

6-22-2015

Measuring Teacher Dispositions Toward Teaching Sustainable Systems: A Mixed Methods Approach for Instrument Development and Psychometric Testing

Jennifer Morales

Florida International University, jmora004@fiu.edu

DOI: 10.25148/etd.FIDC000112

Follow this and additional works at: <https://digitalcommons.fiu.edu/etd>

 Part of the [Curriculum and Instruction Commons](#)

Recommended Citation

Morales, Jennifer, "Measuring Teacher Dispositions Toward Teaching Sustainable Systems: A Mixed Methods Approach for Instrument Development and Psychometric Testing" (2015). *FIU Electronic Theses and Dissertations*. 2188.
<https://digitalcommons.fiu.edu/etd/2188>

This work is brought to you for free and open access by the University Graduate School at FIU Digital Commons. It has been accepted for inclusion in FIU Electronic Theses and Dissertations by an authorized administrator of FIU Digital Commons. For more information, please contact dcc@fiu.edu.

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

MEASURING TEACHER DISPOSITIONS TOWARD TEACHING SUSTAINABLE
SYSTEMS: A MIXED METHODS APPROACH FOR INSTRUMENT
DEVELOPMENT AND PSYCHOMETRIC TESTING

A dissertation submitted in partial fulfillment of the

requirements for the degree of

DOCTOR OF PHILOSOPHY

in

CURRICULUM AND INSTRUCTION

by

Jennifer Morales

2015

To: Dean Delia C. Garcia
College of Education

This dissertation, written by Jennifer Morales, and entitled Measuring Teacher Dispositions Toward Teaching Sustainable Systems: A Mixed Methods Approach for Instrument Development and Psychometric Testing, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

Isadore Newman

Joyce Fine

Erskine Dottin

George O'Brien, Major Professor

Date of Defense: June 22, 2015

The dissertation of Jennifer Morales is approved.

Dean Delia C. Garcia
College of Education

Dean Lakshmi N. Reddi
University Graduate School

© Copyright 2015 by Jennifer Morales

All rights reserved.

DEDICATION

I dedicate this dissertation to my parents Dulce R. Almanzar and Juan A. Morales and to my son Trystan X. Mills. Without their patience, understanding, support, and most of all love, the completion of this work would not have been possible.

ACKNOWLEDGMENTS

Without the guidance and understanding of many individuals, completing this dissertation would have not been possible. With the overwhelming level of support I cannot list the names of everyone who has been influential in completing my doctoral studies; however, I do want to express my gratitude to you all.

I am appreciative of the help provided by Dr. Joyce Fine, Dr. Erskine Dottin, Dr. Isadore Newman, and Dr. George O'Brien, members of my committee who volunteered their time, expertise, and support. Dr. O'Brien, who was instrumental in the process, gave so much of himself in my preparation to enter the world of academia. I thank you for always being there to listen and guide me as I constructed my knowledge through the many learning experiences you extended. The partnership that we developed over the years has truly been the structural foundation of my success.

I would like to thank my adopted parents, Dr. Carole Newman and Dr. Isadore Newman. Without your love, support, and help this research would not have been completed. The encouragement of Dr. Carole Newman brought me to apply into the program while Dr. Isadore Newman provided me with the statistical background and knowledge to conduct this and many other studies. Thank you both for adopting me as your kid, the lessons you have taught me, and your support.

I would like to extend a special thanks to Dr. David Newman. Thank you for your support, guidance, and most of all patience, your friendship is very valuable to me.

Lastly, I would like to thank my professors who always allowed for me to push my assignments to the limit. Thank you, Dr. Linda Bliss, Dr. Mido Chang, and Dr. Robert Farrell, the above and beyond nature of your instruction truly have made Florida

International University worlds ahead. It was a pleasure to work together with all of you and I look forward to our continuing friendship and future collaboration.

ABSTRACT OF THE DISSERTATION
MEASURING TEACHER DISPOSITIONS TOWARD TEACHING SUSTAINABLE
SYSTEMS: A MIXED METHODS APPROACH FOR INSTRUMENT
DEVELOPMENT AND PYSCHOMETRIC TESTING

by

Jennifer Morales

Florida International University, 2015

Miami, Florida

Professor George O'Brien, Major Professor

Sustainability concerns have become prevalent in environmental, societal, and economic systems. To address education towards sustainability the need for explicit instruction in sustainable systems is apparent; however, it is underrepresented in American schools. Despite the emergence of sustainability topics in the literature, few have addressed teacher dispositions about providing this needed instruction and none have reported quantitative measures with acceptable estimates of reliability and validity. Dispositions are defined as the tendency to act in a particular manner that aligns with an individual's belief which can develop and change over time, and are influenced by the experiences and circumstances faced by the individual. To provide the necessary instruction not only are instructors responsible for the curricular content and pedagogical content, but most importantly, they must possess positive dispositions towards providing this instruction.

The purpose of this study was to construct and determine estimates of the validity and reliability of the Dispositions Toward Teaching Sustainable Systems Instrument. Using a sequential mixed methods design and Luyt's Framework for instrument development the researcher, using qualitative methods such as interviews, identified themes that were

supportive of the theoretical construct. The four themes that emerged were: administrative support, outdoor resources, collaboration, and professional development, all receiving above 80% interrater agreement ($n=3$ judges). The quantitative aspect identified the key indicators (items) and their estimates of reliability and validity, and their goodness of fit to the four theoretical constructs.

Content validity estimates of the items generated from the interviews were explored using expert judges ($n=2$) in the area of sustainability education, also achieving above 80% agreement. Participants enrolled in three succeeding semesters of the Teaching Elementary Science Methods course ($n=233$) were then used to further estimate the reliability ($\alpha > .08$), construct validity (normed chi-Square = 2.04; GFI=.86 RMSEA =.076; NFI=.96 CFI =.98 PGFI=.634 binomial index of model fit $p < .001$), measurement invariance (CFI change $p < .011$), discriminant validity ($R = .271$ $R^2 = .073$ (4,227)=4.5 $p < .002$), and measurement sensitivity ($p < .05$), yielding support for the soundness of the instrument's psychometric properties.

TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION	1
Problem	1
Purpose.....	6
General Research Questions	6
Significance of the Study	8
Delimitations.....	9
Definitions and Operational Terms.....	9
Assumptions Underlying the Study	11
Summary	11
II. REVIEW OF THR LITERATURE.....	12
General Background Information	12
Summary	19
III. METHODS	20
Research Design.....	20
Mixed Methods for Instrument Development.....	21
Qualitative Methods.....	25
Trustworthiness.....	28
Quantitative Methods.....	29
Participants.....	30
General Research Hypothesis	31
Sampling Procedures	32
Instrument	33
Data Collection	35
Statistical Treatment	37
Summary	42
IV. RESULTS	44
Participants.....	44
Phase Two Participants	47
Phase Three participants	48
Qualitative Analysis.....	49
Content Validity Estimates	58
Reliability Estimates	59
Construct Validity Estimates	60
Binomial Index of Model Fit	66
Invariance Testing.....	68
Discriminant Validity Estimates.....	72
Measurement Sensitivity.....	73
Summary	75

V. SUMMARY, DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS	76
Summary of the Study	76
Methodology	77
Statistical Analyses	78
Research Questions, Conclusions, and Discussion	79
Implications	87
Limitations	88
Recommendations for Further Research	89
Summary	90
LIST OF REFERENCES	91
VITA	100

LIST OF TABLES

TABLE	PAGE
1. Demographic Information of Participants	49
2. Global Themes and Code Segments that Emerged from the Semi-Structured Interviews with Inter-rater Agreement	58
3. Tos agreement.....	58
4. Internal Consistency for Dispositions Toward Teaching Sustainable Systems total and Subscales at final time point.....	60
5. Model Fit Indices for Theoretical and Adjusted Models Posttest.....	66
6. Direct Effects and Standardized Regression Weights for Theoretical Model	67
7. Direct Effects and Standardized Regression Weights for Modified Model	68
8. Final Model Fit Indices for Time one, Two, and Three.....	69
9. Contribution of Model χ^2 and GFI.....	71
10. Model Invariance Results reporting χ^2 , RMSEA, NFI, CFI and CFI Change	71
11. Means and Standard Deviations of Experience as a Function of Subscales.....	72
12. Correlation of Predictor Variables with Discriminant Functions and Standardized Discriminant Function Coefficients	73
13. Simple Discriminate Function Using Regression Testing Experience against No Experience	73
14. Repeated Measure Analysis of Variance Reporting the Discriminant Capability of the DTTSSI Instrument.....	74

CHAPTER I INTRODUCTION

At the heart of educational reform in the 21st century is the subject area of sustainability education (Morris & Martin, 2011; Saylan & Blumestien, 2011; Stibbe, 2009). As our unsustainable practices elude suspicion due to the commonality of their nature, increasing concern for the state of our environment, economy, and societal health are prevalent (Barr, 2003). These paradoxes between the health of these systems and the practices of everyday choices have developed complex situations which require devoted attention (Colucci-Gray, Camino, Barbiero, & Gray, 2006). The United Nations Educational Scientific and Cultural Organization (UNESCO), the North American Association for Environmental Education (NAAEE) and the Association of Advancement of Sustainability in Higher Education (AASHE), have all focused their efforts in improving the presentation of sustainability education, in all aspects of the field across domains. Their efforts include providing support, professional development, and models of improvement to educational practitioners who desire to include the teaching of sustainable systems. However, even with many support initiatives that are offered by several organizations, the act of teaching about sustainable systems has not been implemented as a common practice amongst educational practitioners in the United States (Feinstine, 2009; Heselink, & Goldstein, 2000; Nolet, 2009 ; UNESCO, 2004).

Problem

According to UNESCO (2015) “citizens of the world need to learn their way to sustainability,” (Education for Sustainable Development, para.1). This statement places an emphasis on educators as a frontline to educating towards teaching for sustainable

systems. Developing aware citizens is only a starting point; these citizens would also need the systemic thinking necessary to deal with the complexity of sustainability problems (Colucci-Gray, Camino, Barbiero, & Gray, 2006; Morris & Martin, 2011; Saylan & Blumestien, 2011; Stibbe, 2009). This calls for a change in the educational culture of schools, to one that places sustainability education at its core. Sterling (2001) refers to this as a “transformative paradigm which values, sustains and realizes human potential in relation to the need to attain and sustain social economic and ecological well-being, recognizing that they must be part of the same dynamic” (p.22). To bring about societal change teachers must possess and positively influence the sustainability literacy among the communities they serve. Sustainability literacy is defined as “the skills, attitudes, competencies, dispositions, and values that are necessary for surviving and thriving in the declining conditions of the world in ways which slowdown that decline as far as possible” (Stibbe, 2009, p.10). As suggested by this definition, to effectively implement the practice of teaching towards sustainable systems these teachers must also possess the positive disposition towards teaching the matter. As Cantor (1990) has simply stated having (the necessary skills) does not equate to doing. Meaning that although colleges of education and arts and sciences prepare teachers with the content and pedagogical content knowledge necessary, it is not certain that they will implement learned skills in the field as has been demonstrated in the case of science education reform (Atkin & Black, 2007; Wilkerson, 2006; Van der Akker, 2003). When presented with new information or curriculum, “... teachers will need to restructure their knowledge and beliefs and, on the basis of teaching experience integrate the new information into practical knowledge” (van Driel, Beijaard & Verloop, 2002, p.140).

Wilkerson (2006) made the point that the willingness of teachers to implement innovative practices is highly dependent on their own dispositions and can be seen as more important than measuring the knowledge and skills alone. In a discussion on dispositions, Wilkerson states

We can think of dispositions as contingent on knowledge and skills; that is, teachers who lack the skills to carry out particular actions will be unable to do so, regardless of their desires. Having the knowledge and skills to teach particular content in particular ways is necessary but not sufficient to ensure that a teacher will employ them in the classroom.

The teacher must also have the disposition to do so.

Therefore, proponents argue, to not include dispositions in the preparation of teachers "is unconscionable and dangerous, since we need to ensure that teachers are likely to apply the skills they have learned in our colleges

(Wilkerson, 2006, p. 3).

It is therefore important to be able to measure these dispositions towards teaching sustainable systems, for a number of reasons. First by measuring dispositions we can predict the likelihood teachers will implement innovative practice in the field. Second a measure of teachers' dispositions serves as a beneficial tool for participants to monitor and reflect on their own dispositions during their program, and finally programs can monitor their effectiveness in supporting positive dispositions towards providing

innovative instruction. A literature search revealed that currently there are limited to no instruments available with good psychometric properties that measure dispositions toward teaching sustainable systems.

Recognizing dispositions as an essential component of effective teaching, the National Council for Accreditation of Teacher Education (NCATE) has included dispositions in the standards for the Accreditation of Teacher Preparation Institutions. Throughout the document readers find professional dispositions are a critical component of practicing skills. NCATE (2002) has defined professional disposition as “professional attitudes, values, and beliefs demonstrated through both verbal and non-verbal behaviors as educators interact with students, families, colleagues, and communities.” Furthermore, NCATE, “expects institutions to assess professional dispositions based on observable behaviors in educational settings,” (p.90) but only dispositions pertaining to fairness and the belief that all students can learn are the dispositions identified as anticipated to assess. Based on the NCATE (2002) mission and conceptual framework, professional education institutions “can identify, define, and operationalize additional professional dispositions.”(p.90) This supports the justified focus of measuring dispositions toward teaching sustainable systems. Even so measuring current dispositions alone does not provide the information needed in attending to teacher dispositions, these efforts should also include the ability to measure changes in dispositions among those currently in teacher preparation programs (Taylor & Wasicsko, 2000).

Similarly the nation’s new accreditor, the Council for the Accreditation of Educator Preparation (CAEP) has placed an emphasis on dispositions (CAEP, 2013). CAEP is concerned with advancing excellence in teacher preparation through evidence

based accreditation and merges the ideas of NCATE and the Teacher Education Accreditation Council (TEAC) to form its basis. Using a peer review based evaluation, CAEP serves two functions: assuring quality and promoting improvement. The focus of dispositions is still apparent in the CEAP framework. For example, in the discussion concerning candidate quality, recruitment and selectivity, CAEP has stated that

Educator preparation providers establish and monitor attributes and dispositions beyond academic ability that candidates must demonstrate at admissions and during the program. The provider selects criteria, describes the measures used and evidence of the reliability and validity of those measures, and reports data that show how the academic and non-academic factors predict candidate performance in the program and effective teaching. (CAEP, 2013)

This statement suggests it is the institutions' responsibility to determine criteria, measure, and report academic and non-academic factors that may predict candidate performance in their current program as well as in their teaching practice. Therefore if effective teaching of sustainable systems is the goal of the 21st century, higher education institutions should focus on identifying indicators such as dispositions towards providing instruction as predictors of effective implementation.

Purpose

The purpose of this study was to develop and investigate estimates of reliability and validity of the Dispositions Toward Teaching Sustainable Systems Instrument (DTTSSI). Both qualitative and quantitative methods were utilized in the construction and testing of this instrument. More specifically, thematic analysis of participant interviews and related literature review were used in the development of the theoretical framework and creation of items. Estimates of reliability using Cronbach's alpha was provided (Cronbach, 1951). Content validity estimates using a Table of Specifications and expert judges was presented (Newman, Lim, & Pineda, 2013). Construct validity estimates were explored using confirmatory factor analysis as well as results from the Binomial Index of Goodness of Fit (Brown, 2012; Clark & Watson, 1995; Frass & Newman, 1994; Fox, 1983; Jöreskog, 1965; Schumaker & Lomax, 2004; Thompson, 2004) In addition to assessing construct validity, estimates of discriminant validity was also determined by comparing initial disposition scores between those with a history of sustainability training and those without training. Lastly, a sensitivity analysis was conducted to assess if the DTTSSI can detect changes in dispositions in pre-service teachers who participate in a one semester three credit class.

