issues in developing a learning progression-based assessment system [PowerPoint slides]. *National Association for Research in Science Teaching - NARST Conference,* Indianapolis, Indiana. Retrieved from http://edr1.educ.msu.edu/environmentallit/publicsite /html/report_2012.html

Du Bois, W. E. B. (1903). The souls of black folk. In C. Kridel (Ed.), *Classic edition sources: Education* (4th ed., pp. 127–131), New York, NY: McGraw Hill.

Elder, J. L. (2003). *A field guide to environmental literacy: Making strategic investments in environmental education*. Reston, VA: Environmental Education Coalition.

Environmental Education Resource Directory. Retrieved from http://directory.eeao.org/ organizations

Excellence in Environmental Education: Guidelines for Learning (K-12). (2010). North American Association for Environmental Education, Washington, D.C. Retrieved from http://resources.spaces3.com/89c197bf-e630-42b0-ad9a-91f0bc55c72d.pdf

Fasheh, M. (2015). Over 68 years with mathematics: My story of healing from modern superstitions and reclaiming my sense of being and well-being. In S. Mukhopadhyay and B. Greer (Eds.), *Mathematics education & society: Proceedings of the eighth international mathematics education & society conference: Vol. 1.* (pp. 33–60). Portland, OR: Ooligan Press, Portland State University.

Farmer, J., Knapp, D., & Benton, G. M. (2007). An elementary school environmental education field trip: Long-term effects on ecological and environmental knowledge and attitude development. *Journal of Environmental Education*, *38*(3), 33-42.

Fonagy, P., Luyten, P., Campbell, C., & Allison, L. (2014, December). Epistemic trust, psychopathology and the great psychotherapy debate. [Web Article]. Retrieved from http://societyforpsychotherapy.org/epistemic-trust-psychopathology-and-the-great-psychotherapy-debate/

Friends of Outdoor School. (2013). *HB 2516 – Outdoor school for everyone.* Retrieved from http://www.friendsofoutdoorschool.org/HB-2516.html

Friends of Outdoor School. (2013). *HB 2516 – Why outdoor school?* Retrieved from http://www.friendsofoutdoorschool.org/Why-Outdoor-School.html

Glenn, J. L. (2000). *Environment-based education: Creating high performance schools and students* [Report]. Washington, DC: National Environmental Education and

Training Foundation.

Golley, F. B. (1998). *A primer for environmental literacy*. New Haven, CT: Yale University Press.

Gopnik, A. (2009). *The philosophical baby: What children's minds tell us about truth, love and the meaning of life.* New York, NY: Farrar, Straus and Giroux.

Gotwals, A. W. & Songer, N. B. (2013). Validity evidence for learning-progression-based assessment items that fuse core disciplinary ideas and science practices. *Journal of Research in Science Teaching*, *50*(5), 597–626. doi:10.1002/tea.21083

Gough, A. (2002). Mutualism: A different agenda for environmental and science education. *International Journal of Science Education*, *24*(11), 1201-1215. doi:10.1080/09500690210136611

Gough, A. (2013). Historical, contextual, and theoretical orientations that have shaped environmental education research: Introduction. In R. B. Stevenson, M. Brody, J. Wals, & E. J. Arjen (Eds.), *International handbook of research on environmental education* (pp. 9-12). New York, NY: Routledge. eISBN: 9781136699313.

Gough, A. (2013). The emergence of environmental education research: A history of the field. In R. B. Stevenson, M. Brody, J. Wals, & E. J. Arjen (Eds.), *International handbook of research on environmental education* (pp. 44-62). New York, NY: Routledge. eISBN: 9781136699313.

Guba, E. G. & Lincoln, Y. S. (1998). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *The landscape of qualitative research: Theories and issues* (pp. 195- 220). Thousand Oaks, CA: Sage.

Guba, E. G. & Lincoln, Y. S. (2005). Paradigmatic controversies, contradictions, and emerging confluences. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (3rd ed.) (pp. 191-216). Thousand Oaks, CA: Sage.

Harlen, W. (2001). The assessment of scientific literacy in the OECD/PISA project. In H. Behrendt, H. Dahncke, R. Duit, W. Gräber, M. Komorek, A. Kross, & P. Reiska (Eds.), *Research in science education - Past, present, and future* (pp. 49-60). Dordrecht, The Netherlands: Kluwer Academic.

Hancock, J. (2005). Taking stock of research and pedagogy in L2 writing. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 597-613). Mahwah, NJ: Lawrence Erlbaum.

Haluza-DeLay, R. (2013). Educating for environmental justice. In R. B. Stevenson, M. Brody, J. Wals, & E. J. Arjen (Eds.), *International handbook of research on environmental education* (pp. 394-403). New York, NY: Routledge. eISBN: 9781136699313.

Hollweg, K. S., Taylor, J. R., Bybee, R. W., Marcinkowski, T. J., McBeth, W. C., & Zoido, P. (2011). *Developing a framework for assessing environmental literacy*. Washington, DC: North American Association for Environmental Education. Retrieved from http://www.naaee.net/sites/default/files/framework /DevFramewkAssessEnvLitOnlineEd.pdf

Hollweg, K., Taylor, J., McBeth, B., Bybee, R., Taylor, J., Marcinkowski, T., & Zoido, P. (2011, December). *Developing a framework for assessing environmental literacy.* Slideshow presented at the Developing a Framework for Assessing

Environmental Literacy meeting, National Press Club, Washington, DC. Retrieved from http://www.naaee.net/framework

Homestead Act of 1862. (n.d.). *Major acts of Congress*. Retrieved August 11, 2009, from http://www.answers.com/topic/homestead-act-of-1862

International Baccalaureate. (2008). MYP: From principles to practice. *Middle Years Programme.* Cardiff, United Kingdom: International Baccalaureate.

International Baccalaureate Middle Years Programme. (2013). Retrieved from http://www.ibo.org/myp/index.cfm

International Baccalaureate Organization. (2016). *Appendix 3: MYP Command Terms.* Retrieved from https://ibpublishing.ibo.org/server2/rest/app/tsm.xql?doc= m_0_mypxx_guu_1409_1_e&part=9&chapter=3

Juntunen, J. R., Dasch, M. D., & Rogers, A. B. (2005). *The World of the Kalapuya: A native people of Western Oregon.* Philomath, OR: Benton County Historical Society and Museum.

Kearney, A. R. (2009). *IslandWood evaluation project: Assessment of student outcomes from IslandWood's school overnight program.* Bainbridge Island, WA: IslandWood. Retrieved from http://meera.snre.umich.edu/reports-and-case-studies/islandwood-evaluation-project-assessment-student-outcomes-islandwood’s

Keeny, S., Hasson, F., & McKenna, H. (2011). *The Delphi technique in nursing and health research*. West Sussex, UK:Wiley-Blackwell.

Kent, K. (1968, renewed 2001). The paradoxical commandments. Retrieved from http://www.paradoxicalcommandments.com/

Kimmerer, R. W. (2013). *Braiding sweetgrass: Indigenous wisdom, scientific knowledge, and the teachings of plants.* Minneapolis, Minnesota: Milkweed Editions.

Kleckner, C. (2010). *Moving science education forward – Day 1* [PowerPoint slides]. Retrieved from http://www.ode.state.or.us/search/page/?=2978

Kriebel, D., Tickner, J., Epstein, P., Lemons, J., Levins, R., Loechler, E. L., et al. (2001). The precautionary principle in environmental science. *Environmental Health Perspectives*, *109*(9), 871–876.

Landis, J. R. & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159-174. Retrieved from http://www.jstor.org/stable/2529310

Lieberman, G. & Hoody, L. (1998). *Closing the achievement gap* [Executive summary]. Retrieved from http://www.seer.org/extras/execsum.pdf

Maddock, T. (1999). The nature and limits of critical theory in education. *Educational Philosophy and Theory*, *31*(1), 43-61. doi:10.1111/j.1469-5812.1999.tb00373.x Retrieved from http://www.tandfonline.com.proxy.lib.pdx.edu/loi/rept20

Marcinkowski, T., Bucheit, J., Spero-Swingle, V., Linsenbardt, C., Engelhardt, J., Stadel, M., & Guzmon, K. (2012). Selected trends in thirty years of doctoral research in environmental education in *Dissertation Abstracts International abstracts* from collections prepared in the United States of America. In R. B. Stevenson, M. Brody, J. Wals, & E. J. Arjen (Eds.), *International handbook of research on environmental education* (pp. 44-62). New York, NY: Routledge. eISBN: 9781136699313

McBeth, B., Hungerford, H., Marcinkowski, T., Volk, T., Cifranick, K., Howell, J., & Meyers, R. (2011). *National environmental literacy assessment phase II: Measuring the effectiveness of North American environmental education programs with respect to the parameters of environmental literacy* [Final Research Report]. Retrieved from North American Association of Environmental Education website: http://www.naaee.net/sites/default/files/programs/research/NELA_Phase_II_Report.pdf

McLeod, S. (2007). *Lev Vygotsky.* Retrieved from http://www.simplypsychology.org/vygotsky.html

McCrae, E.J. (2010, December). Leading the way to environmental literacy and quality: National guidelines for environmental education [Post Media]. Retrieved from http://eelinked.naaee.net/n/guidelines/posts/National-Project-for-Excellence-in-Environmental-Education

Meadows, D. H. (2008). *Thinking in systems: A primer* (D. Wright, Ed.). White River Junction, VT: Chelsea Green Publishing.

Michigan State University. (2010). *MSP Carbon Assessment [Form A]*. Retrieved from http://edr1.educ.msu.edu/EnvironmentalLit/publicsite/files/CarbonCycle/CC%20Assessment/12-13/MSP%202011%20Carbon-A%20+%20QR.pdf

Michigan State University. (2010). *Environmental literacy.* Retrieved from http://edr1.educ.msu.edu/EnvironmentalLit/

Michigan State University. (2010). *Introduction of assessment*. Retrieved from http://edr1.educ.msu.edu/EnvironmentalLit/publicsite/html /assessment.html

Mingé, J. M. (2013). Mindful Autoethnography, Local Knowledges: Lessons from My Family. In S. Jones, T. Adams, & C. Ellis (Eds.), *Handbook of Autoethnography,* (pp. 425-443). Walnut Creek, CA:

Left Coast press, Inc.

Nagel, N. (1996). *Learning through real-world problem solving: The power of integrative teaching.* Thousand Oaks, CA: Corwin Press, Inc.

NASA. (2016). Carbon Dioxide. Retrieved from http://climate.nasa.gov/vital-signs/carbon-dioxide/

Native American words. Retrieved from http://www.native-languages.org/vocabulary.htm

NGSS Lead States. (2013). *Next Generation Science Standards*: *For States by States.* Washington DC: The National Academies Press. Retrieved from http://www.nextgenscience.org/next-generation-science-standards

NGSS Lead States. (2013). *NGSS Public Release II – Crosscutting Concepts in the Next Generation Science Standards*. Retrieved from http://www.nextgenscience.org

NGSS Lead States. (2013). *Next Generation Science Standards*: *For States by States.* Washington DC: The National Academies Press. Retrieved from http://www.nextgenscience.org/search/node/make%20a%20claim%20using%20evidence

NGSS Lead States. (2013). *Next Generation Science Standards*: *For States by States, 3-ESS2 Earth and human activity.* Washington DC: The National Academies Press. Retrieved from http://www.nextgenscience.org/dci-arrangement/3-ess3-earth-and-human-activity

NGSS Lead States. (2013). *Next Generation Science Standards*: *For States by States, 2-LS2 Ecosystems: Interactions, energy and dynamics.* Washington DC: The National Academies Press. Retrieved from http://www.nextgenscience.org/dci-arrangement/2-ls2-ecosystems-interactions-energy-and-dynamics

NGSS Lead States. (2013). *Next Generation Science Standards: For States, By States.* Washington, DC: The National Academies Press.

NGSS Lead States. (2013). *Next Generation Science Standards*: *For States by States, Appendix G - Crosscutting Concepts.* . Washington DC: The National Academies Press. Retrieved from http://www.nextgenscience.org/sites/default/files/Appendix%20G%20%20Crosscutting%20Concepts%20FINAL%20edited%204.10.13.pdf

Nicolson, M. (2013) *MYP: The next chapter* [Video file]. Retrieved from http://www.ibo.org/myp/index.cfm

Nuthalpaty, F. (2010). *Educational Research 102: Selecting the Best Study Design for your Research Question. APGO, Faculty Development Seminar.* Retrieved from http://www.slideshare.net/fnuthalapaty/educational-research-102-selecting-the-best-study-design-for-your-research-question-3062530

Oregon Biodiversity Project. (1998). Oregon's living landscape: Strategies and opportunities to conserve biodiversity. Defenders of Wildlife. Retrieved from https://openlibrary.org/works/OL8693690W/Oregon's_Living_Landscape

Oregon Department of Education, Oregon Environmental Literacy Task Force. (2011). Educational standards and graduation requirements. In Oregon Department of Education, [*Oregon Environmental Literacy Plan*]. Retrieved from http://www.ode.state.or.us/teachlearn/specialty/alt/eltf-oregon-environmental-literacy-strands---chapter-3-december-2011-rev-1.pdf

Oregon Department of Education. (2013, April 18). *Next Generation Science Standards released* [Announcements]. Retrieved from http://www.ode.state.or.us/news/announcements/announcement.aspx?=9013

Oregon Department of Education, Environmental Literacy Task Force. (2013). *Oregon Environmental Literacy Plan: Toward a Sustainable Future*. Retrieved from http://eeao.org/files/OELP_4-20-13.pdf

Oregon Department of Education, Environmental Literacy Task Force. (2010). *Oregon Environmental Literacy Plan*. Retrieved from http://www.ode.state.or.us/

Oregon Department of Education, State Board of Education. (2013). *State Board of Education Meeting Minutes for May 16, 2013.* Retrieved from http://www.ode.state.or.us/search/page/?=3849

Oregon Environmental Literacy Task Force, House Bill 2544. 75th Oregon Legislative Assembly, Regular Sess. (2009). Retrieved from http://www.leg.state.or.us/09reg/measpdf/hb2500.dir/hb2544.intro.pdf

Organisation for Economic Co-operation and Development. (2009). *Green at 15? How 15-year-olds perform in environmental science and geoscience in PISA 2006 assessment: 120* [OECD publication]. Retrieved from http://www.oecdilibrary.org/content/book/9789264063600-en

Orr, D. W. (1992). *Ecological literacy: Education and the transition to a postmodern world.* Albany, NY: State University of New York Press.

Osborne, J. (2007). Science education for the twenty-first century. *Eurasia Journal of Mathematics, Science and Technology Education, 3*(3), 173-184. Retrieved from http://www.ejmste.com/v3n3/EJMSTE_v3n3_Osborne.pdf

Pangrac, A. C. & Christensen, L. (2012). *Fifth and sixth grade student participation in outdoor school programs in Oregon.* [Report]. Portland, OR: Oregon Community Foundation. Retrieved from http://www.friendsofoutdoorschool.org/uploads/2012_ODS_Survey_Report_-_Final_Version.pdf

Paul F-Brandwein Institute. (2013). *The Paul F-Brandwein Lecture Series.* Retrieved from http://brandwein.org/node/1388#roth

Pipher, M. (2013). *The Green Boat: Reviving Ourselves in Our Capsized Culture.* New York, NY: Riverhead Books.

Portland Public Schools. (Producer). (2013). *A town hall meeting with Governor John Kitzhaber and the Portland Public Schools Board of Education: An important conversation about the future of public education in Oregon* [Video file]. Available from http://www.youtube.com/watch?v=niY1lSm9QfA

Rice, J., Doherty, J. H., & Anderson, C. W. (in press). Principles, first and foremost: A tool for understanding biological processes. *Journal of College Science Teaching*. Retrieved from http://edr1.educ.msu.edu/EnvironmentalLit/publicsite/html/report_2013.html

Rose, S. L. & Barton, A. C. (2012). Should Great Lakes City build a new power plant? How youth navigate socioscientific issues. *Journal of Research in Science Teaching, 49*(5), 541-567. doi:10.1002/tea.21017

Ryan, G. & Bernard, H. (2003) Techniques to Identify Themes in Qualitative Data. Retrieved from Analytic Technologies website: http://www.analytictech.com/mb870/readings/ryan-bernard_techniques_to_identify_themes_in.htm

Senge, P. (2014, Feb. 17). Peter Senge on teaching systems thinking in schools [Video file]. Retrieved from https://www.youtube.com/watch?v=rMkktp8uj9s

Shah, N. (2012, January 10). *Award-winning teen-age science in action* [Video file]. Retrieved from http://www.youtube.com/watch?v=sHJZRPWXQ9s

Sneider, C. (2015). How can environmental education support the NGSS? Unpublished manuscript.

Singleton, G. & Linton, C. (2006) *Courageous conversations about race.* Thousand Oaks, CA: Corwin.

Sjulin, J. (2011). Still resonating. In K. Nichols & L. Adams (Eds.), *Counting the rings: Stories, testimonials, and photographs of Multnomah Education Service District Outdoor School* (pp. 30–31). Portland, OR: Odin Ink.

Sobel, D. (2008). *Childhood and nature: Design principles for educators.* Portland, ME: Stenhouse Publishers.

Sosu, E. M., McWilliam, A., & Gray, D. S. (2008). The complexities of teachers' commitment to environmental education: A mixed methods approach. *Journal of Mixed Methods Research*, *2*(2), 169-189. doi:10.1177/1558689807313163

Springer, M. (Producer). 2013. *Listen as Albert Einstein reads 'The common language of science' (1941)* [Audio podcast]. Retrieved from http://www.openculture.com/2013/03/listen_as_albert_einstein_reads_the_common_language_of_science_1941.html

State Board of Education. (2014). *Administrative rule summary: Instructional time/OAR 581-022-0102, required instructional time/OAR 581-022-1620, credit options/OAR 581-022-1131.* Salem, OR: Department of Education.

Stevens, D., and Levi, A. (2005). *Introduction to rubrics: An assessment tool to save grading time, convey effective feedback and promote student learning.* Sterling, VA: Stylus Publishing, LLC.

Stevenson, R. B. (2007). Schooling and environmental education: Contradictions in purpose and practice. *Environmental Education Research, 13*(2), 139-153. doi:10.1080/13504620701295726

Svitak, A. (2010, April 2). *Adora Svitak: What adults can learn from kids* [Video file]. Retrieved from http://www.youtube.com/watch?v=V-bjOJzB7LY

Seymour, M. (2004). Educating for Humanity: Rethinking the Purposes of Education. Boulder, CO: Paradigm Publishers, pp. 89-96.

Tbilisi Communiqué: Educate Today for a Sustainable Future. TBILISI+35: Intergovernmental Conference on Environmental Education for Sustainable Development, September 2012, Tbilisi, Georgia. Outcome Document. IUCN, International Union of Conservation of Nature, 2016. Retrieved from https://cmsdata.iucn.org/downloads/tbilisi_story_komunike_small.pdf.

Trochim, W. M. (2006). Measurements construct validity. In *The research methods knowledge base* (2nd ed.). http://www.socialresearchmethods.net/kb/constval.php

Trochim, W. M. (2006). Measurements validity types. In *The research methods knowledge base* (2nd ed.). Retrieved from http://www.socialresearchmethods.net /kb/measval.php

Trochim, W. M. (2006). Nomological network. In *The research methods knowledge base* (2nd ed.). Retrieved from http://www.socialresearchmethods.net/ kb/nomonet.php

Trochim, W. M. (2006). Pattern matching for construct validity. In *The research methods knowledge base* (2nd ed.). Retrieved from http://www.socialresearchmethods.net /kb/reltypes.php

Trochim, W. M. (2006). Qualitative validity. In *The research methods knowledge base* (2nd ed.). Retrieved from http://www.socialresearchmethods.net/kb/qualval.php

Trochim, W. M. (2006). Threats to Construct Validity. In *The research methods knowledge base* (2nd ed.). Retrieved from http://www.socialresearchmethods.net /kb/consthre.php

Trochim, W. M. (2006). Types of reliability. In *The research methods knowledge base* (2nd ed.). Retrieved from http://www.socialresearchmethods.net/kb/consthre.php

UNESCO. (1977). *The Tbilisi declaration: Intergovernmental conference on environmental education.* Retrieved from http://resources.spaces3.com/a30712b7-da01-43c2-9ff0-b66e85b8c428.pdf

U.S. Environmental Protection Agency (EPA). (2015). EJSCREEN Technical Documentation. Retrieved from http://www.epa.gov/ejscreen

Washington, B. (1901) Up from slavery. In C. Kridel (Ed.), *Classic edition sources: Education* (4th ed., pp. 132–137), New York, NY: McGraw Hill.