General Research Questions

Research questions below are followed by a parenthetical comment indicating the phase of the study in which the question was answered.

1. Will there be themes that emerge from the interviews of participants in a sustainability workshop that could lead to the generation of items for the DTTSSI? (Qualitative)

2. Do the items (that are generated from the interviews) that comprise the Dispositions Toward Teaching Sustainable Systems Instrument support the content validity as measured by a table of specifications? (Qualitative and Qualitative)
3. Do the items (that are generated from the interviews) that comprise the Disposition Toward Teaching Sustainable Systems Instrument demonstrate adequate estimates of reliability as measured by Cronbach's alpha? (Quantitative)
4. Do the items (that are generated from the interviews) generated from the interviews that comprise the Disposition Toward Teaching Sustainable Systems Instrument demonstrate adequate estimates of construct validity as measured by a confirmatory factor analysis? (Quantitative)
5. Does the Disposition Toward Teaching Sustainable Systems Instrument demonstrate acceptable estimates of construct validity as measured by the Binomial Index of Model of Fit? (Quantitative)
6. Does the Disposition Toward Teaching Sustainable Systems Instrument support stability of the factor structure across time as measured by model validation using invariance testing? (Quantitative)
7. Does the DTTSSI demonstrate adequate discriminant validity in measuring disposition differences between pre-service who have had previous sustainability training from those who have not as measured by a simple discriminant function analysis? (Quantitative)
8. Does the DTTSSI demonstrate adequate measurement sensitivity (known group validity) in measuring disposition change in pre-service teachers across time as

measured by pre-test and posttest scores using a repeated measures analysis of variance? (Quantitative)

Significance of the Study

The significance of this research study is to extend the work concerning dispositions into the area of sustainability education. Sustainability concerns are widespread, and have been identified in the economic, societal, and environmental systems of today. Advancement towards sustainability in these areas requires that colleges of education and arts and sciences educate practitioners in providing instruction of sustainable systems with people, profit, and the planet in mind. Current initiatives are slow to manifest in the educational setting and are often under represented due to lack of evaluation efforts and teacher buy in. Teacher dispositions towards practice and inclusion of sustainable systems instruction are critical in determining if implementation is likely to occur in the classroom setting. Having the availability of instruments with acceptable estimates of reliability and validity can be used to measure current efforts in program development. By measuring a program's ability in influencing dispositions for increased inclusion in the school setting identification concerning areas of improvement can be addressed and program effectiveness can be determined. Through the use of psychometrically sound instruments, evaluation of programs can provide insight in program transferability into the educational setting furthering the advancement towards sustainable systems over time.

Delimitations

Three population samples were utilized in this study. The first phase of participants was delimited to educators who self-selected participation in an education for sustainability workshop and agreed to participate in recorded interviews and a self-reporting demographic questionnaire. The sample from the second phase of this research was delimited to those participants who were enrolled in the Teaching Elementary Science methods course at a university in south Florida. This sample was delimited to participants who completed all data collection points and had regular attendance in the course. The third sample included expert judge participants who completed the table of specifications (ToS) and are established in field of either science education, sustainability, or psychometrics.

Definitions and Operational Terms

1. Construct validity is defined as an assessment of a measure's ability in representing the construct being investigated as measured by a confirmatory factor analysis, table of specifications, and binomial index of model fit (Fraas & Newman, 1994; Newman et. al 2013; Polit & Beck, 2012).
2. Content validity is defined as the adequacy with which the specified domain of content is sampled (Nunnally, Bernstein, & Berge, 1967; Polit & Beck, 2006). For the purpose of this study, content validity was measured through expert judges' completion of a table of specification (ToS) that assessed not only the specific domains addresses in the DTTSSI but also asked if there are any additional facets that they felt should be included.

3. Discriminant analysis has been defined as an assessment of a measurement's ability to predict group membership based on some weighted linear set of variables (Newman, Newman, Brown, & McNeely 2006).
4. Dispositions are defined as the tendency to act in a particular manner that aligns with an individual's belief that can develop and change over time, and has been influenced by previous experiences and the particular circumstances faced by the individual (Brown & Cooney, 1982; Freeman, 2007; Katz & Rath, 1985; Richardson, 2003; Rokeach, 1968; Tabachnik & Zeichner, 1984; Villegas, 2007; Diez, 2006).
5. Measurement sensitivity analysis can be defined in many ways, and in this study it has been defined as a procedure to determine the ability of the instrument to discriminate among individuals with varying levels of the attribute of interest as measured by pre-test, posttest differences (Ferketich, 1991).
6. Sustainability education has been defined as the act of presenting knowledge to the current generation to meet their needs without jeopardizing the needs of future generations (Nolet, 2009).
7. Sustainability literacy has been defined as "the skills, attitudes, competencies, dispositions, and values that are necessary for surviving and thriving in the declining conditions of the world in ways which slowdown that decline as far as possible." (Stibbe, 2009 p.10).
8. Table of specifications (ToS) has been defined as a set of procedures that attempt to align a set of items, with a set of concepts to estimate the content validity (Newman et al., 2013).

9. Sustainability experience was defined as the self-reported familiarity of the participant with the topic or related themes of sustainability.

Assumptions Underlying the Study

There are five major assumptions that underlie this study. The first assumption is that the comments of participating practicing teachers and professors are reflective of others in practice. The second assumption is the data collection from the self-reported Dispositions Toward Teaching Sustainable Systems Instrument (DTTSSI) was assumed to be filled out accurately and without bias. The third assumption highlights the point that the data from the DTTSSI were collected across multiple semesters and assumes that there are no meaningful differences between the students from one semester to another. It was also assumed that this sample of students is representative of other pre-service elementary science teachers. Lastly, the underlying assumption that dispositions are measurable was made in this study.

Summary

This chapter presented background information concerning sustainability education and dispositions. Next the problem of enhancing sustainability education was presented followed by the purpose of the study in developing and reporting reliability and validity estimates. Finally, the general research questions and operational definitions were provided along with assumptions that underlie this study.

CHAPTER II REVIEW OF THR LITERATURE

This chapter presents a focused review of previous work pertaining to sustainability education and the development of disposition inventories. Finally this section will also explore if there is a need for the development of a Dispositions Toward Teaching Sustainable Systems Instrument (DTTSSI) exists on the topic of sustainability education and dispositions.

General Background Information

The National Academy of Science, National Academy of Engineering, Institute of Medicine, and National Research Council all acknowledge that environmental and sustainability concerns are becoming prevalent in economic, environmental, and societal systems (Hassan, 2001; National Research Council, 1997). These organizations and many more have suggested sustainability be a major focus for the 21st century (United Nations, 1992). There are multiple definitions of sustainability in the literature. For the purpose of this study sustainability education is defined as the act of presenting knowledge to the current generation to meet their needs without jeopardizing the needs of future generations (Nolet, 2009).

Differing from environmental education, sustainability education focuses on achieving sustainable systems across the three pillars; environment, economics, and society (Nolet, 2009; Orr, 2007; Saylan & Blumstien, 2011). In an effort to increase awareness, environmentalist and educators alike have made numerous efforts in addressing the degradation of these systems within the context of schools. This Education for Sustainable Development, a term coined by the United Nations, aims to increase not only awareness but also the skills attitudes and dispositions necessary to thrive for a

sustainable future (Stibbe, 2009; UNESCO, 1992). Currently institutions of learning are falling short on the translation of sustainability education into measurable outcomes (Feinstein, 2009; Lozano, Lukman, Lozano, Huisingsh & Lambrechts, 2013; Saylan & Blumstien, 2011) Dealing with the complexity of sustainability problems has commonly been seen as going beyond a reductionist approach of teaching to one that is systemic and dynamically represents the complexity of sustainability problems (Heselink & Goldstein, 2000; Morris & Martin, 2011; Rosefsky & Opfer, 2012; Saylan & Blumstien, 2011; Stibbe, 2009). One major component in addressing the complex nature of sustainability in education is the role of the teacher. Historically studies have shown that beyond knowledge and pedagogical knowledge, a teachers' personal dispositions towards particular content areas are critically important because possessing positive dispositions towards given subjects, reforms, or strategies are more likely to result in improved classroom practice (Bamberg & Möser, 2007; Kaiser & Shimoda, 1999; Nespor, 1987).

Sustainability Education

Sustainability education is interpreted in many ways in the literature. The task of identifying the exact history and origins of education for sustainability is complex due to its multiple potential beginnings. Environmental sustainability was first established as national policy in the late 1960's with the National Environmental Policy Act (NEPA). The purpose of this act was to support and promote the ability to create and maintain conditions people can exist in productive harmony with the natural environment and fulfill the social, economic requirements of the present as well as future generations (Cohen & Warren, 1971). Protection for the natural environment was a concern for the public who demanded cleaner air, water, and land use. Sustainability Education,

Education for Sustainability, and Education for Sustainable Development (ESD) are all terms associated with preserving prospects for the future. Arguably, the most common term is education for sustainable development that was developed by the United Nations with the creation of Agenda 21 in 1992. Agenda 21 was the first international document that identified educational practices as a vital component for sustainable development (McKeown, R., Hopkins, C., Rizi, R., & Chrystalbridge, M., 2002). However, many do not consider the Agenda 21 or the NEPA Policy act the beginning of sustainability education in the United States. Disinger (2005) suggests that ESD in America has roots that trace back for over a century to the 1890's nature study movement. During this time nature was a critical component of the curriculum until the gradual progression into the current age of accountability. However, resurgence towards a focus on the environment termed Conservation Education emerged from the agricultural reform projects of the 1920's and 30's. However some authors attribute William Stapp and his students works as founders of the environmental movement (Feinstein, 2009). Yet, after reviewing the literature, it is apparent that regardless of if one traces ESD back to the 1890's or to the 1920's in America, and even with many researchers and educators amplifying conceptual differences between Environmental Education and ESD, (Bonnet, 2002; McKeown & Hopkins, 2003), there is little debate that the current Sustainability Education fields in the United States and the world at large is a direct descendent of the unsustainable practices becoming prevalent in society. In other words sustainability ideas in the 21st century have emerged as more citizens have become more aware of the outcomes of unsustainable practices.

Dispositions

Two major schools of thought exist concerning dispositions in teacher education: those who take the stance where dispositions are fixed or unchangeable, while others believe that dispositions are flexible and can be influenced. For the purpose of this study dispositions are defined as the tendency to act in a particular manner that aligns with an individual's belief which can develop and change over time, and is influenced by previous experiences and the particular circumstances faced by the individual (Brown & Cooney, 1982; Katz & Rath, 1985; Freeman, 2007; Richardson, 2003; Rokeach, 1968; Tabachnik and Zeichner, 1984; Villegas, 2007; Diez, 2006). From this definition we find that tendencies are the inclination towards a particular behavior, and from its observable pattern these behaviors can be predicted. These actions are in essence the driving reason as to why under this definition, dispositions can be measured. Since the 1980's the field of education has experienced a considerable amount of study and development on teacher's beliefs, a component of dispositions. Richardson (2003) has highlighted the work over the last thirty years in this area, and has found that similar to preconceptions students bring to the learning environment, teacher candidates bring in their beliefs. These beliefs to have been shaped by an individual's experiences and hold a critical relationship with the actions teachers take (Richardson & Placier, 2001). Furthermore, the inclusion of both experiences and beliefs are essential in defining dispositions, as they shape the action or implementation of innovative practice. In 2008, Shephard conducted a study identifying affective outcomes such as values, attitudes, and dispositions which serve a mediating role between knowledge and behavior. Shephard's findings, combined with the work of Bamberg and Moser (2007) in the area of behavioral intention and the

work of Kaiser and Shimboa (1999) in the area of personal responsibility influences on pro-environmental behavior, support further investigation in the area of dispositions towards teaching sustainable systems.

The importance of dispositions is well documented in the literature, and has been included in the Interstate New Teacher Assessment Support Consortium (INTASC, 2002) and National Council for Accreditation of Teacher Education standards (NCATE, 2002). NCATE (2002) accredits a significant amount of teacher preparation programs nationwide, and those institutions seeking accreditation are expected to include a performance based system for assessment supportive of the assessment of dispositions, thus, increasing the focus on improving teacher dispositions. More specifically NCATE (2002) standards, postulate that teacher candidates preparing for practice should demonstrate the “professional knowledge skills and dispositions,” to facilitate the learning of all students. However, dispositions as stated in the operational definition are influenced by individuals’ beliefs and experiences and many influential factors exist in the school setting.

Four influential factors have been identified from interviews conducted from the first phase of this study are considered in the scope and sequence of the research. These are: dispositions implementing innovation as it relates to administration, availability of resources, teacher collaboration and professional development. It is apparent that teachers gain experiences that shape their belief from these four influential factors as they mature in the profession and this influences their level of commitment to professional growth (Burden, 1980). Level of commitment to teaching as described by Firestone and Pennell (1993) is not only related to student achievement but also it’s necessary for teachers to

possess the motivation to become professionals and “pursue changes in their practice while dealing with the complex demands these changes present” (Firestone & Pennell, p. 493). In their research, commitment is linked to collaboration, resources, learning opportunities (professional development) and is heavily influenced by the support of colleagues and administrators. This relationship and a direct link to teacher dispositions are underrepresented in the literature and should be investigated further.

Also the call to focus on dispositions is apparent in the field of sustainability education. According to Stibbe (2009), sustainability literacy is defined as “the skills, attitudes, competencies, dispositions, and values that are necessary for surviving and thriving in the declining conditions of the world in ways which slowdown that decline as far as possible” (p. 10) . This definition encompasses not only the term of dispositions, but also related terms such as attitudes and values. Similar to NCATE (2002) who also includes values and attitudes in the definition terminology, dispositions are seen as the preliminary requirement for effective teaching and should be addressed with the necessary skills, content, and pedagogical content in teacher preparation programs.

Sustainability Instruments

The original search using the keywords of sustainability education, dispositions, and instruments resulted in finding no quantitative instruments. Although there were several qualitative questionnaires none of the topics addressed pre-service teacher instruments. Due to the lack of quantitative instruments, psychometric information such as reliability and validity estimates were not reported. There were several searches conducted on using these key words. The first search utilized Google Scholar. The original search word was Sustainability Education and resulted in 1,540,000 hits. When

the word Dispositions was included as a search word the total hits dropped to 639. When the last word instrument was added the total hits dropped to 92. After reviewing these 92 articles it was determined that zero contained quantitative sustainability education dispositions measurement. However the majority of the articles did suggest the need for a disposition inventory designed for sustainability education (Kopnina & Meijers, 2014).

The next database searched was ERIC ProQuest. There were 4273 hits for sustainability education, of which 2937 were peer reviewed. When the key word dispositions was added to the search the total hits dropped to 23 where 14 were identified as being peer reviewed. With the addition of the key word instrument added to the search there were only one hit and zero that had been peer reviewed. The one that was not peer reviewed (O'Farrell, 2010) was determined as having no ties to the topic of focus of sustainability education.