Waters Foundation. (2008). Iceberg plan [Template]. Retrieved from Waters Foundation website http://www.watersfoundation.org

Waters Foundation. (2010). Systems sampler [Booklet]. Systems thinking in Schools – A Waters Foundation Project.

Waters Foundation. (2016). Big picture. Retrieved from http://www.watersfoundation.org/webed/habits/bigpicture.htm

Waters Foundation. (2016). *Systems Thinking: Connections to Curricular Standards.* Retrieved from http://watersfoundation.org/systems-thinking/standards/

Wirth, A. (1972). Charles A. Prosser and the Smith-Hughes Act. *Educational Forum, 36*(3), 365–371. Retrieved from ERIC database (EJ055679).

Zint, M. T., Dowd, P. F., & Covitt, B. A. (2011). Enhancing environmental educators' evaluation competencies: Insights from and examination of the effectiveness of the My Environmental Education Evaluation Resource Assistant (MEERA) website. *Environmental Education Research, 17*(4), 471-497. doi:10.1080/13504622.2011.565117

Zukoski, A. & Luluquisen, M. (2002). *Participatory Evaluation. What is it? Why do it? What are the challenges?* Retrieved from http://depts.washington.edu/ccph /pdf_files/Evaluation.pdf

Appendix A
Iceberg Model Template

Figure A. Iceberg Model Template

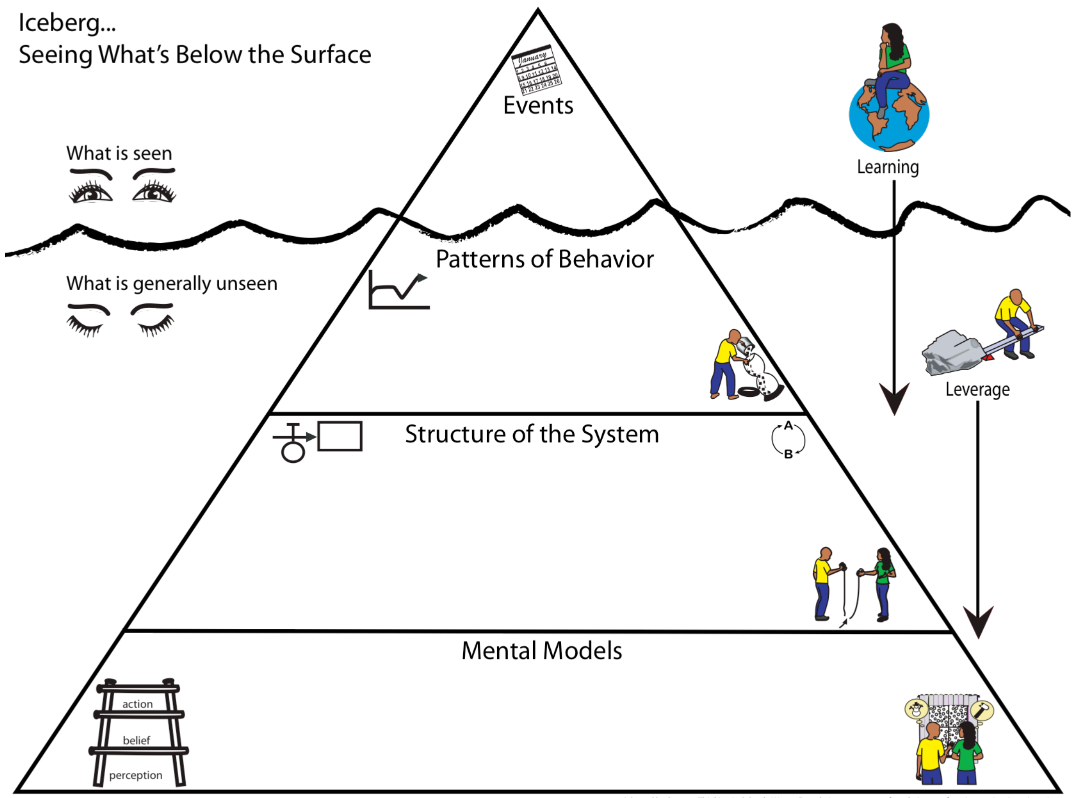
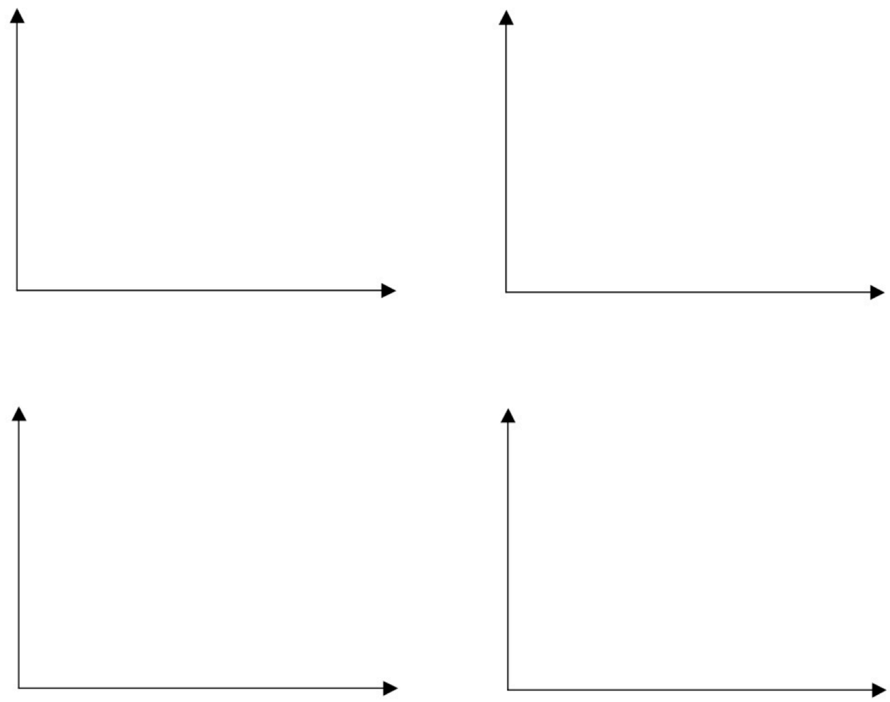


Figure A. Iceberg Model. Adapted by Systems Thinking in Schools, Waters Foundation. www.watersfoundation.org from Innovation Associates, Inc. Reprinted with permission.

Appendix B
Behavior-Over-Time Graph Template

Figure B. Behavior-Over-Time Graph Template

Behavior-Over-Time Graphs (BOTGs) with Questions



Some sample questions to ask when analyzing parts of a system that change over time:

1. What important elements have changed over time?
2. How has ___________ changed over time?
3. During what period of time have the changes occurred?
4. Where on the y-axis should the graph start and why?
5. How would you label the bottom/middle/top of the y-axis?
6. What evidence supports the graph being created?

Questions to consider once BOTGs have been created:

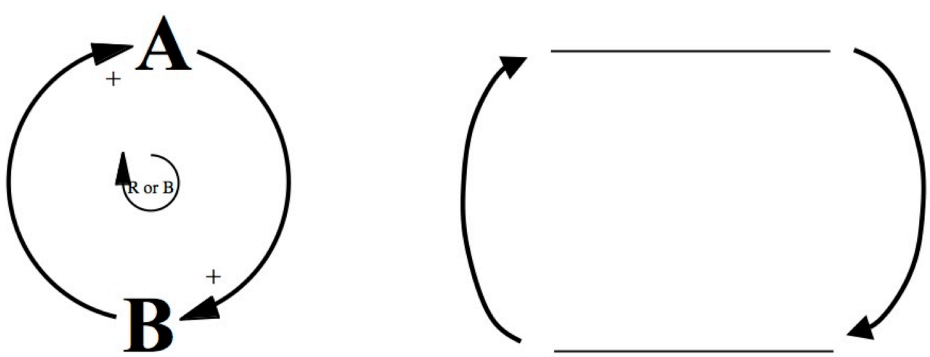
1. What caused any changes in direction or slope?
2. How are interpretations of a graphed element the same or different?
3. What changes may happen in the future based on what has been happening?
4. Do you see any connections (interdependencies or causal relationships) between/among graphs?

Figure B. Behavior-Over-Time Graphs. ©2015 Waters Foundation, Systems Thinking in Education, waterfoundation.org. Reprinted with permission.

Appendix C
Causal Loop Diagram Template

Figure C. Causal Loop Diagram Template

Causal Loop Diagram Template with Questions



Questions to ask when creating and analyzing causal loop diagrams
Creating a loop:

- Are there two elements in the system that affect one another?
- Are there just two elements affecting one another or are there more?
- Between any two parts of your loop, is the relationship…
- Does the first part increase or decrease the second part?
- What kind of loop is it – reinforcing (R) or balancing (B)? Label the center of the loop.
- Can I talk around the loop more than once?
- Are there any other loops or elements that might be connected to this loop?

Analyzing a loop:

- Given the label of 'R' or 'B', what kind of behavior would this loop produce over time?
- If multiple loops are connected, what loop(s) might be more or less influential over time?
- How do you know when to stop adding elements of the system to your diagram? (Consider meaningful boundaries.)

Figure C. Causal Loop Diagram Template.