Due to the unavailability of quantitative hits on sustainability education disposition inventories, it was determined that the search would be broadened to just education disposition inventories. There were 215 hits in ProQuest with 129 being peer reviewed. These inventories ranged from the California Critical Thinking Disposition Inventory, who reported psychometrics on this instrument, and found that it did not have sufficient psychometric properties to assess individual abilities nor sufficient stability reliability (Bondy, Koenigseder, Ishee, & Williams, 2001) to colleges, like the Southern Illinois University College of Education Edwardsville, that require every student that applies to the college to take the Student Disposition Inventory. Although the psychometrics for this inventory were not reported, The Texas San Antonio Future Disposition Inventory did report a $\rho < .08$ on all of the subscales and significant

discriminate validity that differentiated between known groups (Osman, Gutierrez, Barrios, Wong, Freedenthal & Lozano, 2010).

The literature review conducted on established instruments measuring dispositions in the area of sustainability revealed a gap in the availability of psychometrically sound measurements. However the search did identify a need for such an instrument (Kopnina & Meijers, 2014). Although there were instruments from the qualitative perspective, there were no quantitative instruments that provide estimates of reliability or validity. These findings indicate there is a need for the development of a sound psychometrically tested instrument with estimates of reliability and validity.

Summary

The aim of this study was to develop an instrument measuring dispositions towards teaching sustainable systems and to establish the instrument's psychometric properties for validity and reliability. A general background and a review of the literature concerning sustainability education and dispositions were presented. Finally, a search was conducted using multiple search engines to identify the availability of established instruments measuring dispositions towards teaching sustainability. This search resulted in determining that no such instrument was available. Finally, a broader search on dispositions in teaching instruments was conducted and presented. From the literature review the researcher concluded that a need exists in the area of measuring dispositions of teachers towards teaching sustainable systems and educating for sustainability. Ethical approvals for all stages of this study were obtained from a university institutional review board prior to data collection.

CHAPTER III METHODS

Research Design

The Dispositions Toward Teaching Sustainable Systems Instrument (DTTSSI) was constructed and developed using a sequential mixed methods approach as presented by Tashakori and Newman (2010). The first phase focused on the construction of the instrument and utilized qualitative techniques and archival data; the second phase employed quantitative measures to estimate reliability (Cronbach's alpha) and construct validity estimates using confirmatory factor analysis.

From Figure 1, a sequential mixed method research design is identifiable by two sets of research questions. In this particular study qualitative research questions were explored and analyzed. Based on thematic analysis items of interview data items were developed to construct the Dispositions Toward Teaching Sustainable Systems Instrument. The second set of research questions was concerned with the psychometric testing of the instrument. Analysis of both quantitative and qualitative data resulted in inferences. From the qualitative inferences coded themes aided in the development of constructs tested in the quantitative inferences. These inferences were then combined to form a meta inference on the psychometrics of the instrument.

Where:

1. QN is quantitative
2. QL is Qualitative

Figure 1: Sequential Mixed Methods Research Design

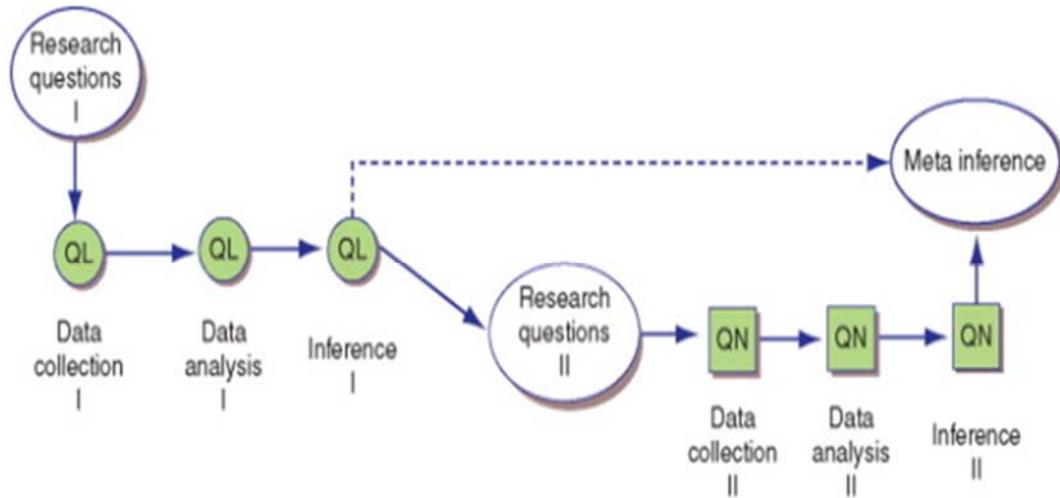


Figure 1 was adapted from Tashakkori, A., & Newman, I. (2010). Mixed methods: Integrating quantitative and qualitative approaches to research. In B. McGaw, E. Baker, & P. P. Peterson (Eds.), *International encyclopedia of education* (3rd ed., pp. 514-520). Oxford, UK:

Using both the sequential mixed methods design and Luyt’s Four Level Framework for instrument development, this research determined estimates of reliability and validity of the Disposition Toward Teaching Sustainable Systems Instrument. A summary of this procedure is presented in the following text.

Mixed Methods for Instrument Development

The aforementioned database searches conducted in chapter two suggested a need for the development of an instrument measuring dispositions towards providing sustainable instruction. Although qualitative instruments exist, no quantitative instruments that report estimates of reliability and validity were found. This section discusses using the mixed methods for instrument development framework as presented by Luyt (2012).

Extending the work of Adcock and Collier (2001), Luyt (2012) recommended that a four level cyclical framework be utilized for measurement development and determining estimates of validity and reliability. The four levels described by Luyt (2012) have been identified as (a) background content (b) systematized concept (c) indicators (items), and (d) quantitative and qualitative analyses. The first level is concerned with finding the broader meaning and understandings associated with the topic of interest. By conducting a focused literature review, special attention is given to the theoretical underpinnings that exist on the topic and the identification of themes are further explored in a related focus group or individual interviews. Focus group or singular interviewed participants are selected using purposive sampling (Patton, 2002), and transcripts from these sessions are coded and the determination of themes using interrater agreement is developed. Finally, a search for previous instruments that have been developed and report acceptable psychometric properties is conducted. From these three sources the researcher identifies commonalities as well as discrepancies in theme development and the analysis informs the second phase. Level two, content systemization, is the narrowing down of specific concepts, and is used to conduct a content analysis. Overarching themes are aligned to named constructs. Once again using a table of specifications, content analysis is determined by how well the constructs are in part supported by the literature, focus group participants/ interviews, and existing instruments' data.

After extensive revision and the identification of specific indicators the third phase is concerned with item development. Using the artifacts generated from the interview statements identified for each construct, items are developed. More specifically focus group statements are transformed into items that elicit responses measuring each

construct. A table of specifications that uses expert judges to determine how well items align with the intended construct is then explored. Furthermore, an opportunity for expert judge feedback is given, and items are revised until 80% agreement among the expert judges is achieved.

From the final iteration of items following expert judge feedback, the final fourth stage is conducted. In this stage both quantitative and qualitative analysis are utilized.

For this study participants enrolled in a 4-year program completed the preliminary instrument to produce numerical scores. These participants were monitored as they progressed throughout the course and as the researcher concurrently investigated the reliability and validity estimates. This is an iterative cyclical process, where more than one feedback loop was sought before finalization into the next phase. By using such a framework as the one presented, the researcher was involved in a mixed method approach that was designed to measure the phenomenon of interest.

Mixed Methods Procedures

Level 1 Background content

Stage 1.

A content review of literature on factors that have been linked to teacher dispositions was conducted.

Stage 2.

Analysis of data concerning purposive sampled interviews conducted with teachers who self-selected into North American Association for Environmental Education sustainability education workshop. The themes that emerge from the interviews were

analyzed using modified grounded theory since there was prior knowledge based on the literature. The emerging themes have been compared to the literature to ascertain overlaps and modifications were completed. Four influential factors were identified and compared to the literature administration, outdoor resources, professional development, and collaboration. Themes were coded using focused coding. The coding of the interviews was conducted using MaxQDA.

Stage 3.

Instrument Review was the final process in this phase and consisted of the comparison between the literature review and the interviewed participant responses.

Level 2 Systematized Concepts

Stage 4.

Based on the themes/domains that emerged and that were confirmed across the previous stages 1-3, constructs were developed.

Level 3 Item Development

Stage 5.

Using the artifacts identified in each construct, items were developed. Interview statements were transformed into items that elicited responses measuring each construct.

Stage 6.

A ToS that utilized expert judges determined how well stage 6 items aligned with the intended construct. Furthermore opportunities for expert judge feedback was given, and items achieved 80% agreement without revision.

Level 4 Quantitative and Qualitative (Validity and Reliability)

Stage 7.

Cronbach's alpha was conducted on the initial administration of data ($N=97$) to estimate internal consistency.

Stage 8.

Estimates of construct validity were assessed using a confirmatory factor analysis (CFA) including overall model fit, measurement invariance, construct invariance and the Binominal Index of Model fit.

Stage 9.

Estimates of discriminant validity was assessed by testing to see if the instrument's subscales account for a significant proportion of unique variance in discriminating prior sustainability knowledge and or training.

Stage 10.

Estimates of measurement sensitivity was assessed by conducting a repeated measures analysis of variance of participants with varied levels of the sustainability education exposure.

Stage 11.

Refinement.

Qualitative Methods

During phase one, qualitative development of themes using archival data was sought through the use of the phenomenology approach presented by Moustakas (1994) and Creswell (1998). This approach is based on a paradigm of personal knowledge and subjectivity that emphasize the importance of personal perspective and interpretation. As

such, this method is powerful for understanding subjective experience, gaining insights into people's motivations and actions. (Creswell, 1998) Phenomenological research has overlaps with other essentially qualitative approaches including ethnography, hermeneutics and symbolic interactionism. However, true phenomenological research seeks to describe rather than explain, and to start from a perspective free from hypotheses or preconceptions (Husserl 1970). Phenomenological research defined by Bogdan and Biklen (2007) states that this approach to research "...is concerned with describing the point of view of the subjects (p.26)." Lester (1999) states that humanist and feminist researchers' influences have promoted the acceptance that starting without preconceptions or bias may not be a possibility, resulting in a move towards emphasizing the importance of transparency in how interpretations and meanings have been placed on the findings. From this perspective it is important to make it known that the researcher is an interested and subjective participant rather than an impartial observer (Lester, 1999). In this study, the researcher's relationship to the situation and participants is that of a participant observer, as the researcher attended and participated in the workshop and conference where interview data were collected. Also, the researcher took the role of the primary data collection instrument as she conducted semi-structure interviews with three workshop participants. Furthermore, the researcher is closely connected to this topic due to continued efforts in researching topics concerning sustainability education as well as including it in practice at various educational levels. Identifying this connection, the research monitored subjectivity using a variety of techniques. Among the techniques practiced was the use of a field journal. The field journal which served as a reflection journal was used to record and recall information from the participant observations as

well as from the semi-structured interviews. The field journal acted as a useful tool in monitoring subjectivity and biases that emerged, by paying close attention to additive comments of the observations and interviews. Furthermore, the researcher used the journal to answer the interviewing questions as a means to identify personal experiences and biases that exist concerning the topic of sustainability. Another strategy used to monitor subjectivity, was the use of the peer review. As the researcher went through the process of data collection and analysis for this portion of the study, she continuously consulted with colleagues who facilitated the logical analysis of data an interpretation of data. Finally, monitored data analysis and interpretation techniques were utilized by coding and recoding the data with colleagues as well as implementing consistent member checking opportunities for the informants to monitor interpretation of data.

This presentation of phenomenological methods has been utilized in this study due its effectiveness at portraying experiences and beliefs of individuals, gathered from their own perspectives and thus enabling it to be used as the basis for practical theory. From this approach practical theory is developed through the use of a nomonological net and networks. A nomonological net is similar and best described as a connection between an item, construct, or artifact to the phenomenon of interest or dependent variable and can be interpreted similar to a path analysis model, where a nomonological network shows multiple relationships such as one would encounter in a structural equation model (Newman, Ridenour, Newman, Smith, & Brown, 2013).

Trustworthiness

The trustworthiness of this qualitative research phase was addressed by considering four criteria commonly used in naturalistic work. The topics of credibility, transferability, dependability, and confirmability are addressed in this section.

Credibility

According to Lincoln and Guba (1985) credibility refers to the congruency between the research findings and the meaning of the participants. In quantitative research this is similar to validity estimates. This study employed the following measures to establish credibility of the findings. Peer debriefing refers to the obtainment of different perspectives from other professionals. By using this technique the researchers can determine if their own subjectivity has presented a barrier in data analysis and/or interpretation. Weekly discussions of the research study were conducted, with various university colleagues. Identified areas of concern or weakness were addressed accordingly. Another strategy, member checking, allowed for the monitoring of accurate data. Participants were emailed transcripts for review for accurate reporting of what was said during interviews. The final strategy that was used to establish credibility in this study is thick rich description. Thick rich description referred to by Lincoln and Guba (1985) is the process of including the “widest possible range of information reported in the study (p.80)”. This research attempted to include rich detail in the setting and scenarios of the study for the reader.

Transferability

Merriam (2002) wrote that external validity is concerned with the extent to which the findings of one study can be applied to other situations. The rich thick description used to convey the findings of this study is provided for the readers'/ practitioners' ease of deciding if the findings relate to their own positions.

Dependability

Dependability is seen as the extent the study is reliable, and replicable (Newman, Mcneil, & Fraas, 2004). The process in this research was reported transparently enabling a future researcher to repeat the work. This is done through the research design and implementation. Also the researcher reflected on the effectiveness of the data collection techniques, and provided a rationale for their use (Shenton, 2004)

Confirmability

Finally confirmability is a measure of how well the findings are supported by the data (Lincoln & Guba, 1985). By answering the interview questions before data collection, the researcher was able to admit to pre-dispositions concerning sustainability education, and exposed bias.

Quantitative Methods

Incorporating the qualitative findings, developed through thematic analysis the second phase of this study was concerned with instrument development. As recommended by Cronbach (1951) and Nunnally (1967), estimates of reliability and validity should be determined before an instrument is implemented in practice. Estimates of reliability and validity are key determinants of the soundness of findings when it is

used in an informative setting. Having multiple estimates of validity and reliability were addressed by the researcher (Cronbach, 1951; Nunnally, 1967). To add to the robustness of the DTTSSI, multiple estimates of validity were provided. Content validity, was explored using a table of specifications and determined by expert judges (Newman et al., 2013). Another estimate, construct validity, is concerned with the DTTSSI's ability to consistently measure thematic constructs that emerge from the theoretical model. Discriminant validity estimates were explored in this study, investigating if the DTTSSI discriminates between participants with prior sustainability experience from their counterparts. Finally a sensitivity analysis were conducted to explore if the DTTSSI is capable of determining differences between pre and posttest scores from participants enrolled in a 3 credit course.

Participants

Qualitative Phase Participants

In the qualitative phase of this study the researcher developed themes that emerged from participants' experiences with teaching sustainability in their fields and their beliefs concerning sustainability literacy. The researcher also identified the various influences in attending the additive workshop as a means of understanding common themes of participants' affiliation and interest in the topic of sustainability. Finally this researcher was interested in participants' experiences concerning teaching sustainability in the classroom, and their beliefs about the characteristics and outcomes of providing sustainability education. Therefore, the participants responses in this phase of the study have been obtained from data solicited in a sustainability workshop conducted during the 2013 NAAEE yearly conference.

Quantitative Phase Participants

A convenience sample was utilized for the psychometric testing of the DTTSSI. Registered pre-service teachers of a public university in South Florida enrolled in the Teaching Elementary Science methods course were required to participate. Participants were actively seeking a bachelor's degree in elementary, early childhood and exceptional student education and are expected to be at various stages of completing their degree. These participants were required to enroll in this course as partial fulfillment of their degree seeking program.

General Research Hypothesis

General Research Hypothesis One (GH1): There are themes that emerge from the interviews of participants who attended a sustainability workshop that lead to the generation of items for the DTTSSI.

General Research Hypothesis Two (GH2): The items (that are generated from the interviews) that comprise the Disposition Toward Teaching Sustainable Systems Instrument are supportive of the content validity as measured by a table of specifications.

General Research Hypothesis Three (GH3) The items (that are generated from the interviews) that comprise the d Disposition Toward Teaching Sustainable Systems Instrument have acceptable internal consistency as measured by Cronbach's alpha.

General Research Hypothesis Four (GH4): The (items generated from the interviews) that comprise the Disposition Toward Teaching Sustainable Systems Instrument support the construct validity estimates as measured by a confirmatory factor analysis.

General Research Hypothesis Five (GH5): The Disposition Toward Teaching Sustainable Systems Instrument supports the construct validity as estimated by the binomial index of goodness of fit.

General Research Hypothesis Six (GH6): The Disposition Toward Teaching Sustainable Systems Instrument supports the stability of the factor structure across time as measured by invariance testing using multi-sample analysis (Schumacker & Lomax, 2010).

General Research Hypothesis Seven (GH7): The DTTSSI demonstrates adequate discriminant validity in measuring disposition differences between those who have had previous sustainability training from those that don't as measured by a simple discriminant analysis and monitored across the 3 assessment points.

General Research Hypothesis Eight (GH8): The DTTSSI demonstrates adequate sensitivity in measuring disposition change in per-service teachers across time as measured by pre and posttest scores using a repeated measures analysis of variance.

Sampling Procedures

Sampling procedures were divided into three periods. The first period of sampling refers to the qualitative aspect of the study. Purposive sampling was utilized in collecting those data and represents participants who self-selected into the workshop on sustainability presented at the 2013 North American Association of Environmental Education and who were in the position of providing instruction to students. Within this phase interrater agreement was explored using expert judges

The quantitative phase of the study used both the second period and third period participants. For the second period, sampling procedures were concerned with contacting

expert judges to complete the table of specifications. This is a purposive sample of professors in the field of science education, sustainability, and or psychometrics. Experts were contacted by email. Two expert judges were used for this study and are delimited to those who agreed to participate and completed ToS responses.

Finally for the third sampling period a convenience sample was utilized. All students enrolled in a Teaching Elementary Science methods course in a university in south Florida were given the opportunity to participate. According to Plichta and Kelvin (2012), a convenience sample is a nonrandom sampling technique. The use of this technique is highly dependent on the researcher's ability in subjectively judging if the sample is representative of the target population. In this study, pre-service teachers who serve south Florida schools are the target population. By including all pre-service teachers enrolled in a science methods course the probability of including members of all segments of the target population are increased. Students, who have graduated from this university, constitute 50% of practicing teachers in the area.

Instrument

Typically this section reviews the psychometric properties of established instruments that was utilized in data collection. Since the purpose of this research study is to develop and report estimates of reliability and validity a general overview of the process is reported.

Qualitative and quantitative techniques were used to develop and establish psychometric properties of the Disposition Toward Teaching Sustainable Systems Instrument (DTTSSI). The DTTSSI was developed in two phases prior to reliability testing. In the first phase qualitative interviews were conducted with practicing educators

who were actively seeking and implementing sustainability education into practice. Data for this phase was collected in the North American Association of Environmental Educators conference. Member checking of the transcripts was extended to the interviewed participants. Member checking allows for the monitoring of accurate data. Participants were emailed transcripts for review to improve accurate interpretation. From these interviews, the second phase will consist of thematic analysis and latent constructs were identified. From the coded themes, literature was explored and used to develop items. From the literature and interviews, items were then organized using a table of specifications (Newman et al., 2013). A table of specifications (ToS) is defined as a set of procedures that attempt to align a set of items, with a set of concepts that are to be assessed. This process was used to establish content validity, and facilitates the development of the confirmatory factor analysis. Finally, estimates of reliability, construct validity, discriminant validity, and a sensitivity analysis were conducted and reported.

Variable List and Instrument Format

The variable list associated with this research was determined after the analysis of qualitative data.

Prior sustainability experience 1= yes 0 = no

Male = 1

Female = 0

In addition to the demographic variables there are 19 items generated from the interviews that addressed the dispositions that comprise the DTTIS. These items are measured on a 5-point disposition scale.

The following is an example of an item:

Given direct administrative support I,

[would ----5----4----3----2----1----would not]

include the topic of sustainability in my curriculum presentation as a generative (overarching) theme in my teaching.

Data Collection

This study used a combination of data gathering techniques and was divided into qualitative and quantitative procedures.

Qualitative Data Collection

Archival data for the qualitative phase is comprised of a document that includes questions concerning demographic as well as follow up contact information, a detailed participant observation of the participants during the workshop as well as transcripts of three semi-structured interviews. Bogdan and Biklen (2007) present the Participant/Observer continuum, which illustrates the importance of the researcher's actions during this data collection period (p.91). The intention of the researcher is to act as a complete observer in the workshop. The complete observer position was not possible, and changes from this technique to one described by Bogdan and Biklen (2007) as fieldwork occurred. Field work describes the researcher's relationship as somewhere between complete observer and complete involvement. One of the strengths of the role of the participant observation as supported by Rubin and Rubin (2012) and Bogdan and Biklen (2007), is that it helps in establishing a rapport with the participants, and facilitated better relationships during the interviewing process. As best stated by Rubin and Rubin (2012), "participant observation easily flows into doing research interviews (p.80)." During the participant observation the researcher was able to record the

vocabulary used in the workshop, describe the major topics discussed, and establish meaningful relationships with the participants, thus increasing her ability to obtain relevant responses during the semi-structured interviews. Finally the use of semi structured interviewing was used in this study. Participants selected were individually interviewed and recorded for approximately 30 minutes. Each interview was conducted within three days of the workshop experience, and was recorded for accurate interpretation and transcription. After each interview, the voice recording was reviewed to inform and tailor the next interview.

Quantitative Data Collection

Similar to the qualitative phase demographic information was collected during the quantitative phase of the study. Data used in this section of the study was collected using the DTTSSI. Participants completed the DTTSSI at three time points in the semester. This repeated measures design is useful in reporting growth trends and stability of constructs across time as well as for the reporting of reliability estimates (McNeil, Newman, & Kelly, 1996). The first administration of the DTTSSI established base line data and was conducted prior to participation in the course. The second administration of the DTTSSI was conducted at the midpoint of the 16 week semester and after the sixth class of the 6 week semester. Finally the last administration was conducted at the completion of the course. Participants were given 15 minutes to respond to the survey items and this was conducted in a group setting. Participants were instructed to respond to all items on the instrument, and the survey administrator provided clarification of the items as needed.

Statistical Treatment

In this section the statistical treatment for all validity and reliability estimates was described. Content validity, construct validity, Binominal Index of Model Fit, measurement invariance, construct invariance, discriminate analysis, and measurement sensitivity analysis were discussed. Finally a discussion of finding estimates of reliability analysis using Cronbach's alpha was presented.

Content Validity

Content validity is defined as the adequacy with which the specified domain of content is sampled (Newman et al., 2013; Nunnally, Bernstein, & Berge, 1967). When finding estimates of validity attention is given to the plan and procedures of construction. The validity of the measure is then judged on the plan and procedures of its development. If there is agreement among the expert judges on the soundness of the plan and procedures it is determined that the instrument is estimated to have a high level of content validity (Nunnally et. al., 1967).

Construct Validity

Construct validity is referred to as an assessment of a measures ability in representing the construct being investigated (Polit & Beck, 2012). Constructs are abstract representations that are created by the researcher as a conceptualization of an unobservable variable using item scores. In essence construct validity is the perceived measure of the overall quality of the instrument, and its ability to measure the latent constructs presented. One method used to test the consistency of a measure's construct and the proposed theoretical framework is through a confirmatory factor analysis (CFA). The CFA is designed to test the multi-dimensionality of theoretical constructs and is

considered a special case of factor analysis (Kline, 2014). A confirmatory factor analysis was conducted using Lisrel 9.1 on the items developed from the interviews that comprise the instrument measuring dispositions toward teaching sustainable systems using preservice teachers. The proposed model was evaluated by three absolute fit measures (chi-square, GFI, and RMSEA). Two relative fit measures (NFI and CFI), and one parsimonious fit measure (PGFI). A significant chi-square means that the reproduction of the correlation matrix based upon the path coefficients of the sample data is significantly different from hypothesized model. A significant chi-square implies that the observed and implied variance-covariance matrices differ and may be due to sampling variation. Due to the chi-square model fit criterion being sensitive to sample size and model complexity, the norm chi-square was also presented (Tabachnick & Fidel, 2007). The goodness of fit (GFI) and root mean square error of approximation (RMSEA) was also interpreted. GFI values of .95 and greater are generally considered acceptable (Hooper, Coughlan, & Mullen, 2008). The RMSEA is a measure of the average size of the residuals between the observed correlation from the sample and expected model estimated for the population, and values below .80 for continuous data are considered acceptable (Schreiber, Stage, King, Nora, & Barlow, 2006). Norm fit index (NFI) assesses the model by comparing the χ^2 value of the model to the χ^2 of the null model, and comparative fit index (CFI) is a revision that takes into account sample size (Hooper, Coughlan, & Mullen, 2008). Values for CFI and NFI range between 0 and 1, 1 indicating perfect fit. Finally, parsimonious fit, which is equivalent to an adjusted R^2 in multiple regression was investigated. The parsimonious fit statistics was used to compare the number of parameters used to identify the model, and closer to 1 indicates a better fit. A

large number of parameter estimates will produce a more likely fit, and acceptable model, due to the number of possible paths.

Shur (2014), suggests an estimate of 5 to 20 participants for each parameter estimate should be considered when determining sample size for a confirmatory factor analysis. The χ^2 general model goodness of fit statistic is problematic with large datasets, but is equally problematic with using a small sample since it might lack the power required to discriminate between good fitting models and poor fitting models (Kenny & McCoach, 2003). In 1977 Wheaton et al. developed a χ^2 that was not affected by sample size. This statistic is called the relative/normed χ^2 and is the χ^2 divided by the degrees of freedom (χ^2 / df). According to Wheaton et al (1977) the recommended range for an acceptable normed χ^2 goes as high as 5.0. However, Tabachnick & Fidel (2007) suggest that a lower maximum value of 2.0 is acceptable. In addition to sample sizes according to findings of a Monte Carlo study on the effects of model complexity and measurement invariances, Cheung and Rensvold (2002) found that despite the correction for degrees of freedom in the GFI, both the GFI and CFI were negatively impacted as the complexity of the model increased. The only measure that was not affected by sample size was RMSEA. This finding should caution researchers when drawing a hard line delineating between a good and bad model fit using a subjectively determined cut scores without taking model complexity into consideration. Both the model complexity and sample size are potential limitations in this study that have to be considered when assessing overall model fit. Therefore these considerations were reported and transparencies of the research findings were presented.

Binomial Goodness of Fit

Another way to assist in determining good model fit is the use of the binomial index of model fit. Most fit indices determine how well the sample data is able to reproduce the correlation matrix. Using these indices alone can be considered sample specific. When testing the model the researcher should be concerned with how well the data supports the theory that underlies the model. According to Newman & Fraas (1994), a researcher should judge whether a path is supported by the data using the following criteria; (a) the parameter estimate for a path exceeds a prior effect size. (b) the parameter estimate is statistically significant (c) the parameter estimate reflects the hypothesized sign and finally (d) a combination of these criteria. For the purpose of this study, two criteria was utilized. The number of paths in the right direction is indicative of overall model fit. To analyze the binomial model fit one must test the actual number of paths supported by the data using the following formula.

$$p(x) = \frac{n!}{x! (n-x)!} (.5)^x (.5)^{(n-x)}$$

Where :

1. p is equal to the probability of obtaining x paths supported by the data out of n number of paths
2. x is equal to a series of numbers ranging in value from the number of paths supported by the data to the total number of paths represented by the model, inclusive.
3. n is equal to the number of paths

Therefore the binomial index is used to determine over all theoretical model fit.

Invariance Testing

A multi-sample analysis to determine invariance of the factor structure was conducted in four phases. An exploration of model invariance over time was conducted to investigate model stability over developmental time points (Little, Preacher, Selig, & Card, 2007). Model A compared three samples by testing the equality of all parameters (factor loadings, error variances, and factor correlations). Model B maintains the assumption of equal variances and factor correlations, but allows the factor loadings of the two samples to be different. Model C maintains equal correlations between the samples, but allows for differences between factor loadings and error variances. Finally Model D specifies that the factor loadings and correlations be the same and allows for error variances to be different. Initially, the quality of all parameters was tested, and the χ^2 , RMSEA, NFI, and CFI were reported. Furthermore individual values contributing to the chi-square were reported.

Discriminant Analysis

Discriminant Analysis was used to test how well a set of variables differentiate between two groups, and how well the identified variables each account for a unique variance over and above the other variable (McNeil et.al., 1996). Running a simple discriminant analysis is comparable to a multiple linear regression when the dependent variable is dichotomous (Kelly & Veldman, 1964; Newman, 1988). Research hypothesis seven is concerned with how well the DTTSSI is able to discriminate between participants with prior sustainability knowledge to those who identify having no knowledge on the subject. This analysis was conducted using a discriminant function analysis and multiple linear regression that tests if there are statistically significant

differences between those who have had prior sustainability experience and those who have not on the total DTTSSI.

Measurement Sensitivity

The final hypothesis is concerned with instrument sensitivity in detecting differences between pre and post testing after treatment. A repeated measures analysis of variance was conducted to determine, if the instrument is able to detect gains in dispositions. Measurement sensitivity analysis is a useful technique when attempting to determine the ability of the instrument to discriminate among participants who represent varying levels of the attribute of interest (Ferketich, 1991). For this research sensitivity was determined if this instrument is sensitive enough to detect significant changes in dispositions using a pre-test-mid test- posttest assessment during a three credit preservice teacher science methods class. This analysis can be conducted in several ways; for the purpose of this study sensitivity analysis was conducted using a repeated measures analysis of variance.

Summary

Chapter 3 described the sequential mixed methods approach that was utilized in this study. Both the qualitative and quantitative methods of participant selection, data collection, and data analyses were described. The qualitative phase utilized a phenomenological approach while the quantitative data analysis has multiple stages and is using a single time point approach to determine estimates of reliability and validity as well as a multi-time point approach to investigate not only the stability of the reliability and validity estimates over time, but also the sensitivity of this instrument to detect changes over time. The general research hypotheses describes the process of how

analysis of the qualitative data was designed to lead to the development of items for the creation of the survey and the psychometric testing of reliability, content validity, construct validity, discriminate validity, and sensitivity analysis. Details of the specific data analyses were provided along with a variable list of known variables at this stage in the research process.

CHAPTER IV RESULTS

The following section or discusses descriptive and inferential statistics. It presents the results by research hypothesis for each procedures identified in the research design section of the previous chapter.