Appendix D
Tragedy of the Commons Archetype Template

Figure D. Tragedy of the Commons Archetype Template

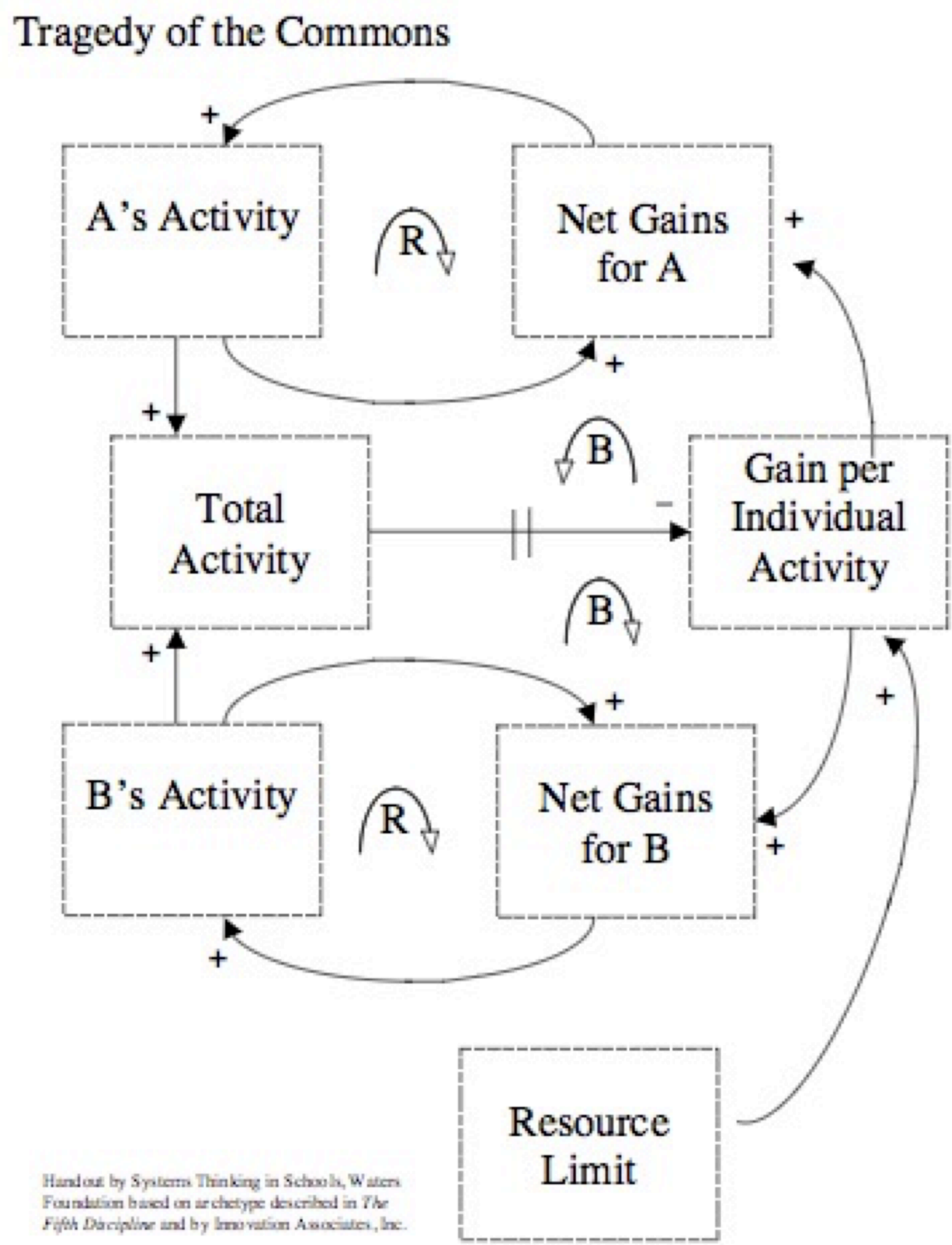
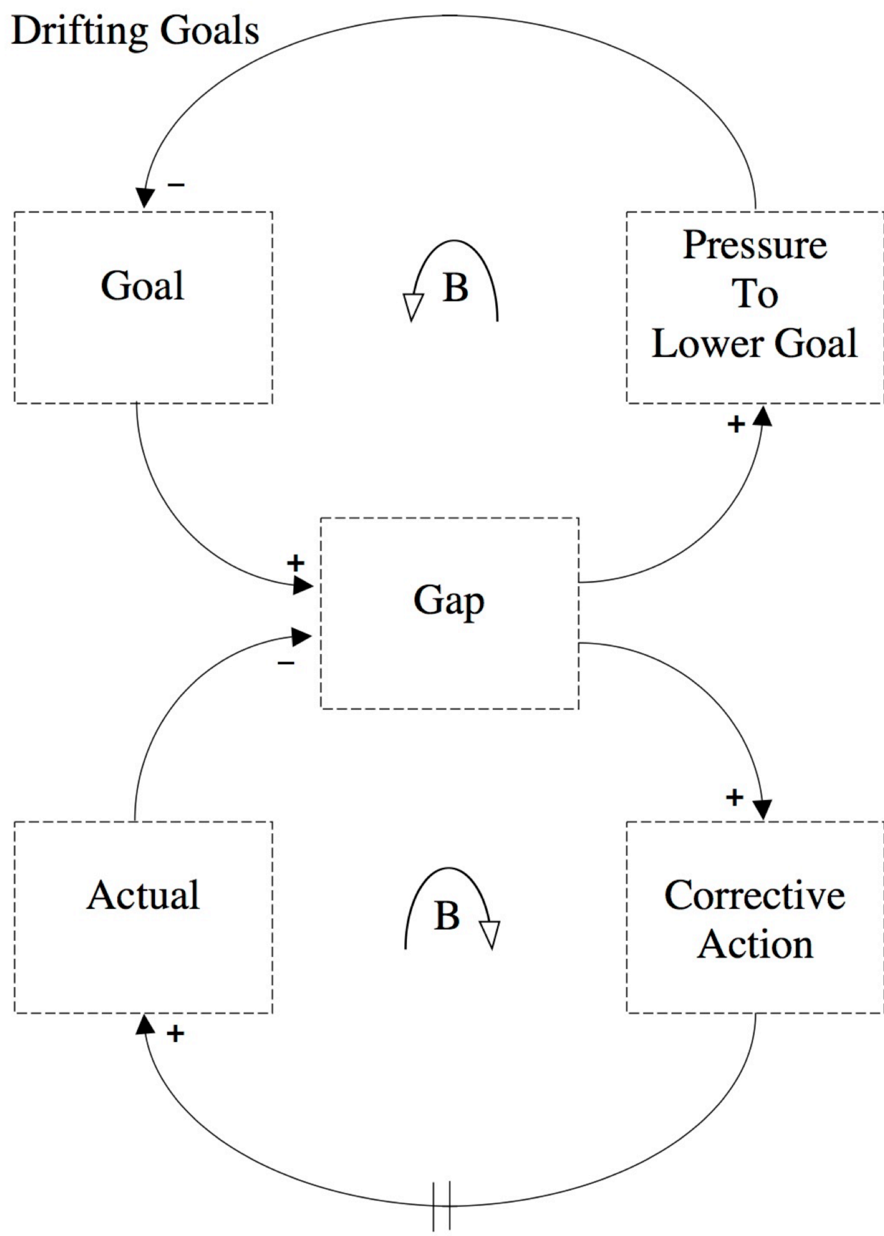


Figure D. Tragedy of the Commons Archetype Template. Handout by Systems Thinking in Schools, Waters Foundation based on archetype described in *The Fifth Discipline* and by Innovation Associates, Inc. Reprinted with permission.

Appendix E
Drifting Goals Archetype Template

Figure E. Drifting Goals Archetypes



Handout by Systems Thinking in Schools, Waters Foundation, www.watersfoundation.org, based on archetype described in *The Fifth Discipline*, Senge and by Innovation Associates, Inc.

Figure E. Drifting Goals Archetype Template. Handout by Systems Thinking in Schools, Waters Foundation, www.watersfoundation.org, based on archetype described in *The Fifth Discipline,* Senge and by Innovation Associates. Reprinted with permission.

Appendix F
Ladder of Inference

Figure F. Ladder of Inference

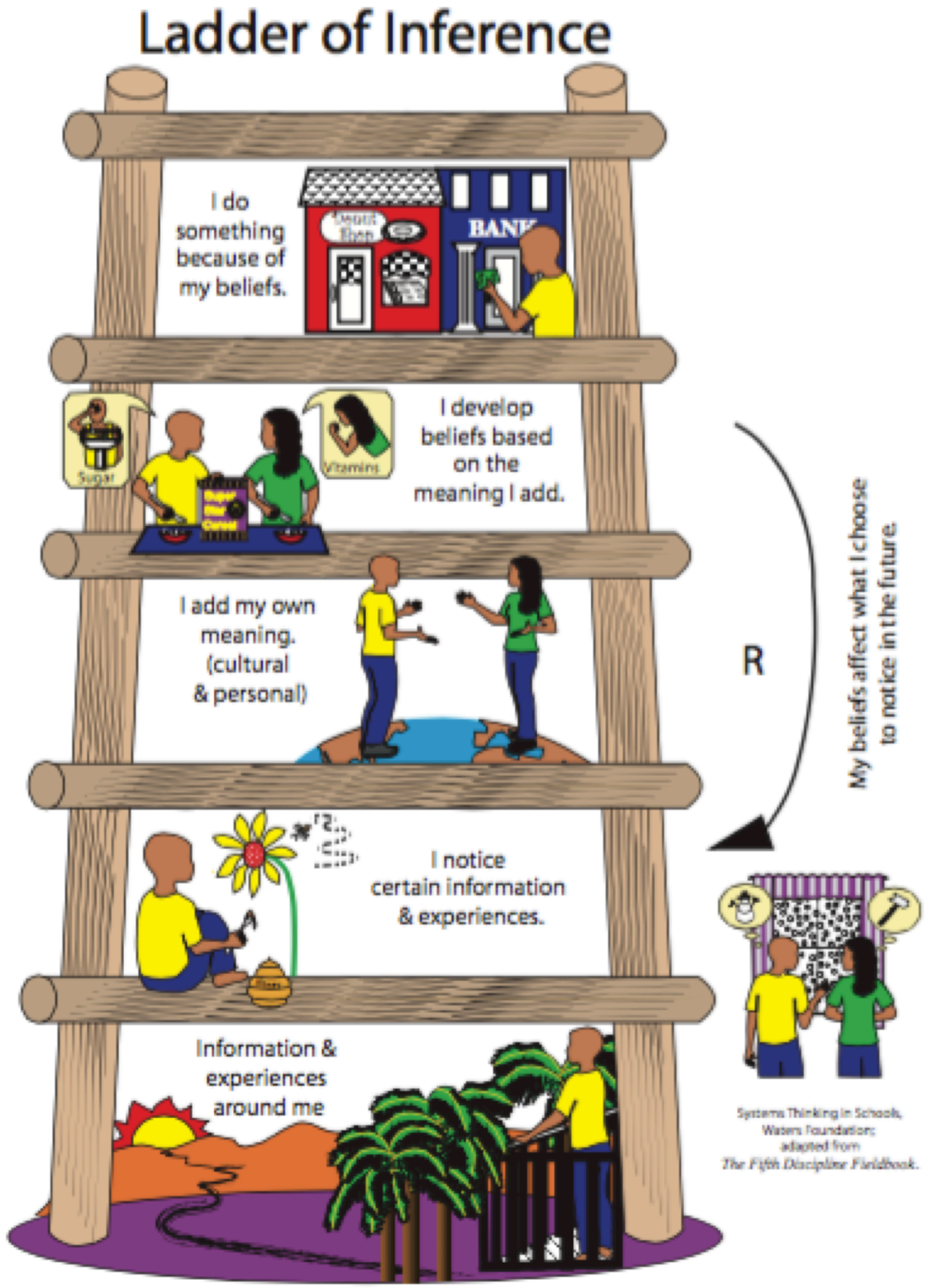


Figure F. Ladder of Inference. Systems Thinking in Schools, Waters Foundation, adapted from *The Fifth Discipline Fieldbook.* Reprinted with permission.