This study was conceptualized in three participant phases. Phase was considered phenomenological in that participants who attended the 2013 NAAEE conference sustainability workshop were interviewed to elicit their beliefs and experiences concerning teaching for sustainability. These interviews were conducted using three purposefully selected participants. The second phase includes expert judge participants who were requested to explore both the alignment of the coded segment for the interviews with the global themes ($n=3$) and the content validity of items generated that measures the latent constructs that were created from the global themes using a table of specifications. Finally the third phase utilized 234 pre-service teachers to empirically identify underlying constructs in estimating its confirmatory support of the theoretical base using a confirmatory factor analysis. The next section describes the participants for each of the phases.

Participants

Phase One Participants

Phase one participants were purposefully selected from the Education for Sustainability Workshop participants at the 2013 annual conference of North American Association of Environmental Educators. There were a total of 10 workshop participants including myself who attended the session. Five of the participants were men and the other five were women. A wide range of ages and levels of experience were represented

in the group. A mixture of professionals attended from those who worked in the field of education to others who worked in environmental education programs in various settings such as fish and wildlife, parks, and recreation services. Only participants who were incorporating sustainability education and worked in the traditional educational field were considered for this study.

Interviewed participants

Ethan is a 29 year old white man. He holds a bachelor's degree in education and has seven years teaching experience. Ethan teaches 5th grade science and his school does not provide a sustainability curriculum. Ethan entered the teaching career through a residency program he describes as "a localized version of Teach for America." When asked about his experiences in this program Ethan describes, "There was no reference whatsoever to sustainability and hardly to science, in general. It's a fast-paced program, much like Teach for America (TFA) to get people in the classroom, and that's what it did." Ethan has an interest in including sustainability in his classroom and describes his efforts as "loosely termed sustainability." He is responsible for the sustainability efforts at his school which includes 735 students across k-8 grades. Ethan's school does have access to outdoor learning environments, and describes this space as "...a fairly good size outdoor space considering the fact that we're a city school." A garden space existed before Ethan's involvement with the program.

Our school five years ago, this is before I was at the school, was entirely pavement in the back of the school, so about 2 acres worth of pavement. Through over a \$100,000 worth of grants, they changed that to roughly football size grass play space. Then, added an acre worth of native beds and vegetable gardens.

Although this space was available for use before Ethan's involvement little or no interaction with garden space was included in the curriculum.

Jason is a 42 year old African-American male. He holds a bachelor's degree in biology, and a master's degree in education with 12 years of teaching experience. He works in the south region of the United States and teaches high school science. Jason's school does not provide a sustainability curriculum but does use an environmental science curriculum which has an embedded sustainability component. When asked about his teaching preparation experiences said;

It was all about the core courses you need for the pedagogy. I have a background, because teaching was a second profession for me; a second career. My degree is in biology, so I have the content but I didn't have the pedagogy. That's really what the teacher training did; that's what my Masters' was really for, just to get me the certification. Nowhere even in my undergrad, I took ecology, but we weren't even using the term sustainability. This is something, for me, relatively new.

Although sustainability is relatively a new term for Jason he is interested in including sustainability in his science classes. Jason's school has access to programs such as Audubon Watershed Experience (AWE). This program offers outdoor learning experiences for students, and is supported at Jason's school. During the interview year 2013 Jason's school was in the process of adding a garden to the schools availability of outdoor learning resources.

Chris is a fifty year old Caucasian male. He holds a Master of Education (M.Ed.) in TESOL, as well as an M.A. in anthropology. He has 22 years teaching experience. He works in the south region of the United States. Chris is a professor of English as a second language and reading at his university, where he is not provided with a sustainability curriculum. When asked about his teaching preparation he states that he

...got the Certificate in Professional Campus Sustainability Leadership at University of Vermont, which wasn't what I thought it was going to be. I thought it was going to be how you reduce your electricity use, and make more efficient use of heating and cooling and stuff, but it was about being a change agent from like solving things. Things like that, which was interesting. It was fine, but it wasn't what I was hoping about energy use and things like that.

He also took an online science course from the American Neurological Society and is a member of the National Oceanic Atmospheric Association, climate steward's education project. Although Chris is involved in continuous professional development, the topic of outdoor resources did not emerge from the semi-structured interview.

Each of the three participants demonstrated his interest in and affinity to include sustainability in his teaching practices. Results of the themes that emerged from the semi-structured interview are presented later in this chapter.

Phase Two Participants

Phase two participants were solicited via email to complete the table of specifications used for content validity. Solicited participants worked in the area of science education, sustainability, or psychometrics. Fifteen expert judges were sent the table of specifications and only two agreed to participate. Participant one is an associative professor in the area of civil engineering with a focus on sustainability at Clemson University. His courses include sustainable restoration, sustainable energy innovation, and sustainable construction. His research interests also extend into sustainability including systems thinking for sustainability, gender and sustainability in engineering, and decision making for sustainability in the built environment. He is also a Leadership in Energy and Environmental Design (LEED) accredited professional. The second expert judge participant served as an assistant professor in science education at

Florida International University and adjunct professor at Florida Atlantic University in the area of environmental education. He currently works as an environmental education consultant in the areas of science and environmental education, project based and constructivist approaches to learning, and geographic information systems (GIS). His involvement in local community projects such as community gardens, research at the Pine Jog Environmental Education center, and Bok Tower Gardens has provided him with many years of experience in sustainability education efforts, initiatives, and critical issues.

Phase Three participants

Phase three participants were selected based on enrollment in the teaching science methods course at Florida International University during the spring 2014 to fall 2015 semesters. Four sections in fall and spring, as well as one section in the summer were utilized. Students in the selected semesters experienced the presentation of sustainability topics as part of their course requirements.

There were 234 participants in this study and as one can see from Table 1 participants were predominantly Hispanic females who reported having no previous experience with the topic of sustainability.

Table 1:

Demographic Information of Participants

Demographic	N	%
Sex		
Male	15	6.4
Female	219	93.2
Ethnicity/Race		
White	40	17.02
African American	24	10.21
Afro-Caribbean	5	2.13
Hispanic	160	68.09
Native American	1	0.43
Other	5	2.13
Experience		
No	189	80.4
Yes	46	19.6

Note. Only 234 out of 235 participants identified sex.

Qualitative Analysis

To explore research question one the transcribed interviews were coded and thematic analysis was conducted. After coding the interviews inter rater agreement among three university education faculty was sought and reported. Raters one and two both had previous sustainability education experience while rater three did not. Each of the raters did have extensive experience in teacher preparation. And as one can see from Table 2 inter rater agreement was above 80% for each of the four latent constructs across raters.

There were four global themes that emerged from the semi-structured interviews that crossed all three of the study participants. These themes are (a) Administrative support (b) Outdoor resources (c) Collaboration and (d) Professional Development.

Operational Definitions of Qualitative Data

- Administrative Support refers to administrative influences to include the school, county, district, state, curriculum, provided resources and school administration such as a principal.
- Outdoor resource refers to outdoor learning environments and outside community resources such as parks, preserves, and community gardens.
- Collaboration refers to the experiences working with other teachers, faculty, and professionals in various fields as well as other community members.
- Professional Development is the ideas, topics, statements toward seeking, evaluating, or participating in experiences to improve practice.

Global Theme 1: Administrative Support

The first global theme administrative support was comprised of three different subthemes that emerged from the interviews. These three segments were, outside support, financial support, and curriculum.

The first coded segment is outside support refers to the availability or lack of availability of provided resources, knowledge of administrative initiative or lack thereof as well as other authoritative influences. The second subtheme under administrative support is financial support which refers to identified support or deficiency in outside monetary support. The third and final subtheme curriculum completes the coded segment that comprises administrative support. Under this subtheme statements concerning the use of provided materials or availability of teacher sought materials were collected and analyzed.

Both Ethan and Jason experienced support of their administration.

Ethan describes coming into his position he had a sense of what was expected from administration. He states “The goal was to get the kids hands- on in that garden space, and get them learning about sustainable systems and sustainability, and how they fit into that.”

Jason saw a less direct role; he aligned his sustainability education initiatives with the provided district materials,

“Within our school district, we have an actual ... We actually have two curriculums. We have one that's for strictly environmental science that has some sustainability components embedded in it, and there's actually one for sustainability. Some teachers actually teach a course in sustainability.”

Chris on the other hand felt there is no outside administrative influence in providing sustainability education to his students. He states “As far as I know, I'm the only professor who has an entire reading course dedicated to global warming.” There were varying levels of outside support.

There were many different revenue streams that the participants navigated to make these educational sustainability programs possible. Only two participants spoke of financial support. Ethan’s interview contained the most of the comments concerning financial support. He stated that there was “over a \$100,000 worth of grants”, and “a lot of donors choose, kind of mini, tiny grants to get more science content books that are not textbooks.” While Jason in contrast only mentioned that financial resources are important and it “gives them those opportunities, to give teachers, because we're all under financial constraints.”

Curriculum resources finish the coded segments of resources. All three of the participants revealed that there were differences in the curriculum resources provided to them. Jason stated that the curriculum for sustainability in his district is,

Embedded within our curriculum for environmental science we have actual kits that are pre-assembled to help kids understand what sustainability is. How they can be a part of it, and it looks at different aspects from acid rain to water quality testing; things of that nature.

In contrast to that Ethan stated that the curriculum “Is it mostly teacher made. Largely, I'd say close to 50/50.” However he mentioned,

The University of Maryland Extension Office has provided me with some resources. They've been very helpful. They've been an Urban Gardener Program, so they will come out to the school and help us weed if nothing else. They have a lot of resources online. Beyond that, I can't think of any off the top of my head. I'm sure they're all bookmarked on my computer.

While both Ethan and Jason have some administrative support with curriculum Chris is developing his resources on his own without outside input. One major resource he utilizes in his course is “Cooler Smarter from the Union of Concerned Scientist.”

Global Theme 2: Resources

As with most educational changes they are not possible without a substantial amount of resources. These resources and support systems are needed for both the initiation of a new program as well as for maintenance of the program. In this research resources emerged as the second global theme. The coded segments that composed the theme resources were student experiences, lesson examples, and outdoor resources environment when presenting the sustainability topics. For example Ethan's school provided a garden.

Outdoor resources were discussed mostly by Ethan and Jason. Each of these two participants was involved in using the outdoor learning environment. Although Ethan's

school has had the garden for five years little to no interaction with it was done until his involvement.

Yeah, it's an exciting new step at our school. We've had the garden for 5 years, but the kids really haven't had that much interaction with it. It's going to be something to look out the window at. That's not the point.

He describes the area as having "... tables and chairs, and describes the student experiences as "They enjoy it. We're still working on the procedures around that. Convincing 8 year olds that being outside doesn't mean it's playtime has been a challenge, and that screaming when they see a bug isn't okay. We're working on being comfortable outside. Part of the pre-k to 4 goal, I see each one of those classes once every 2 weeks. Part of the goal is to just get them engaged and interested in the outdoors."

Jason shares a similar goal however his school is in the process of adding a garden, but prior to its inclusion his outdoor learning environment lessons involved community parks and surrounding area programs.

He believes "The best way to get them involved is with what we call the, AWE program; the Audubon Watershed Experience. It's out of a local park, from Patterson Park in Baltimore. Basically, someone comes out, talks to them about the Chesapeake Bay, what is a watershed, trains them and then, they go to North Point State Park.... we've been doing this for about five or six years."

Jason describes the connection from the AWE program to sustainability themes such as environmental stewardship. He states "That's the environmental stewardship part, and previous experiences, we've canoed and picked up trash. We've gone to a local park; picked up trash. We have planted trees around the school campus."

Although Chris recognizes the importance of outdoor environments, he also recognizes he has limited time to develop this in his presentation. He describes his approach

It's about how individuals can reduce their own carbon dioxide, and heating and cooling, transportation, electricity use at home, heating and cooling at home, what they eat, and their shopping. Then, it goes into how you take local action, and how you can do this at work and get people in your neighborhood. Get other people to join you. We mainly focus on the chapters of reducing the pollution. Just basic time in the semester. There's not a lot, but we have only 20 weeks.

Global Theme 3: Collaboration

Collaboration is the third global theme and is comprised of the coded segments other teacher's efforts, teacher support, and community volunteers. Each of the three participants experienced working with other professionals to present their sustainability topics. Chris describes his experience with helping other faculty in preparing materials and his involvement in various support groups that address sustainability. He is aware of other's efforts in identifying sustainability such as "...putting together a list of classes that have to do with the environment and doing something about the environment, so we have what they call Workforce Development and Continuing Education. You can take a course on designing this whole solar panel arrangement and installing it." Once again Jason and Chris had the most detailed involvement with collaboration. Their experiences are both positive and negative. Ethan explains that he has a limited opportunity to work with others on sustainability. He states "I don't know what a lot of my colleagues are doing at my school. I am loosely connected with a group of teachers in Baltimore City that are more invested in sustainability and science education in general."

Ethan's experience is both positive and negative as he describes,

One or two teachers that are a little more invested than others in recycling. I think is just that they kind of have that in their own lives, and are more tuned to it than others. Still working on getting everybody fully onboard for that kind of thing, of course.

Some teachers, I've kind of gathered semi-resent my position this year and that I'm spending a lot of time outside in the garden and I'm paid the same amount they are. So to speak, not that we get the same amount. It's not the nuts and bolts math and reading, so they don't believe in it as much. Some teachers are extremely narrowly focused on math and reading for testing reasons, primarily.

In contrast Jason experience is mostly positive as he describes

We work together on it. We'll take groups of kids out. If it's something campus wide, we'll get all of the science classes involved and we'll just make a schedule so that they're not missing instructional time, but they get to go out when they would be scheduled for your class.

In Jason's school it is common to see others take the lead in various sustainability efforts.

Jason states

One teacher, who teaches the regular biology students, is actually forming a partnership with the local community garden. It's literally up the street from the school. The kids can get the service learning hours that Maryland requires. You have to have 75 in order to get a high school diploma. The kids can work in the field, literally they can turn over the soil, grow crops and that would be the perfect place for her class and mine, to talk about issues such as sustainability; we do a lot of collaborative planning.

Global Theme 4: Professional Development

The final global theme that emerged from the interviews was professional development. The coded segment that make up professional development are content knowledge, pedagogy what you need, learning goals, and transferability. Each of the participants evaluated their own professional development needs; they also discussed ways that they were involved in improving their own practice. Ethan describes his content knowledge is lacking in the area of science as he states

Well, building my content knowledge about sustainability. That's a big missing piece in probably all teachers, but I focus on elementary science. Within Baltimore City Schools, we're really trying to build science content knowledge, and still we don't have the sustainability piece within our efforts. We're trying to build physics and chemistry.

He further explains how this is an area of focus for him as he is implementing sustainability as common practice in his classroom. When he discusses the district initiative he states

We're not including environmental science area, sustainable science right now. My goal is to further develop mine because it's a day-to-day use for me, but also to hopefully move towards a district model. I can say we are far from sustainable as a system, but also in teaching. We don't have that.

When Chris reflects on his professional development he is fully aware of his strengths and weakness he states

I wasn't always a trained teacher, but my opinion is that people are in the classroom, teaching classes could use teaching training. How to teach. How people learn. Activities. I think that's a good thing for anybody. Now, for some people it might come naturally.

Chris feels preparation is necessary, and spends much of his free time creating his own professional development experiences by attending conference sessions. However he feel these experiences are not practical when describing a session he attended he states

It said the emergence of fracking as an issue offers a teachable moment for building under energy understanding, relating to environmental sustainability, economic development, social justice, and educational dynamics. Ok, so they made their point with that, but no, it's basically people citing other people and researchers and so-and-so wrote this book and this theory and that, no. It was not... but, it did provide an example if you wanted to teach about all those different stuff, but after that, they could have done that in 10 or 15 minutes.

Chris not only identified his own need for professional development, but also addressed others' by providing learning opportunities. He describes "I gave a workshop, we have something called the center for teaching learning, on how to teach about global warming

in class, in any discipline. Maybe it was less “how-to” than showing them it could be done. “

Jason was the only participant who felt he had adequate content knowledge but identified his focus of professional development was in pedagogy. “My degree is biology, so I have the content so I have the content but I didn't have the pedagogy.”

He felt that his hope for professional development would provide more. He describes his expectations for professional development as

If you're gonna be an environmental science teacher dealing with sustainability, then we're gonna show you ... Give you some ... Start you with a bag of tricks." Because, a lot of times, as a year one teacher, you have none. You just come in with the curriculum, a book, and you have to make something happen.

The coded interviews provided many statements which support the global themes of administrative support, resources, collaboration and professional development. Each of the participants had varying experiences and degrees of influence concerning each of the global themes; however they all possessed and demonstrated the disposition towards providing sustainable systems instruction. These statements were then grouped to develop behavior differential items used to comprise the Disposition Towards Teaching Sustainable systems instrument. The following research hypotheses are formulated to provide the estimates of reliability and validity of these items intended to measure the latent constructs.

Table 2:

Global Themes and the Code Segments that Emerged from the Semi-Structured Interviews with Inter-rater Agreement

Global Themes	Coded Segments	Rater One	Rater Two	Rater Three
Administrative Support	<i>Curriculum</i>			
	<i>Outside Support</i>	100%	100%	100%
	<i>Financial Support</i>			
Resources	<i>Student Experiences</i>			
	<i>Lesson Example</i>	100%	97.50%	97.50%
	<i>Outdoor Resource</i>			
Collaboration	<i>Community Volunteers</i>			
	<i>Teacher Support</i>	100%	82.14%	82.14%
	<i>Other teacher efforts</i>			
Professional Development	<i>Content Knowledge</i>			
	<i>What you know</i>			
	<i>Identified need</i>			
	<i>Integrated curriculum</i>	100%	100%	93.75%
	<i>Transferability</i>			
	<i>Learning Goal</i>			
	<i>Pedagogy</i>			

Note. Percentage of interrater agreement was calculated from the number of agreed upon statements divided by total number of statements in each global theme.

Content Validity Estimates

Content validity was explored using a table of specifications. Each expert judge was directed to align survey items to the identified latent constructs. The 19 item instrument was then examined and each rater had the opportunity to provide feedback for each latent construct. As can be seen from Table 3 the overall agreement determined by correctly identified alignment of the item to the construct was over 80% for each of the raters. Therefore no further modifications were necessary.

Table 3:

Table of Specifications Comparing Subscales and Overall Expert Judge Ratings

	Expert Judge 1		Expert Judge 2	
	Number of items Correctly identified	Percentage	Number of items Correctly identified	Percentage
Administrative Support	4 out of 5	80%	3 out of 5	60%
Resources	4 out of 4	100%	4 out of 4	100%
Collaboration	5 out of 5	100%	5 out of 5	100%
Professional Develop.	5 out of 5	100%	4 out of 5	80%
Overall	18 out of 19	94.70%	16 out of 19	84.20%

Reliability Estimates

Cronbach's Alpha was used to measure reliability and assess the internal consistency of the theoretically derived latent constructs as well as the overall instrument at the final data collection point. Four latent constructs were identified; administration, outdoor resources, collaboration, and professional development According to Kline (1999) alpha levels running from 0.7 – 0.80 are acceptable while alphas ranging from 0.6 – 0.7 are questionable, and alphas of 0.5 -0.6 are considered poor. The overall 19 item instrument reported a Cronbach's alpha ranging from $\alpha = .914$ indicating an acceptable internal consistency (Kline, 1999; Nunnally, 1964). The four latent constructs all reported an internal consistency greater than .70.

Table 4:

Internal Consistency for Dispositions Toward Teaching Sustainable Systems total and Subscales at final time point

Subscales	posttest	
	N of Items	α
Administration Support	5	0.829
Resources	4	0.89
Teacher Collaboration	5	0.888
Professional Development	5	0.833
Disposition Total	19	0.949

Construct Validity Estimates

To test the model fit (Figure 2) of a hypothesized Confirmatory Factor Analysis (CFA) model, structural equation models were conducted in LISREL 9.1. The hypothesized CFA model indicates that the items (1,2,3,19,20) measure the latent variable Administrative Support, and items (7,8,9,10) measure the latent variable Outdoor resources, the third latent construct of Collaboration is measured by items (11, 12, 21, 22, 23) and finally Professional Development was measured using items (13,14,16, 17, 18). All latent variables were allowed to correlate. The observed variables were measured by asking pre-service teachers to respond to each specific statements measuring either their dispositions towards teaching sustainable systems as they relate to administrative support, outdoor resources, collaboration, and professional development. Participants responded to the items by placing a response on a behavioral differential instrument. The students were instructed to use any number from 1 to 5, where 1 equals “Would Not” and 5 equals “Would”.

Figure 2: Confirmatory Factor Analysis for Hypothesized Model

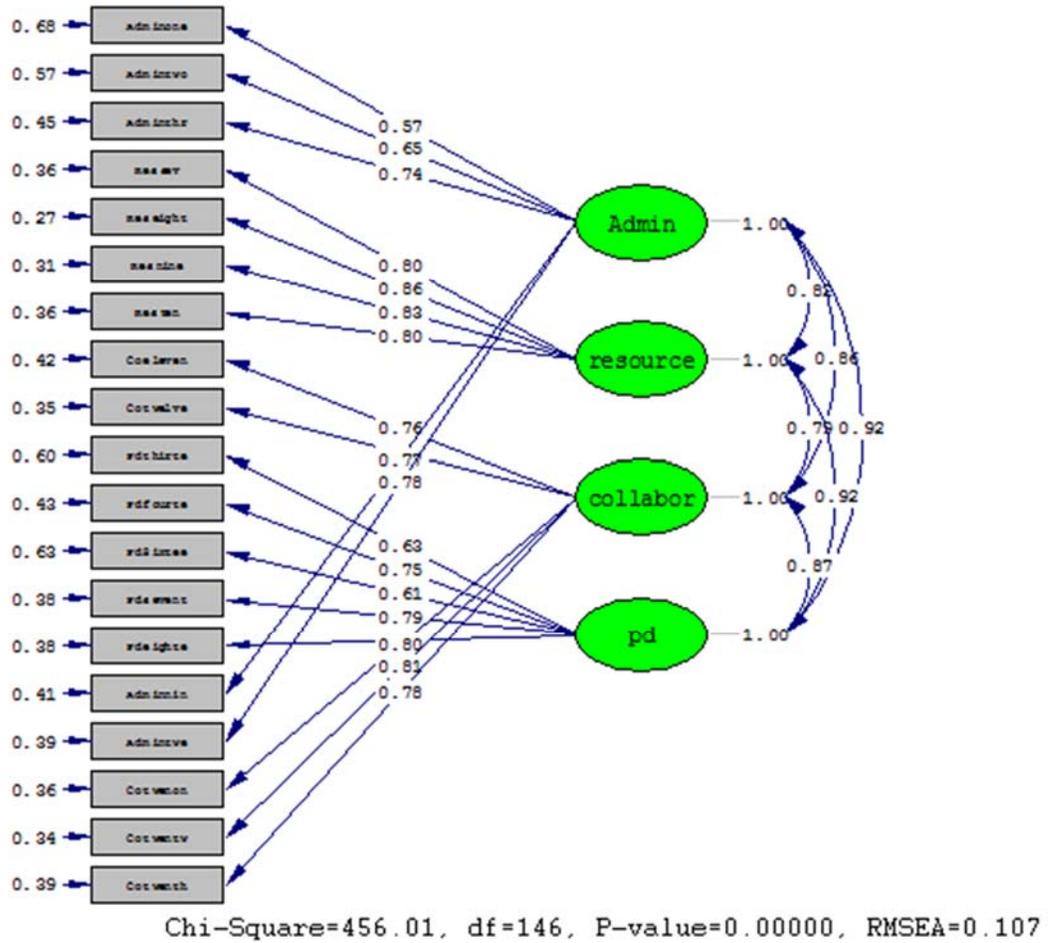


Figure 2. Confirmatory factor analysis of the theoretical model of four correlated latent constructs.

This model was evaluated by three absolute fit measures (chi-square, GFI, and RMSEA). Two relative fit measures (NFI and CFI), and one parsimonious fit measure (PGFI). The model presented in Figure 1 has a significant chi-square ($\chi^2 = 456.01, p > .05$). This means that the reproduction of the correlation matrix based upon the path coefficients of the sample data is significantly different from hypothesized model. A significant chi-square implies that the observed and implied variance-covariance matrices

differ and may be due to sampling variation. Due to the chi-square model fit criterion being sensitive to sample size and model complexity, the norm chi-square is also presented the $\chi^2/Df = 3.04$ indicating an acceptable model fit (Wheaton et al, 1977). The Goodness of Fit (GFI) and Root Mean Square Error of Approximation (RMSEA) were also interpreted. GFI values of .95 and greater are generally considered acceptable (Hooper, Coughlan, & Mullen, 2008). The GFI reported by this model is .786 indicating the proportion of variance in the sample correlation is not accounted for by the predicted model and is not acceptable (See Table 3). The RMSEA is a measure of the average size of the residuals between the observed correlation from the sample and expected model estimated for the population, and values below .80 for continuous data are considered acceptable. The data reported RMSEA of . 0.107, which is an unacceptable fit (Schreiber, Stage, King, Nora, & Barlow, 2006). Norm Fit Index (NFI) assesses the model by comparing the χ^2 value of the model to the χ^2 of the null model, and Comparative Fit Index (CFI) is a revision that takes into account sample size (Hooper, Coughlan, & Mullen, 2008). Values for CFI and NFI range between 0 and 1, 1 indicating perfect fit. The data for the hypothesized model reported CFI of 0.963, which indicates acceptable fit, and NFI of 0.946, which also represents acceptable fit threshold.

The absolute fit measures indicate an unacceptable model fit for the hypothesized model while the relative fit measures are provided to describe the incremental fit of the hypothesized model. Parsimonious fit, which is equivalent to an adjusted R^2 in multiple regression, was also investigated. The parsimonious fit statistics is used to compare the number of parameters used to identify the model, and closer to 1 indicates a better fit. A

large number of parameter estimates will produce a more likely fit, and acceptable model, due to the number of possible paths. The PGFI reported is .6 indicating an acceptable fit.

All the paths in the hypothetical model were direct and significant. None of the error variances were correlated. All paths were in the hypothesized direction and statistically significant. Therefore, all the paths were preserved in the modified model. Further exploration of the direct effect of the hypothesized model was discussed using the binomial index of model fit in research hypothesis

Since the error of the items within the constructs is theoretically correlated modification of the original model was conducted. After the suggested modification was implemented the model was rerun prior to correlating the subsequent error term. Six modifications were made to the model. The first modification correlates the error terms of item 19 and 20. By correlating the error term there was a decrease in chi-square of 48.2. The second modification made to the model suggests correlating the error between items 11 and 12 resulting in a decrease of chi-square of 30.1. The third modification made to the model, as suggested by the modification index is to correlate items 17 and 18, resulting in a decrease in chi-square of 22.6. The fourth modification of the model correlates the error of items 16 and 17, resulting in a decrease of chi-square of 29.8. The fifth modification made to the model, as suggested by the modification indices is to correlate the errors of 13 and 14 resulting in a decrease in chi-square of 10.9. Finally the final modification made to the model, as suggested by the modification indices is to correlate the errors of 13 and 14 resulting in a decrease in chi-square of 8.1

Again three absolute fit measures (chi-square, GFI, and RMSEA), two relative fit measures (NFI and CFI), and one parsimonious fit measure (PGFI), were used to measure

the model fit of the adjusted model. The model presented in Figure 3 has a significant chi-square ($\chi^2 = 296.36$, $p < .05$). The normed chi-square of the adjusted model 2.06 indicating a good fit. The GFI reported by this model is .86 indicating the proportion of variance in the sample correlation is accounted for by the predicted model and is not acceptable but is influenced by low sample size and model complexity (See Table 3). The data reported RMSEA of .07, which is an acceptable fit (Schreiber, Stage, King, Nora, & Barlow, 2006). The data for the adjusted hypothesized model reported CFI of .981, which indicates acceptable fit, and NFI of .965, which is also considered within the acceptable fit threshold. The absolute fit measures indicate an acceptable model fit for the hypothesized model and the relative fit measures are provided to describe the incremental fit of the hypothesized model. Finally the PGFI reported is .634 indicating an acceptable level.

Figure 3: Confirmatory Factor Analysis Adjusted Model

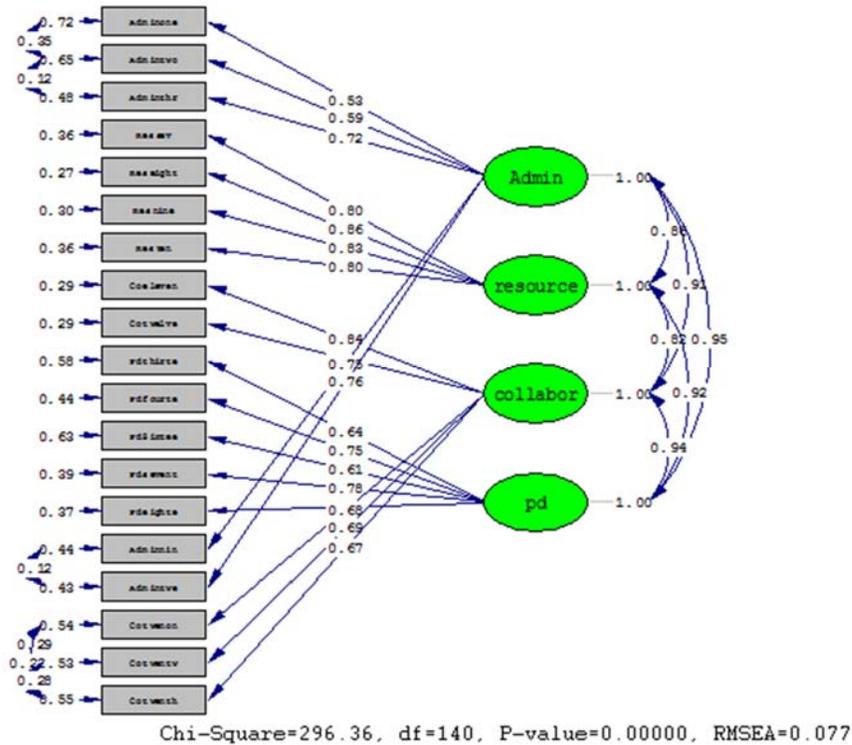


Figure 3. Confirmatory factor analysis of the adjusted theoretical model of four correlated latent constructs.

The adjusted model suggested a good model fit where all the paths in the hypothetical model were direct and significant. Six of the error variances were correlated in the adjusted model. The direct effects for the adjusted model are discussed and reported under research hypothesis three along with the discussion of the results for the binomial index of model fit.