Appendix G
Correlation Between OELP Strands and NGSS

Table G1

Correlation of Oregon Environmental Literacy Plan Strands and Next Generation Science Standards

2013 Oregon Environmental Literacy Strand	Next Generation Science Standards
Systems Thinking Students apply systems thinking skills to study various types of systems and issues from a holistic perspective, striving to understand the relationships and interactions among the systems' parts. Students use the knowledge gained to consider the implications and consequences of choices on the economic, ecological, and social systems within which they live, in order to optimize outcomes for all three systems (p. 2).	**Ecosystems: Interactions, Energy, and Dynamics** MS-LS2-2. Construct an explanation that predicts patterns of interactions among organisms across multiple ecosystems. [Clarification Statement: Emphasis is on predicting consistent patterns of interactions in different ecosystems in terms of the relationships among and between organisms and abiotic components of ecosystems. Examples of types of interactions could include competitive, predatory, and mutually beneficial.]
Physical, Living, and Human Systems Students understand Earth systems' characteristics, including physical, living, and human systems (p. 3).	**Ecosystems: Interactions, Energy, and Dynamics** MS-LS2-1. Analyze and interpret data to provide evidence for the effects of resource availability on organisms and populations of organisms in an ecosystem. [Clarification Statement: Emphasis is on cause-and-effect relationships between resources and growth of individual organisms and the numbers of organisms in ecosystems during periods of abundant and scarce resources.]
Interconnectedness of People and the Environment Students understand the interdependence between the environment and humans, including the interconnectedness of human well-being and the environment (p. 5).	**Earth and Human Activity** MS-ESS3-3. Apply scientific principles to design a method for monitoring and minimizing a human impact on the environment. [Clarification Statement: Examples of the design process include examining human environmental impacts, assessing the kinds of solutions that are feasible, and designing and evaluating solutions that could reduce that impact. Examples of human impacts can include water usage (such as the withdrawal of water from streams and aquifers or the construction of dams and levees), land usage (such as urban development, agriculture, or the removal of wetlands), and pollution (such as of the air, water, or land).]

Note: The Oregon Environmental Literacy Strands are cited from the 2013 Oregon Environmental Literacy Plan, and related resources can be downloaded online from the Oregon Department of Education at http://www.ode.state.or.us/search/page/?id=2886. The Next Generation Science Standards were found online using a search of the specific section title at http://www.nextgenscience.org/search-standards. Reprinted with permission.

Appendix G (continued)
Correlation Between OELP Strands and NGSS

Table G2

Correlation of Oregon Environmental Literacy Plan Strands and Next Generation Science Standard

2013 Oregon Environmental Literacy Strand	Next Generation Science Standards
Personal and Civic Responsibility Students understand the rights, roles, responsibilities and actions associated with leadership and participation that lead toward healthy, sustainable environments and communities (p. 6).	**Ecosystems: Interactions, Energy, and Dynamics** MS-LS2-5. Evaluate competing design solutions for maintaining biodiversity and ecosystem services. [Clarification Statement: Examples of ecosystem services could include water purification, nutrient recycling, and prevention of soil erosion. Examples of design solution constraints could include scientific, economic, and social considerations.]
Investigate, Plan, and Create a Sustainable Future Students apply the civic action skills that are essential to healthy, sustainable environments and communities (p. 7).	**Earth and Human Activity** MS-ESS3-4. Construct an argument supported by evidence or how increases in human population and per-capita consumption of natural resources impact Earth's systems. [Clarification Statement: Examples of evidence include grade-appropriate databases on human populations and the rates of consumption of food and natural resources (such as fresh water, mineral, and energy). Examples of impacts can include changes to the appearance, composition, and structure of Earth's systems as well as the rates at which they change. The consequences of increases in human populations and consumption of natural resources are described by science, but science does not make the decisions for the actions society takes.]

Note: The Oregon Environmental Literacy Strands are cited from the 2013 Oregon Environmental Literacy Plan, and related resources can be downloaded online from the Oregon Department of Education at http://www.ode.state.or.us/search/page/?id=2886. The Next Generation Science Standards were found online using a search of the specific section title at http://www.nextgenscience.org/search-standards. Reprinted with permission.

Appendix H
Pugh Chart used to Evaluate Existing Environmental Literacy Assessment Instruments

Table H1

Section One: Comparison of Assessment Instruments for Construct Validity

Evaluation Criteria	OAKS (Datum)	PISA	MYP	ODE Work Sample	NELA	MEERA	AESL	EUGENE
Validity: Operational construct is framed before use.	Eligible content approved by expert panel formed by ODE, which is shared with public. (0)		Teaching, learning & assessment evaluated by experts visiting setting. (+1)	Students scored by educators who were present in the context of instruction. (+1)	Findings only generalizable to a few programs. (-1)	Provides instruction for how to establish validity in the design of the instrument by linking educators' reports. (+1)	Construct defined using a concept map for learning progression and scoring guide to indicate levels of proficiency. (+1)	Eligible content in test item bank limited to environmental science topics. (-1)
Validity: Interaction between testing and treatment includes outdoor experience.	Environmental literacy strands match seven of the existing science content standards. (0)			Field studies recommended by ODE for science inquiry. (+1)	Schools were selected to participate in research with the instrument if 2 classes had participated in environmental education activities for 2 years. (+1)	Examples include use of cognitive maps and phone interviews with students who are asked to retell understandings based on outdoor experiences. (0)	Construct defined using a concept map for learning progression and scoring guide to indicate levels of proficiency. (+1)	Eligible content in test item bank limited to environmental science topics. (-1)

Note: PISA = Programme for International Student Assessment; MYP = Middle Years Programme; ODE = Oregon Department of Education Work Sample; NELA = National Environmental Literacy Assessment; MEERA = Measuring the Evaluation Competency of Non-Formal Educators; AESL = Assessments for Environmental Science Literacy—Michigan State University, and EUGENE = Ecological Understanding as a Guideline for Evaluation of Non-formal Education. Each instrument was compared for the criteria against the datum of the OAKS and score with a +1 if the instrument's characteristics appeared stronger than the datum, and -1 if they appeared weaker.

Appendix H (continued)

Pugh Chart used to Evaluate Existing Environmental Literacy Assessment Instruments

Table H2

Section Two: Comparison of Assessment Instruments for Reliability

Evaluation Criteria	**OAKS (Datum)**	**PISA**	**MYP**	**ODE Work Sample**	**NELA**	**MEERA**	**AESL**	**EUGENE**
Reliability: Provides measures of proficiency in environmental literacy that are meaningful to educators from formal and non-formal settings.	Blueprint and released items available at ODE website for practice. (0) Formal and non-formal educators can access the Science Content & Assessment Panel. (0)	Demographic and survey questions allow for correlations and generalization. (0)	Training in scoring and unit planning, approval an oversight provided to educators whose organizations apply to be part of the fee-based. (+1)	Teachers trained to score through moderation and calibration meetings twice yearly. (0) Matches data collection and analysis skills learned by field researchers who become educators. (0)	Web-based training for those administering the survey. (0) High content validity allows comparison of programs. (+1)	Educators can spend as many hours as they wish learning how to generate evaluation instruments with the on-line tutorial. (0) Designed for use by K-12, university and non-formal educators. (+1)	Multiple scorers for same student's work. (0) Reliability checks are done to ensure the students' work has all the characteristics for a particular level on the rubric. (+1)	Multiple scorers for same student's work. (0) Question bank and on-line test generator created through partnership between educators in formal and non-formal settings. (+1)

Note: PISA = Programme for International Student Assessment; MYP = Middle Years Programme; ODE = Oregon Department of Education Work Sample; NELA = National Environmental Literacy Assessment; MEERA = Measuring the Evaluation Competency of Non-Formal Educators; AESL = Assessments for Environmental Science Literacy—Michigan State University, and EUGENE = Ecological Understanding as a Guideline for Evaluation of Non-formal Education. Each instrument was compared for the criteria against the datum of the OAKS and score with a +1 if the instrument's characteristics appeared stronger than the datum, and -1 if they appeared weaker.

Appendix H (continued)

Pugh Chart used to Evaluate Existing Environmental Literacy Assessment Instruments

Table H3

Section Three: Comparison of Assessment Instruments for Equity with Total Scores for Validity, Reliability, and Equity

Evaluation Criteria	OAKS (Datum)	PISA	MYP	ODE Work Sample	NELA	MEERA	AESL	EUGENE
Equity: Allows educator to provide specific supports to individual students.	On-line test. (0) Can stop & continue later. (0) Multiple attempts permitted. (0) Immediate tabulation of quantitative score given to student. (0)	Open responses. (+1) Assumes learning progressions. (+1) Developed for International Audience. (0) Qualitative scores measured by researchers and reported by demographics. (0)	Availability of eAssessment in 2015. (0) Collection of multiple work samples used to make summary judgment for each student. (+1) Students can be given freedom to research topics of individual interest. (+1) Proficiency level determined by educator. (+1)	Collection of multiple work samples used to make summary judgment for each student. (+1) Students can be given freedom to research topics of individual interest. (+1) Proficiency level determined by educator. (+1)	Measures knowledge, affect, cognitive skills, and behavior. (0) Qualitative scores measured by researchers and reported by demographics. (0)	Assessments are generated by specific educators to evaluate specific programs. (+1) Assessment instruments can be generated on a case-by-case basis. (+1) A variety of assessment instruments are recommended. (+1) Qualitative and quantitative data provided to various stakeholders for a range of purposes. (0)	Open ended response. (+1) Index of Discrimination used to correlate items with overall proficiency. (+1) Qualitative scores measured by researchers and reported by demographics. (0)	Designed for students to complete pre and post test on-line for calculating t-test results. (0) Qualitative scores measured by researchers. (0)
Total Scores	0	2	5	6	1	5	5	3

Appendix I
Scoring Guide Revealing Researcher Bias

Systems thinking: [I can] apply systems thinking skills to study various types of systems and issues from a holistic perspective, striving to understand the relationships and interactions among the systems' parts. I can use the knowledge gained to consider the implications, and consequences of choices on the economic, ecological and social systems within which they live, in order to optimize outcomes for all three systems (OELP, 2010, p. 2.)

Interdependent Relationships in Ecosystems: [I can] construct an explanation that predicts patterns of interactions among organisms across multiple ecosystems.[Clarification Statement: Emphasis is on predicting consistent patterns of interactions in different ecosystems in terms of the relationships among and between organisms and abiotic components of ecosystems. Examples of types of interactions could include competitive, predatory, and mutually beneficial (NGSS Lead States, 2013, MS-LS2-2).

Table I1

Section One: Scoring Guide for Systems Thinking Environmental Literacy Strand of Oregon Environmental Literacy Plan

Change over time graphs and Connection circle	Causal loops and Stock/Flow maps	Ladder of Inference	Iceberg Model and Archetypes
I can graph how parts of Earth's physical, living and human systems change over time, and describe patterns and trends. I can show the relationships between the elements that are changing over time by drawing a connection circle. (p. 3-4)	I can identify and draw feedback loops that show the interconnectedness of human wellbeing with the environment. I can draw a map showing the inflows and outflows that are causing increases or decrease in the parts of a system where things accumulate (stocks). I can explain how science principles affect the rate of inflow and outflow in a system (p. 5-6).	I can use a ladder of inference to explore existing mental models, which is how people form their perspective before they take action. I can describe how my outdoor experience, culture and available information influences my own, and others', perspectives around an issue (p. 7-8)	I can use the iceberg model to expose the patterns and trends, structures, and mental models lying below the initial events caused by a specific problem. I can identify leverage points that lead to long-term solutions. I can associate an archetype with a system, and draw how it applies to the system I am considering. I can use my understanding of the archetype to optimize how the parts of a systems work together to create a sustainable future (p. 9-13).

Note: Adapted with permission from the Waters Foundation, Systems Thinking in Education, watersfoundation.org

Appendix I (continued)
Scoring Guide Revealing Researcher Bias

Table I2

Section Two: Scoring Guide for Tools of Systems Thinking Environmental Literacy Strand of Oregon Environmental Literacy Plan with Next Generation Science Standards Connections

Change over time graphs and Connection circle	Causal loops and Stock/Flow maps	Ladder of Inference	Iceberg Model and Archetypes
Physical, Living and Human Systems: [I] understand Earth systems' characteristics, including physical, living and human systems (OELP, 2010, p. 3). **Matter and Energy in Organisms and Ecosystems:** Analyze and interpret data to provide evidence for the effects of resource availability on organisms and populations of organisms in an ecosystem. [Clarification Statement: Emphasis is on cause and effect relationships between resources and growth of individual organisms and the numbers of organisms in ecosystems during periods of abundant and scarce resources.] (NGSS Lead States, 2013, MS-LS2-1)	**Interconnectedness of People and the Environment:** Students understand the interdependence between the environment and humans, including the interconnectedness of human wellbeing and the environment (OELP, 2010, p. 5). **Earth and Human Activity:** Apply scientific principles to design a method for monitoring and minimizing a human impact on the environment [Clarification Statement: Examples of the design process include examining human environmental impacts, assessing the kinds of solutions that are feasible, and designing and evaluating solutions that could reduce that impact. Examples of human impacts can include water usage (such as the withdrawal of water from streams and aquifers or the construction of dams and levees), land usage (such as urban development, agriculture, or the removal of wetlands), and pollution (such as of the air, water, or land).] (NGSS Lead States, 2013, MS-ESS3-3.)	**Personal and Civic Responsibility:** Students understand the rights, roles, responsibilities and actions associated with leadership and participation that lead toward healthy, sustainable environments and communities (OELP, 2010, p. 6). **Ecosystems: Interactions, Energy and Dynamics** Evaluate competing design solutions for maintaining biodiversity and ecosystem services (NGSS Lead States, 2013). [Clarification Statement: Examples of ecosystem services could include water purification, nutrient recycling, and prevention of soil erosion. Examples of design solution constraints could include scientific, economic, and social considerations.] (NGSS Lead States, 2013, MS-LS2-5)	**Investigate, Plan and Create a Sustainable Future:** Students apply the civic action skills that are essential to healthy, sustainable environments and communities (OELP, 2010, p. 7). **Earth and Human Activity:** Construct an argument supported by evidence or how increases in human population and per-capita consumption of natural resources impact Earth's systems. [Clarification Statement: Examples of evidence include grade - appropriate databases on human populations and the rates of consumption of food and natural resources (such as fresh water, mineral, and energy). Examples of impacts can include changes to the appearance, composition, and structure of Earth's systems as well as the rates at which they change. The consequences of increases in human populations and consumption of natural resources are described by science, but science does not make the decisions for the actions society takes.] (NGSS Lead States, 2013, MS-ESS3-4).

Note: Reprinted with permission from the Oregon Department of Education and NGSS. Source: NGSS Lead States. 2013. *Next Generation Science Standards: For States, By State*s. Washington, DC: The National Academies Press.

Appendix J

A Scoring Guide for Systems Thinking – An Oregon Environmental Literacy Strand

Modeling and Analysis	Systems Habits	Problem Solving	Refining and Proposing Changes
Make a claim using evidence, and provide your reasoning orally and in writing. Construct an argument from analysis of data.	Identify long and short-term consequences.	Identify the problem of a situation.	Display proposed changes and outcomes via easily understood diagrams. Use visual graphic skills to clearly present how the changes affected the environment.
Predict how changes in one part of a system could affect the rest of the system. Identify variables and differing outcomes with changes to variables.	Identify the purpose of the system and why it is important.	Create solutions that could mitigate the problem, and predict how changes to the system could emulate those solutions. Make inferences from experience.	Analyze data.
Create solutions for systems that are not in balance or unsustainable.	Identify leverage points with greatest impact. Suggest how to use leverage to affect the system.	Use creative thinking. Use the design cycle.	Develop a sound model. Suggest corrective actions by finding leverage points and making slow gradual changes.
Explain if relationships are "correlation" (a mutual relationship between two things) or "causation" (one action causes another).	Develop models. Use an Iceberg model to show what people already know, and the bigger picture of the system.	Use dynamic modeling with stocks and flows, and change variables until the desired outcome is achieved.	Collaborate. Use the design cycle to explore multiple solutions for the same problem and create a +/- chart for each solution.
	Show how a system's structure generates its behavior.		Use strong presentation skills to share ideas in a way that people will understand you. Vocalize the proposed changes and answer questions pertaining to the ideas shared.
	Present the complex inner workings of a system in a simple and succinct way.		
	Make connections between the parts of a system and their outcomes, highlighting the interdependence of each part to make a whole.		
	Show how elements in systems change over time. Track changes over time.		

Note: Created using the Delphi Process with a forum of Oregon teachers, middle school students, and environmental educators in June 2015.

Appendix K
Letter of Consent for Educators

Dear Educator,

You are invited to take part in a study, *Evaluating an Assessment Instrument for the Oregon Environmental Literacy Plan (OELP*), that uses your systems thinking skills and experience in teaching outdoors in a formal or non-formal setting. You would be one of six educators working with three middle school students to generate a scoring guide. The scoring guide will be tested with 100 educators from across Oregon. The study compares how often educators from places like Outdoor School and teachers from different school districts assign the same evaluative scores to the same project.

What will I have to do?

The researcher, Susan Duncan, would like your permission to:

- Include your knowledge, and/or practical experience in a scoring guide to assess a middle school student's work sample for systems thinking.
- Give you permission to access the Google Doc site "Environmental Literacy Forum" using the email that you choose to provide.
- Include comments that you add to three Google Doc forms in the scoring guide and share them with others in the 9-member group.
- Publicly post the scoring guide to the Internet in a Google Docs form and provide the link to other educators so they can test it by scoring a sample of student's work.

In addition, you will be requested to:

- Share your ideas about the tools and skills used by systems thinkers
- Review examples of students' work using the links at http://goo.gl/sMLxM7.
- Access the Oregon Environmental Literacy Plan (OELP) at http://goo.gl/dAcGqk and read pages 16-24.
- Review a 1-page chart at http://goo.gl/D16ijT showing Next Generation Science Standards (NGSS) associated with the OELP.
- Optional: Learn more about systems thinking through self-study using these videos at http://goo.gl/CG7Ixg.

Are there any risks?

A key concern is for your ideas to be taken seriously by the group working on the scoring guide. By maintaining anonymity, not even the researcher will know which comments came from which member of the group. Any English grammar or spelling errors will be corrected and statements will be aggregated and paraphrased as shared meaning develops. You are also encouraged to keep your participation in the project private because I cannot guarantee that other participants will not recognize you or find out that you are participating.

Appendix K (continued)
Letter of Consent for Educators

What are the benefits?
This study is designed to provide information that will assist teachers, administrators, and partners in the education community in evaluating programs like Outdoor Science School that support environmental literacy. The results of this study might become a benchmark for other educators and students. This might potentiate more middle school students having a chance to excel in systems thinking and resulting in a positive impact on the community.

What are you doing to protect me?
All information in this study will be kept confidential to the extent permitted by law, and the names of all people in this study will be kept confidential. Any identifying information will be stored separately from your responses.

Your participation is voluntary. If you decide not to participate in this study, it will not affect a class grade, or your relationship with an instructor, school district, or the University. If you decide to participate in the study, you may withdraw at any time without penalty. By participating, you are not waiving any legal claims, right or remedies.