Table 5:

Model Fit Indices for Theoretical and Adjusted Models posttest

Model	χ^2	Df	χ^2/Df	GFI	RMSEA	NFI	CFI	PGFI
Theoretical Model	443.82	146.00	3.04	0.79	0.11	0.95	0.96	0.60
Adjusted Model 1	392.99	145.00	2.71	0.81	0.10	0.95	0.97	0.62
Adjusted Model 2	359.88	144.00	2.50	0.83	0.09	0.96	0.97	0.63
Adjusted Model 3	336.17	143.00	2.35	0.84	0.09	0.96	0.98	0.63
Adjusted Model 4	306.83	142.00	2.16	0.85	0.08	0.96	0.98	0.64
Adjusted Model 5	296.15	141.00	2.10	0.86	0.08	0.96	0.98	0.64
Adjusted Model 6	288.43	140.00	2.06	0.86	0.08	0.97	0.98	0.63

Binomial Index of Model Fit

To test the binomial index of model fit each path in the theoretical model as well as the adjusted model were examined. As one can see from Tables 6 and 7, all paths were direct and significant in the right direction for both the hypothesized and adjusted models. The direct effect of the level of administration to items 1, 2, 3, 19, 20, the direct effect of the level of outdoor resources to items 7, 8, 9, 10, the direct effect to the level of collaboration to items 11, 12, 21, 22, 23, as well as the direct effect of the level of professional development to items 13, 14, 16, 17, 18, all had a statistically significant relationship. To determine the binomial index of model fit the number of statistically significant paths in the right direction was considered in comparison to the number of total items. Using a binomial table the significance is determined where the p value was determined to be less than one out of a thousand ($p < .001$).

Table 6:

Direct Effects and Standardized Regression Weights for Theoretical Model

Endogenous Variable	Factor	Estimate	S.E	P <	Direction
Item 1	← Administrative	0.569	0.07	0.001	+
Item 2	← Administrative	0.655	0.06	0.001	+
Item 3	← Administrative	0.742	0.06	0.001	+
Item 19	← Administrative	0.769	0.06	0.001	+
Item 20	← Administrative	0.78	0.06	0.001	+
Item 7	← Resources	0.803	0.06	0.001	+
Item 8	← Resources	0.856	0.05	0.001	+
Item 9	← Resources	0.833	0.06	0.001	+
Item 10	← Resources	0.798	0.06	0.001	+
Item 11	← Collaboration	0.765	0.06	0.001	+
Item 12	← Collaboration	0.805	0.06	0.001	+
Item 21	← Collaboration	0.798	0.06	0.001	+
Item 22	← Collaboration	0.813	0.06	0.001	+
Item 23	← Collaboration	0.783	0.06	0.001	+
Item 13	← Professional	0.633	0.06	0.001	+
Item 14	← Professional	0.753	0.06	0.001	+
Item 16	← Professional	0.611	0.06	0.001	+
Item 17	← Professional	0.785	0.06	0.001	+
Item 18	← Professional	0.788	0.06	0.001	+

Table 7:

Direct Effects and Standardized Regression Weights for Modified Model

Endogenous Variable	Factor	Estimate	S.E	P <	Direction
Item 1	← Administrative	0.533	0.07	0.001	+
Item 2	← Administrative	0.594	0.07	0.001	+
Item 3	← Administrative	0.718	0.06	0.001	+
Item 19	← Administrative	0.748	0.06	0.001	+
Item 20	← Administrative	0.757	0.06	0.001	+
Item 7	← Resources	0.802	0.06	0.001	+
Item 8	← Resources	0.856	0.05	0.001	+
Item 9	← Resources	0.834	0.06	0.001	+
Item 10	← Resources	0.798	0.06	0.001	+
Item 11	← Collaboration	0.841	0.06	0.001	+
Item 12	← Collaboration	0.841	0.06	0.001	+
Item 21	← Collaboration	0.681	0.06	0.001	+
Item 22	← Collaboration	0.687	0.06	0.001	+
Item 23	← Collaboration	0.671	0.06	0.001	+
Item 13	← Professional	0.645	0.06	0.001	+
Item 14	← Professional	0.745	0.06	0.001	+
Item 16	← Professional	0.609	0.06	0.001	+
Item 17	← Professional	0.781	0.06	0.001	+
Item 18	← Professional	0.797	0.06	0.001	+

Invariance Testing

To test the invariance of the model across time a multi-sample analysis was conducted in four phases using the modified model using each of the three time points Sample A pretest (N = 233) Sample B mid test (N=187) and Sample C post test scores (N = 182). Model A compared the three samples by testing the equality of all parameters (factor loadings, error variances, and factor correlations). Model B maintains the assumption of equal variances and factor correlations, but allows the factor loadings of the two samples to vary. Model C maintains equal correlations between the samples, but

allows for differences between factor loadings and error variances. Finally Model D specifies that the factor loadings and correlations be the same, and allows for error variances to be different.

A CFA was conducted to evaluate a model fit for each time point (see Table 8). This cross validation was evaluated by four absolute fit measures (chi-square, norm chi-square, GFI, and RMSEA), two relative fit measures (NFI and CFI) and one parsimonious fit measures (PGFI). As one can see from Table 8 each of the respective time points indicate acceptable model fit for the exception of the GFI which approaches acceptable model fit at the .90 threshold. Results indicated the adjusted modified model supports the theoretically derived constructs.

Table 8:

Final Model Fit Indices for Time one, Two, and Three

Model	χ^2	Df	χ^2/Df	GFI	RMSEA	NFI	CFI	PGFI
Pre	269.2	140	1.92	0.892	0.063	0.935	0.968	0.657
Mid	259.5	140	1.85	0.88	0.068	0.945	0.974	0.648
Post	288.4	140	2.06	0.86	0.076	0.965	0.982	0.634

Initially, the quality of all parameters were tested and the $\chi^2 = 1169.417$ ($p > .05$), $df = 520$, $RMSEA = .00790$, $NFI = .0.932$, and $CFI = 0.961$ (see Table 9). The individual values for Sample A contributed to $\chi^2 = 4.896$, percentage of contribution = 0.419 and $GFI = .85$. The individual values for Sample B contributed to $\chi^2 = 33.526$, percentage of contribution = .302, and $GFI = .857$. Sample C contributed to $\chi^2 = 2.27$, percentage of contribution = 0.195 and $GFI = .772$ (see Tables 9 and 10).

Secondly, the quality of factor structures were tested and the $\chi^2 = 1049.3$ ($p > .05$), $df = 470$, $RMSEA = .00790$, $NFI = .0.939$, and $CFI = 0.965$ (see Table 6). The

individual values for Sample A contributed to $\chi^2 = 4.801$, percentage of contribution = .458, and GFI = .87. The individual values for Sample B were contribution to $\chi^2 = 3.49$, percentage of contribution = .333, and GFI = .0864 (see Tables 9 and 10). The individual values for Sample C contributed to $\chi^2 = 2.202$, percentage of contribution = .21, and GFI = .799.

A third model was constructed to test the quality of factor structures, and the $\chi^2 = 917.9$ ($p > .05$), $df = 451$, RMSEA = .0072, NFI = .0946, and CFI = 0.972 (see Table 9 and 10). The individual values for Sample A contributed to $\chi^2 = 4.752$, percentage of contribution = 0.518, and GFI = 0.879. The individual values for Sample B were contribution to $\chi^2 = 3.501$, percentage of contribution = 0.381, and GFI = 0.864 (see Table 5 and 6). Finally Sample C contributed $\chi^2 = 2.026$, percentage of contribution = 0.221, and GFI = 0.851. The following model will test the fit by setting the errors free.

A fourth model was constructed to further test the quality of factor structures and the $\chi^2 = 999.8$ ($p > .05$), $df = 482$, RMSEA = .0073, NFI = .0941, and CFI = 0.969 (see Table 9 and 10). The individual values for Sample A contributed to $\chi^2 = 4.856$, percentage of contribution = .486, and GFI = .854. The individual values for Sample B were contribution to $\chi^2 = 3.487$, percentage of contribution = .349, and GFI = .867. Finally sample C contributed to $\chi^2 = 2.073$, percentage of contribution = .207, and GFI = .833. (see Tables 9 and 10).

To determine invariance across time CFI change was considered. A Monte Carlo study conducted by Cheung & Rensvold (2002) found that that the CFI change are a more reliable method for testing group invariance than the likelihood ratio χ^2 change. Their suggestion was to use CFI change with a change of less than 0.01 to indicate model

invariance. As one can see in table 6 there was an invariant structure between factor correlations and errors (Model B) and Factor Loadings & Correlations (Model D) compared to the unconstrained model containing all of the parameters (Model A). Factor Correlation (Model C) approached the acceptable level stated by Cheung & Rensvold (2002) reporting a CFI change = .011. These indices suggests support for the hypotheses that states that there are no significant variances between the unconstrained model, the factor correlations, errors and factor loadings is acceptable.

Table 9:

Contribution of Model χ^2 and GFI

Model	contribution of χ^2	%	GFI
Model A: Group A	4.896	0.419	0.85
Model A: Group B	3.526	0.302	0.857
Model A: Group C	2.276	0.195	0.772
Model B: Group A	4.801	0.458	0.87
Model B: Group B	3.496	0.333	0.864
Model B: Group C	2.202	0.21	0.799
Model C: Group A	4.752	0.518	0.879
Model C: Group B	3.501	0.381	0.864
Model C: Group C	2.026	0.221	0.851
Model D: Group A	4.856	0.486	0.854
Model D: Group B	3.487	0.349	0.867
Model D: Group C	2.073	0.207	0.833

Table 10:

Model Invariance Results reporting χ^2 , RMSEA, NFI, CFI and CFI Change

Global Goodness of Fit Model	χ^2	df	RMSEA	NFI	CFI	CFI Δ
All Model Parameter	1169.4	520	0.079	0.932	0.961	
Factor Correlation & Error	1049.3	470	0.079	0.939	0.965	0.004
Factor Correlation	917.9	451	0.072	0.946	0.972	0.011
Factor Loadings & Correlations	999.8	482	0.073	0.941	0.969	0.008

Discriminant Validity Estimates

To test research hypothesis 7 a discriminate analysis was utilized to test how well differentiate between two groups, those participants who identified as having prior sustainability knowledge and those who reported not having prior experience. When the dependent variable is dichotomous a simple discriminate analysis is comparable to a multiple linear regression is dichotomous (Kelly & Veldman, 1964; Newman, 1988). A multiple linear regression was conducted to investigate if the instrument could discriminate between participants who reported having sustainability prior knowledge from those who did not. As can be seen from tables 11, 12 and 13 this investigation was found to be significant with a reported $R=.271$ $R^2=.073$ $F(4,227) = 4.5$ $p=.002$. As can be seen from Tables 12 and 13 Administrative support was the only variable accounting for a significant proportion of unique variance with a reported $p < .001$.

Table 11:

Means and Standard Deviations of Experience as a Function of Subscales

Experience	No (n=187)		Yes (n=45)		Total (n=232)	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Administrative	18.28	3.75	20.64	3.28	18.74	3.77
Resources	16.47	2.79	16.91	8.19	16.56	4.37
Collaboration	20.88	3.81	21.53	3.07	21.01	3.68
Professional	20.19	5.3	20.67	2.81	20.28	4.91

Table 12:

Correlation of Predictor Variables with Discriminant Functions and Standardized Discriminant Function Coefficients

Predictor variable	Correlation with discriminant functions	Standardized discriminant function coefficients
	Function 1	Function 1
Administrative Support	.912	1.231
Collaboration	.250	.053
Resources	.142	-.407
Professional Development	.136	-.206

Note. Lambda .927 $p=.002$

Table 13:

Simple Discriminant Function Using Regression Testing Experience against No Experience

Variable	B	SE B	β	t	P
(Constant)	-.151	.167		-.905	.367
Administrative Support	.035	.009	.332	4.085	.000
Resources	.001	.006	.014	.207	.836
Collaboration	-.011	.009	-.107	-1.261	.209
Professional Development	-.004	.006	-.054	-.728	.467

Note. $R=.271$ $R^2=.073$ $F(4,227)=4.5$ $p=.002$

Measurement Sensitivity

To explore the final research hypothesis a repeated measures analysis of variance was conducted. A repeated measures analysis of variance was used to answer if the instrument was able to detect significant difference between participants at each time point. This technique can be used when samples are measured under a number of different conditions and allows for the dependent variable to be measured repeatedly (Field, 2013). A well supported strength of repeated measures is its ability to reduce errors associated with the within-group variance and eliminate systematic bias resulting

from individual differences among participants. Another advantage of repeated measures analysis of variance is that a smaller sample can be used (Stevens, 2009) because this method measures the same subjects using more than one data collection point and increases power. As can be seen from Table 14 the instrument was able to detect statistically significant differences across time points as a whole as well as in each subscale.

To estimate the sensitivity of the instrument in detecting change in dispositions attention is given to the reported effect size of each subscale as well as the disposition total score. Effect size refers to the magnitude of the relationship of one variable on some other variable. Although there are many ways to estimate effect size the use of eta squared is utilized in this study. Eta squared is a measure of variability of time points that can be attributed to the change in disposition. Small, medium and large eta squares are determined as .01, .06, and .14 respectively (Cohen, 1977; McNeil, Newman & Kelly, 1996). As one can see from table 14 the disposition total reports a large eta squared along with the subscales administrative support, collaboration, and professional development. Resource is the only subscale to report an eta that approached a medium effect size (see Table 14).

Table 14:

Repeated Measure Analysis of Variance Reporting the Discriminant Capability of the DTTSSI Instrument

	Pretest		Mid Test		Posttest		F	p	η^2
	M	SD	M	SD	M	SD			
Total Score	75.85	12.48	79.82	11.22	84.13	11.11	34.52	<.001	0.31
Administrative	18.48	3.84	20.87	3.08	22.26	2.84	73.3	<.001	0.49
Resources	16.52	5.00	16.59	2.89	17.2	3.26	4.46	0.013	0.05
Collaboration	20.72	3.96	21.9	3.15	22.66	2.85	17.98	<.001	0.19
Professional	20.01	3.78	21.1	2.96	22.1	3.34	20.15	<.001	0.21

Summary

Chapter 4 began with a description of participants in each of the three phases followed by the results of the eight research questions. Qualitative analysis was then described along with inter-rater agreement of the coded themes to the global themes that make up the latent constructs. Content validity using a table of specifications and expert judges' feedback was presented. Reliability estimates were then explored using Cronbach's alpha. The next phase presented validity estimates to include estimates of construct validity using a CFA and binomial index of model fit were explored along with model invariance testing, discriminant validity using discriminant function analysis, and measurement sensitivity tested through a repeated measures analysis of variance.

CHAPTER V SUMMARY, DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

This chapter presents a brief summary of the study. The sections include the restating of problem and purpose of the study. Furthermore it provides an overview of the methodologies, conclusions, and discussion of the eight research questions. This is then followed by implications, limitations, and finally recommendations for further research.

Summary of the Study

Increasing concerns towards addressing sustainability issues have led to a heightened focus on sustainability education. Although many grassroots initiatives have made their way to the school setting, widespread adoption towards teaching sustainable systems are not yet common practices in American schools (Feinstine, 2009; Heselink & Goldstein, 2000; Nolet, 2009 ; UNESCO, 2004). One key determinant in school reform is practitioner buy in and dispositions towards adopting innovative methods, programs, or teaching strategies. In the case of sustainability education, a major issue in evaluating programs and increasing the likelihood of implementation in the classroom is the lack of available disposition measures. Without disposition measures it is unlikely to determine if practitioners have the positive disposition towards adopting its practice. Furthermore current efforts in improving dispositions toward the inclusion of sustainable teaching have no way of evaluating outcomes or monitoring disposition change over time. This research developed and investigated estimates of validity and reliability of the Disposition Toward Teaching Sustainable Systems Instrument. A sequential mixed methods research design was used to construct and test the psychometric properties of the

instrument. The initial qualitative stage included three purposefully selected participants who were actively involved in sustainability education. The phenomenological qualitative interviews were coded resulting in the emerging of four global themes: administrative support, resources, collaboration and professional development. Interrater agreements of the qualitative coded segments were explored and the generated items were given to expert judges to align to the global themes using a table of specifications exploring content validity.