Whom Can I Call with Questions?
If you have questions about this study, please contact me at (503) 422-2853, duncans@pdx.edu. You may also contact the Principal Investigator, Professor Swapna Mukhopadhyay, by email at swapna@pdx.edu, or by phone at (503) 725-8495. If you have concerns or problems about your participation in this study or your rights as a research participant, please contact the PSU Office of Research Integrity, 1600 SW 4th Ave., Market Center Building, Ste. 620, Portland, OR 97207; phone (503) 725-2227 or 1 (877) 480-4400.

If you wish to participate, please sign the attached consent form, indicate your decision by checking the options, and return the consent form to the researcher by mail:

Environmental Literacy Forum
c/o Susan Duncan
PO Box 82912
Portland, OR 97282

Please keep the first 2 pages for your records.

Appendix K (continued)
Letter of Consent for Educators

Consent Form for Participation in the Study:
Evaluating an Assessment Instrument for Environmental Literacy

Educator

______ Yes, I agree to participate in the study. I allow the ideas and opinions that I submit in three Google docs form to be used in a scoring guide for systems thinking. I give permission for the researcher to make the scoring guide available to the public via a Google docs form in order to measure whether 50 educators from a formal education setting and 50 educators from a non-formal setting score students work the same way.

Email address: ____

______ No, I prefer not to participate in the research study.

____ ____________
Educator Signature Date

Appendix L
Letter of Consent for Students

Dear Student and Parent,

You are invited to participate in a study called, *Evaluating an Assessment Instrument for the Oregon Environmental Literacy Plan (OELP*), that proposes to help people learn to use systems thinking in finding possible solutions to environmental challenges. You would be one of three middle school students, along with six educators, to decide what successful students do when they use systems thinking. The ideas about systems thinking that the forum identifies are important will be added to a scoring guide. The scoring guide will be used later by 100 Oregon educators from across the State to evaluate systems thinking project(s) that you and other students share with us.

What will I have to do?
The researcher, Susan Duncan, would like your permission to:

- Include the ideas that you add to three Google Doc forms for the scoring guide and share them with the 9-member group called, a "forum".
- Give you permission to access the Google Doc site "Environmental Literacy Forum" using the email that you and your parent choose to provide.
- Post the scoring guide that you help create to a Google docs form that can be publicly accessed by Oregon educators via an email or electronic newsletter.
- Post a project that you have done with systems thinking as a Google doc so that it can be publicly accessed and scored by educators who are given the web address via an email or electronic newsletter.

In addition, you will be asked to:

- Share your ideas about how you use the tools and skills recommended by systems thinkers.
- Review examples of students' work using the links at http://goo.gl/sMLxM7.
- Open the Oregon Environmental Literacy Plan (OELP) at http://goo.gl/dAcGqk and read pages 16-24.
- Review a 1-page chart at http://goo.gl/D16ijT showing Next Generation Science Standards (NGSS) associated with the OELP.
- Optional: Learn more about systems thinking through self-study using the videos at http://goo.gl/CG7Ixg.

Are there any risks?
You might be concerned that your ideas will not be taken seriously because of your age. Since you will be given a code name to protect your privacy, not even the researcher will know which comments come from which member of the group. Any English grammar or spelling errors will be corrected. Your ideas will be combined with the ideas of others who agree with you. You and your parents might be worried about posting your systems thinking project to a Google doc site where anyone who has the link can access it. Your work will only be available for the duration of the study and then removed. Although educators are asked to agree not to download, copy or share your project, I, as the researcher, cannot guarantee that educators will not violate their agreement. To safeguard against copyright infringement, you and your parents are encouraged to

send a photo of the original work or an electronic copy in .pdf format. I will not publish a copy of your project as part of my report for the dissertation because my question is about whether educators can use the scoring guide, not how well students' projects score. You are also encouraged to keep your participation in this study private because I cannot guarantee that others who participate will not recognize you or find out that you are participating.

What are the benefits?
This study is designed to provide information that will assist the education community with evaluating outdoor learning and science centers that support environmental literacy. The results of this study might set a benchmark for future educators and students to reach. More middle school students might learn how to find potential solutions to challenges in their community and natural systems.

What are you doing to protect me?
All information in this study, including your name, will be kept confidential unless some one is unsafe or treated unfairly. In which case, I am ethically and legally required to notify the appropriate authorities. Any identifying information will be removed from your systems thinking project before it is used in this study. Educators are asked to refrain from downloading, copying, sharing or distributing any copies of your work, but I cannot guarantee that they will. I encourage you to sign and date the original copy of your work and have a witness sign it, too. That is one way to show that you are the copyright owner. The scores that educators assign to the systems thinking projects will be used to calculate their level of agreement. Individual scores on a particular project will not be shared. Please note that by signing this consent form, you agree to be supervised by your teacher or parent while using the Internet to participate in this project, and you will abide by the agreements in your family, or at your school regarding safe use of the Internet.

Your participation is completely voluntary. If you decide not to participate in this study, it will not affect your grade, or relationship with your teacher, and your standing at your school. If you decide to participate in the study, you may withdraw at any time without penalty. You and your parents are not waiving any of your legal rights by signing this consent form, which means, please tell your parent if you think you feel unsafe or are being treated unfairly so they can contact PSU.

Whom Can I Call with Questions?
If you have questions about this study, please contact me at (503) 422-2853, duncans@pdx.edu. You may also contact the Principal Investigator, Professor Swapna Mukhopadhyay, at (503) 725-8495 or swapna@pdx.edu. If you have concerns about your participation in this study or your rights as a participant, please contact the PSU Office of Research Integrity, 1600 SW 4th Ave., Market Center Building, Ste. 620, Portland, OR 97207; phone (503) 725-2227 or 1 (877) 480-4400.

If you wish to participate and/or share your systems thinking project, please sign the attached consent form, indicate your decision by checking the options, and return the consent form on the next page to your teacher. Your teacher will mail it to Ms. Duncan, who will contact you.

Please keep these first 2 pages for your own records.

Appendix L (continued)
Letter of Consent for Students

Consent and Assent Form for Participation in the Study:
Evaluating an Assessment Instrument for Environmental Literacy

Parent/Guardian(s)

______ Yes, I will allow my child/ward (print name) ____________________ to participate in the study, allowing the researchers to use my child's school project, or work sample(s). It can be posted in a Google Doc form on the Internet so 100 Oregon educators can score it.

I also agree for the researcher, Ms. Duncan, to contact my child via email with instructions for sharing their ideas with educators and students participating in a discussion about their skills of systems thinking. I allow the ideas and opinions that my child submits in three Google doc forms to be used in a scoring guide for systems thinking. I give permission for the researcher to make the scoring guide available to the public via a Google Docs form in order to assess whether 50 educators from a formal education setting (e.g. public schools) and 50 educators from a non-formal setting (e.g. outdoor education) score a student's work the same way.

Email address: ____

______ No, I prefer that my child/ward not participate in the research study.

____ _____________
Parent/Guardian Signature Date

Student

______ Yes, I agree to participate in the study, allowing the researcher, Ms. Duncan, to use my school project, or work samples, and post it in a Google Doc form on the Internet so 100 educators from around Oregon can score it.

I also agree for the ideas and opinions that I submit in three Google doc forms to be used in a scoring guide for systems thinking. I give permission to Ms. Duncan to make the scoring guide available to the public via a Google Docs form in order to measure whether 50 educators from a formal education setting and 50 educators from a non-formal setting score a student's systems thinking project in the same way.

______ No, I prefer not to participate in the research study.

____ _____________
Student Signature Date

Appendix M
Invitation to Educators for Electronic Newsletters

A hundred educators are needed to assess a newly created scoring guide for the Oregon Environmental Literacy Strand - Systems Thinking! A practicing teacher and a researcher from Portland State University's Graduate School of Education cordially invite you to participate by scoring a project about salmon completed by middle school students. Your participation will be maintained as anonymous and will take about 15 minutes. The researcher will use the scores to measure inter-rater reliability. To participate, please complete the Google Doc Form at https://goo.gl/3Egf5n, or click the yellow highlighted link at https://sites.google.com/a/pdx.edu/environmental-literacy-forum/

Appendix N
Data Analysis Procedures for Scoring Guide

Complete the following without the written statements, only the numbers to avoid bias.

1. Download file as Excel from Google Docs. Remove "date stamp."
2. Sequence by Statement Number.
3. Convert responses to associated ordinal numbers using "replace."
4. Add Column to label statistics.
5. Enter formula for median to 2 decimal places. Calculate for each statement.
6. Enter formula for standard deviation to 2 decimal places. Calculate for each statement.
7. Add data for statements from Round 2 that reached agreement using the data posted on the Worksheet Round 3 used by the forum members.
8. Enter each person's rating for each statement to SPSS to determine percent agreement for each response to each statement. Rule: If the statement has 80% agreement for 'Very important' or 'Quite important', include it in the scoring guide. If the statement has 80% agreement for 'Very unimportant' or 'Quite unimportant', do not include it in the scoring guide.
    a. Open a new data file in 'Variable View,' and label the variables and their attributes.
        i. Name → Statement#
        ii. Type → Numeric
        iii. Width → 10
        iv. Decimals →2
        v. Label → Statement 1
        vi. Values → None
        vii. Missing → None
        viii. Columns→ 8
        ix. Align → Center
        x. Measure → Ordinal
        xi. Role → Input
    b. Click the tab for 'Data View,' and enter the ratings provided by each member of the forum.
        i. Copy and paste by statement from Excel file that was downloaded from Google.
9. Analyze the Data using SPSS by going to Analyze: Descriptive Statistics: Frequencies.
    a. Select all statements with data entered. Statements not included for Round Three included: 9,12,13,14,20,27,44,52,54.
    b. Click on Statistics. Check boxes for mean, median, range, and standard deviation.
    c. Click on Format: Check box for "Compare variables."
    d. Double click on chart to open "Pivot Table"
    e. Select variable and statistic headers and drag to opposite axes.
    f. Select all data. Copy and paste to Excel spreadsheet.

10. In Round 2, the following sort was used: Sort by category (smallest to largest), 5- Very important (largest to smallest), and Standard deviation (smallest to largest).
    a. Delete rows that are less than 60% for 5- Very important, or a combination of 5 & 4 of <80%. Basically remove all the median 3's.
11. Recheck
12. For Round 3, compare results for Round 2 & Round 3 using the rules for agreement described below:
13. Sort again and allow for 80% of very and quite important added together.
14. Cross reference median and standard deviation with calculation made by Excel.
15. If interested, set up Excel to compute inter-rater reliability percentage. Compare.

Reasoning Behind Scoring Guide

I first made the scoring guide using the rule for 80% consensus rating for very important, but my bias for including a Connection Circle for category one had me reframe the rule. So, I developed these additional rules and compared the data over both rounds rather than just using the final data without looking for changes in voice. I could not use Kappa since I did not code each person separately. Next time, I would ask them to make up a code name so that I would use it to track the data for calculating Cohen's Kappa and Fleiss Kappa. However, I would not report the data in such a way that they could identify themselves.

Category 1

- Statement 10 – Name of Category 1. Had the highest median in both rounds 4.5 and 5 respectively. Although the standard deviation in round 2 was smaller for statement 11, the median was lower and it became clear there was more agreement that statement 11 was ranked below statement 10 because the SD was low at .89 for a median score of 3. The percent agreement for the statement as very important increased by 10% in the third round, which may have been biased in favor of including the statement because one less person responded than in Round 2.
- Statements 12 and 13 were not included in Round 3 because the median scores were 3.5 and 3, which were below quite or very important. Their percentages for neither important, or unimportant were 50% and 67%, which were not yet at the median level of consensus for the group of 80%. However, they were removed to begin to simplify the survey.
- Statement 5 was included because it reached the median level of consensus determined by the group of 80% in the Round 3.
- Statement 3 was added to the scoring guide because an additional member increased its level of importance in Round 3. It also had a median score in Round 2 of 4.5 and Round 3 of 5. The range of data continued to remain small as 2 and 3 respectively. As hoped, the ranges for Round 3 increased over all. Where statements had a smaller range in both rounds, it was clear they were important to the forum. Statement 3 was an example of this. The combined percentage of agreement for very important (60%) and quite important (20%) met the level of consensus for importance in terms of how the worksheet was worded:

> "In order to measure consensus, I will use the median of the level suggested by all the participants, which was 80%. Those statements that gain consensus as 'quite important' and 'very important' will be added to the scoring guide. Those statements that gain consensus as 'quite unimportant' and 'very unimportant' will not be included in the scoring guide. Those statements that did not reach consensus will not be included in the scoring guide, but will be reported as findings for future reference."

- Statements 7 and 1 also had a median of 5 in Round 3. Sixty percent of the forum scored them as very important. The entire forum agreed that Statement 1 was either very or quite important. One person moved favorable towards Statement 7 in Round 3.
- Statement 9 was removed after Round 2 because the median was less than 4.
- Statement 8 was removed after Round 3 because the median was less than 4.
- Statements 2 and 6 were rated as quite unimportant in Round 3 indicating a shift from Round 2. So, both of those were removed.
- Statement 4 was included because it held a total percentage of 80% for very and quite important both rounds.

Category 2

- Statements 28 and 30 are very closely tied for the name of the category. They have held high means both rounds: 4 in Round 2 and 5 in Round 3. The difference between the average and the mean is smallest for Statement 30 and it also has the smallest range and SD.
- Statement 27 was removed after Round 2 because the median was less than 4.
- Statement 29 was removed after Round 3 because the median was less than 4.
- Statement 22 reached consensus as very important in Round 2.
- Statements 18 and 24 reached consensus as very important in Round 3.
- Statements 14 and 20 were removed after Round 2 because the median was less than 4.
- Statement 26 was removed after Round 3 because the median was less than 4.
- The following statements were rated as very or quite important in both rounds so the 60% + 20% rule was used along with the median 5 to include them in the scoring guide: Statements 15, 17, 19, 25.
- Since there were statements in category 1 that did not meet the 60-20 rule, the following statements were removed: Statement 4.
- Statement 16, 21 and 23 were removed because they had a median of 4 but did not reach 80% using the 60% very important and 20% quite important rule.
- Since there were statements in category 1 that did not meet the median of 4 and the 60-20 rule, the following statements were removed: Statement 7 and 1.

Category 3

- Statement 43 reached consensus in Round 3 for the name of the category.

- Statement 44 was removed because it did not reach a median of 4 in Round 2.
- Statements 41 & 42 were removed because they did not reach a median of 4 in Round 3.
- Statements 34 & 35 reached consensus in Round 2.

When this is done I am going to go back and check Round 2 for the statements that met the 67% + 33% or 67% + 17% rule for very and quite important to see if they all made it into the scoring guide.

- Statement 37 was added for the 60-20 rule and a median of 5.
- Even though Statements 31,32, 33, 36, 39 & 40 all had a median of 4 or 5 and one had 60% agreement for very important, they did not meet the 60% + 20% to add to 80%.
- I did included Statement 38 that reached 80% agreement by 40 +40 that had median of 5 if they had high percentages in both rounds.
- I added Statement 4 back to Category 1 based on the 40+40 rule. I reviewed Category 2 for these as well and added State 21 back into the scoring guide.

Next time I would include an opportunity to comment on each one to see if comments increased or decreased and find some way to help the forum share one another's reasoning and respond to it.

Category 4

- Statement 53 became the name of the category using the 60 + 20 rule. It had the highest median both rounds, and
- Statements 54, 55, 56, 57 were removed because only one name is needed.
- Statement 46 was added because it reached consensus for very important in Round 3.
- Statements 49 and 51 were added using the 60+20 rule and both had a median of 5.
- Statement 52 was removed after Round 2 because it did not reach a median score of 4.
- Statement 48 was added because it had a median of 4 reached the consensus for quite important using the 20+60 rule.
- I'll go back now through the whole thing to see if I missed any other 20 + 60 for quite important and then add the statements to see how unwieldy it is.
- Statement 6 was added to Category 1 using the 20+60 rule for consensus for quite important.
- Statements 50 & 45 were removed because they did not reach a median of 4 in Round 3.
- Overall, the final scoring guide created using the descriptive statistics as measures of importance has yielded a fairly lengthy scoring guide. I have a poor idea of which statements the numbers represent at this point except I wish that number 1 had made it since I think that was the one that I combine a number of statements

into one to be sure a Connection Circle would appear before sending the statements out in Round 2. I think for fairness I would not do any combining in the future and let the statements sort themselves out by their wording.

- I will send it out and ask for comments. Maybe I will just do as originally planned and summarize comments from the scoring testing process of Stage 2 and ask them to share their final thoughts. How do I keep myself out and in this process?

Final Thoughts

- This guide has some heft to it now and represents the forum solely using the numbers. I say, test it like it is, and let it ask the questions of the people who get it. I also suggest sending it to students when I ask for their work.
- Maybe take the things that were very important and rank them as highly proficient and quite important as nearly proficient, but NO that was not the question they were given. They were ranking items they already felt were required to be proficient. Maybe in the comment process there is a way to find out what makes a student highly proficient.
- The scoring guide calls for use of inquiry skills and engineering design skills in context. Systems thinking gives equal value to the context as well as the approach to solving the problem. All are just tools and limited to the persons ability to gain insight or extend their senses from using them.
- Once I had listed all these rules I felt like I had been fair to everyone who participated in the forum. 80% felt these statements were important to have in the scoring guide.

Appendix O
Codes for Autoethnography

Adopt a Farmer
Assessment - Interpretive Talk
Assessment - Nonformal
Assessment - Parent
Assessment - Verbal
Assessment -Formal
Asset-based Community Development
Choice
Citizen Science
Commodification
Culture - Japan
Culture - Kalyapuya
Culture - Natural Resources
Adaptation
Advocate
Application of systems thinking
Archetype - Drift to Low Performance (Enhance w/Best Actual)
Archetype - Shifting the Burden
Assessment - Scoring Guide
Assigned
Attention
Belief
Challenges, Needs
Communication Skills
Competition
Confounding
Cooperation
Cultural Competence
Cultural Identity
Cultural Influence - EE
Cultural Influence - School
Culture
Curious
Decision-making Skills
Dialogue
Diversity
Diversity - Gender
Diversity - White
Education - Purpose
Elders
Engineering
Environmental Identity
Environmental literacy
Equity
Equity - Self
Experience
Explore
Extension Service
Family
Field Trips
For profit
Frame of Reference
Framework for Curriculum Development
FTE
Generational Knowledge
Harmony
Health
Immigration
Imminent Domain
in loci
Inclusion/Exclusion
Instructional time
Interpretive Naturalist
Job Pipeline
Language
Limits to Growth
Locus of textual authority
Marmot Dam
Mental Model - Autonomy
Mental Model - Community
Mental Model - Crisis
Mental Model – Decision-making
Mental Model - Employee
Mental Model - Gender Roles
Mental Model - Intercultural Sensitivity
Mental Model - Learning
Mental Model - Sustainability
Natural resources
NGSS
NGSS - Assessment Boundaries
NGSS - Crosscutting concepts
Nonprofit
ODS
OELP Learning Strands
Online teacher
Oppression
Outcomes
Oversimplify
parent in loci
Parents
Peer Mentors
Planning Time
Platinum Rule
Political Structure - Advocacy
Political Structure - Authority
Political Structure - Historical
Political Structure - State Department of Education
Political Structure - State Departments of Natural Resources
Poverty

Preserving
Problem Solving
Public servants
Qualitative Reasoning
Quantitative reasoning
Reciprocity - Gift
Self
Sense of Place
Sense of Role
Social Capital
Social Norms of Conservation Behavior
Social Self
Social Structures - ADA
Social Structures - Nonformal & Formal
Social Structures - Responsibility
Sovereignty
Standards
State Board of Education
States of Being
Statistics
STEM, STEAM
Stories
Straddling Two Worlds
Structural influence - Balance, Harmony
Structural Influence - Funding
Structural Influence - Schooling
Structural Influence - Standards
Student Voice
Students - Middle
Systems Thinking - Archetypes
Systems Thinking - Behavior over Time Graphs
Systems Thinking - Causal Loops
Systems Thinking - Connection Circle
Systems Thinking - Flow
Systems Thinking - Habits
Systems Thinking – Hierarchy
Systems Thinking - Iceberg model
Systems Thinking - Ladder of Inference
Systems Thinking - Leverage
Systems Thinking - Purpose/Function
Systems Thinking - Time Delays
Systems Thinking - Unintended Consequences
Teaching
Theory
Underserved
Value
Violence
Vision
Vocation