The second set of research hypothesis investigated the psychometric soundness of the instrument. A focus on estimates of reliability using Cronbach's alpha was conducted on the overall instrument as well as each individual data collection point. Construct validity estimates were explored using a variety of techniques. A confirmatory factor analysis, binomial index of model fit and model invariance, were utilized to determine construct validity estimates. Using discriminant function analyses, discriminant validity was tested to determine the instrument's ability in differentiating between participants with varying levels of experience with sustainability. Finally, measurement sensitivity was determined using a repeated measures analysis of variance due to its ability to detect changes over time as well as reports of the accounted-for-variance of each construct.

Methodology

Research Design

This investigation utilized a sequential mixed methods approach for instrument development. Both qualitative and quantitative methods and two sets of research questions were identified. The qualitative research question generated the theoretical model of the instrument and was then followed by quantitative research questions to test

the soundness of the theory. This mixed methods approach aligned with a modified approach to Luyt's framework for instrument development allowing for the iterative process of refining the instrument and cross validating the results. This sequential approach was employed to increase the likelihood of measuring the phenomena of interest.

Data Sources

The data sources for this research were derived from three independent resources. In the first phase, a purposive sample of three interviewed participants who attended the NAAEE Conference and who self-selected into the Education for Sustainability workshop was utilized. These participants were actively engaged with adding sustainability in their teaching practice. In the second, a purposive sample of two expert judges was utilized to conduct content validity estimates using a table of specifications. In this phase expert judges in the areas of sustainability education were asked to evaluate the instrument. The final phase used a convenience sample of pre-service teachers who were actively enrolled in a one semester Teaching Elementary Science methods course across three semesters. The participants ($n=233$) were presented a sustainability curriculum as part of their regular course curriculum. Data were collected at three time points throughout their semester.

Statistical Analyses

Descriptive and inferential statistics were utilized in the study. The qualitative research hypotheses were explored using phenomenological methods where member checking and interrater agreement were sought to provide credibility of the findings (Creswell, 1998; Lincoln & Guba, 1985). The quantitative methods explored various tests

to determine estimates of validity. Content validity was explored using expert judge feedback and a table of specifications, allowing for the research to explore content and construct validity (Newman et al., 2013). Further testing of construct validity using a confirmatory factor analysis and Binomial Index of Model Fit was explored to determine the goodness of model fit and theoretical model fit using a variety of fit measures (Newman & Fraas, 1994; Polit & Beck, 2012). The stability of the factor structure over time was also determined using invariance testing. CFI change was calculated to determine the stability of the factors using three different model scenarios compared to an all parameter structure (Cheung & Rensvold, 2002). Discriminate function analysis was utilized to further investigate the instruments ability to determine differences between participants having prior sustainability experience from those who reported none (Kelly & Veldman, 1964; Newman, 1988). Finally, measurement sensitivity was investigated to determine the instruments capability in detecting changes in dispositions using repeated measures of analysis of variance with special attention to the effect size of Eta squared to determine the amount of accounted for variance of each subscale (Field, 2013).

Research Questions, Conclusions, and Discussion

This section is organized by the general research questions. Each research question is addressed individually along with the conclusions and discussions. The research questions are then followed by a global discussion concluding this section.

Research Question 1

Will there be themes that emerge from the interviews of participants in a sustainability workshop that could lead to the generation of items for the DTTSSI?

(Qualitative)

The coded interviews provided many statements which support the global themes of administrative support, resources, collaboration and professional development.

Interrater agreement was over 80% in each of the categories across all three raters. From the analysis it is apparent that each of the participants had varying experiences and degrees of influence concerning each of the global themes; however, they all possessed and demonstrated the disposition towards providing sustainable systems instruction.

These statements were then grouped to develop behavior differential items used to comprise the Disposition Toward Teaching Sustainable Systems Instrument (DTTSSI).

Research Question 2

Do the items (that are generated from the interviews) that comprise the Disposition Toward Teaching Sustainable Systems Instrument support the content validity as measured by a table of specifications? (Qualitative)

To investigate the content validity, expert judge feedback was determined using a table of specifications. Fifteen invitations to participate were extended however only two were returned in the timeframe the study. The overall instrument had an agreement rating of over 80% from the two expert judges' responses. Some similarities between raters existed. For example each rater identified item 3 as a resource item rather than an item of administrative support. This item did not directly identify administrative support but did use resource in the text. This similarity among the judges suggests that administrative

support was not a clear underlying concept measured by item 3, but rather supported the construct of outdoor resources. Furthermore both raters expressed how each item could fit in multiple constructs and this supported the correlation of the constructs in the confirmatory factor analysis. Although raters were to select one construct per item it was common to find that two constructs were selected for some items. To determine the percentage of scores, consideration was given to how the rater aligned items to the theoretical construct. The number of items the rater identified and aligned to the theoretical construct was then divided by the total number of items in the theoretical construct. Feedback opportunities for each construct were also provided by both raters. Administrative support included the comments of exploring the influences of high stakes testing on sustainability initiatives as well as removing the term “generative” to improve readability of the items. Professional development comments were both positive and reflective. Rater 1 suggested looking into creating items related to teacher training programs. While rater 2 extended his ideas of how items seemed to reflect more than one latent construct such as professional development and collaboration. Both raters were insightful towards elaborating the development of the latent construct of outdoor resources. Suggestions such as removing the term “garden,” and specifics such as “monthly” would be beneficial. Furthermore suggestions such as include expanding resources to including more indoor resource topics as well. Finally, collaboration was highly rated by both raters; suggestions at looking into the Piedmont Project and other sustainability initiatives were discussed (Eisen & Barlett, 2006). These considerations will serve as a resource when modifying future versions of the instrument.

Research Question 3

Do the items that comprise the Disposition Toward Teaching Sustainable Systems Instrument demonstrate adequate estimates of reliability as measured by Cronbach's alpha? (Quantitative)

To investigate internal consistency, Cronbach's alpha was conducted on the overall dispositions scale as well as each individual subscale. The overall scale shows an above acceptable estimate of reliability reporting $\alpha=.914$. This indicates that the overall instrument items are consistently measuring the participant responses. A breakdown of each individual subscale was also explored. Each of the subscales reported an α above .8 indicating a high estimate of internal consistency. Using Cronbach's alpha it was determined that the instrument has above acceptable estimates of reliability. However reliability alone does not indicate that a measure will be useful in practice, further investigation into validity estimates were necessary as reliability is only a precursor to having acceptable estimates of validity.

Research Question 4

Do the items generated from the interviews that comprise the Disposition Toward Teaching Sustainable Systems Instrument demonstrate adequate estimates of construct validity as measured by a confirmatory factor analysis? (Quantitative)

To investigate if the instrument items support construct validity, a confirmatory factor analysis was conducted using structural equation modeling. Three absolute fit measures, two relative fit measures, and one parsimonious fit measure was used to determine the overall model fit. In the theoretical model only one of three measures indicated acceptable model fit. This was the normed chi square =3.04. RMSEA and GFI

both indicated an unacceptable fit. Both the relative fit measures NFI and CFI and the parsimonious fit measure indicated that the theoretical model was a good fit. However since items within the subscales are expected to measure the same constructs, modifications to the theoretical model were made to account for these relationships. Six modifications were made to the theoretical model. These modifications allowed for the correlation of the error terms between items within the same construct. Only two constructs received modifications improving the overall fit of the model, where all model fit measures with the exception of GFI were considered acceptable. This indicated that the proposed model had acceptable estimates of model fit. However, further testing of the model for invariance over time is an indicator of how stable the latent constructs are over the developmental growth of the participants.

Research Question 5

Does the Disposition Toward Teaching Sustainable Systems Instrument demonstrate acceptable estimates of construct validity as measured by the Binomial Index of Model Fit? (Quantitative)

The fifth research hypothesis investigated the overall goodness of fit estimate for the Disposition Toward Teaching Sustainable Systems Instrument. For this research question, the estimate was calculated using the Binomial Index of Model Fit on both the theoretical model and adjusted model. All of the theoretically proposed and adjusted paths were found to be statistically significant and in the predicted direction (nineteen out of nineteen). The likelihood that this would occur by chance is one out of a thousand ($p < .001$). Therefore the Binomial Index of Model fit supports the theoretical model as well as the adjusted model (Newman & Fraas, 1994). This measure indicates that the

Disposition Toward Teaching Sustainable Systems Instrument's theoretical and adjusted model can be effective in assessing teacher dispositions towards providing this instruction.

Research Question 6

Does the Disposition Toward Teaching Sustainable Systems Instrument support stability of the factor structure across time as measured by model validation using invariance testing? (Quantitative)

To investigate invariance of the factor structure across time a four phase model testing was employed. To confirm model fit held up across time, a CFA using the modified model was conducted on all three time points and compared. Each of the invariance models were then explored using CFI change as an indicator of acceptable invariance. CFI change below .01 is an indicator of invariance. Model A tested the all parameter structure and was used to compare models B, C, and D. Model B assumes the factor correlations and error variances are equal but allows the factor loadings to vary and reported a CFI change of .004 indicating invariance. Model D allows for the factor loadings and correlations to be equal and for the error variances to vary and reported a CFI change of .008 also indicating invariance. Model C, the most stringent of the models, must maintain the factor correlations however allows only for the differences between the factor loadings and error variances to vary and reported a CFI change of .011. Although this score slightly exceeds the .01 subjectively stated cut score, it was determined to have invariance, since it is not a test of significance.

Research Question 7

Does the DTTSSI demonstrate adequate discriminant validity in measuring disposition differences between pre-service who have had previous sustainability training from those who have not as measured by a simple discriminant function analysis?

(Quantitative)

To investigate the capability of the DTTSSI in differentiating from participants who reported having sustainability experience to those who reported having none, a discriminate function analysis was conducted. The instrument was able to detect differences between these two groups. However, only one subscale, Administrative Support, accounted for a significant amount of unique variance where $p < .001$. This finding may be influenced by the self-reported nature of the question. This is a promising finding despite the following hindsight considerations. The criterion of what constitutes experience versus non-experience was not defined. Furthermore Box's M was significant indicating unequal groups and to determine the discriminate validity estimate these two considerations should be taken into account in future studies.

Research Question 8

Does the DTTSSI demonstrate adequate measurement sensitivity (known group validity) in measuring disposition change in pre-service teachers across time as measured by pre-test and posttest scores using repeated measures analysis of variance?

(Quantitative)

The final research question investigated if the DTTSSI had adequate measurement sensitivity. A repeated measures analysis of variance provided the researcher two pieces of information to determine the instrument's sensitivity. First a look at whether the

instrument was able to detect changes over time and second a look into the magnitude of the relationship. The DTTSSI was capable of determining statistically significant changes over time, and three of the four subscales reported accounted for a large amount of variance. The resource subscale was the only area where the findings indicated a small effect. However due to the population being limited to pre-service teachers inadequate experiences with school resources can be seen as a limitation and should be investigated further.

Global Discussion of the Research Questions

The eight research questions in this study were derived to construct and explore estimates of validity and reliability of the Dispositions Toward Teaching Sustainable Systems Instrument. The theoretically derived latent constructs proposed by the phenomenological analysis was assumed to exist and was tested to explore the estimates of reliability and validity. The results supported the model in that the relationship of the items to each of the four latent construct were found to have interrater agreement, content validity, supported construct validity and stability of the factor structure over time. Also preliminary investigations indicate the instrument is able to detect differences in sustainability experience at the pretest time point, however further investigation to determine to what extent would need to be explored. This is due to the findings that administrative support accounted for a significant proportion of variance in detecting differences between experiences. Finally the results supported the models ability in detecting changes in disposition with a large effect in three of the four subscales. As these results indicate the Dispositions Toward Teaching Sustainable Systems Instrument therefore has a strong overall support of sound psychometrics and can be used for

program evaluation, curriculum development, and monitoring individual disposition change in the area of sustainability education.

Implications

This research is critically important because there are no available instruments measuring dispositions towards providing sustainable systems instruction currently available in a rapidly growing field. Dispositions towards teaching have been reviewed extensively in the literature and can be seen as a mediator in practicing innovative teaching as well as pro-environmental behavior (Bamberg and Moser, 2007; Kaiser & Shimboa, 1999; Katz & Rath, 1985; Freeman, 2007; Richardson, 2003; Rokeach, 1968; Tabachnik and Zeichner, 1984; Villegas, 2007; Diez, 2006). Teacher education programs have focused efforts in measuring dispositions and its influence on the quality of educational practice, and these efforts should be extending in the area of sustainability education. Without the availability of instruments with acceptable reliability and validity estimates, programs lack the key determinants to improve, evaluate, and provide formative and summative feedback of their efforts.

Funding to continue improving the presentation of sustainability education requires these evaluative measures be utilized to provide the insight in attending to the influence of improvement in the field. This being said, this research study was conducted to construct an instrument and to find estimates of validity and reliability to add to the field of sustainability education. There was overwhelming support for the use of the Dispositions Toward Teaching Sustainable Systems Instrument as an indicator of teacher dispositions.

Limitations

The research results had above acceptable estimates of reliability and validity of the Dispositions Toward Teaching Sustainable Systems Instrument; however, there were limitations to this study. One limitation was concerning the participants for each of the phases. For the first phase the availability of practicing teachers who were implementing sustainability was scarce. Out of the 10 participants, five, were practicing teachers and three agreed to participate. Another limitation was related to participants in the second phase. Fifteen emailed invitations for expert judge feedback for content validity were distributed; only two accepted and followed through with participation. Several potential expert judges did not respond, while others expressed not feeling comfortable to be a rater either because they were in science education rather than in sustainability education or because they were unfamiliar with psychometric and instrument development. The fragmentation mind frame ruling disciplines may have hindered the collaboration of professionals in developing this instrument. The next limitation concerning participants was the use of pre-service teachers. These participants have varied experience with the influential factors that exist in the school setting, and including practicing teacher dispositions would further strengthen the practicality of the instrument. Finally, the self-identified sustainability experience data could have been defined with more fidelity. As participants began the course they associated having experience with sustainability with varying degrees from hearing about it and recycling, to having a course on the topic. As suggested by the expert judge differentiating sustainability knowledge using assessment scores such as the Ohio Assessment of Sustainability Knowledge (ASK) survey may

serve as a better indicator of discriminating between experience and non-experience (Zwickle, Koontz, Slagle, & Bruskotter, 2014).

Recommendations for Further Research

The presented data supports many opportunities to extend the research initiatives. This section presents some of the suggested options for further study.

- An investigation exploring predictive validity estimates is recommended as a critical component of the instrument. Data concerning if those who reported positive dispositions toward providing sustainable systems instruction did in fact implement the topic in practice.
- An analysis of diverse subgroups to investigate how well latent constructs hold for different populations.
- A critical look into discriminant validity estimates using the Ohio Sustainability Assessment of Sustainability Knowledge (ASK) as an indicator of sustainability experience, and determine if the instrument is able to detect differences between participants' scores (Zwickle, et.al, 2014).
- An investigation into the relationships between instruction from teachers with positive dispositions towards providing sustainable systems instruction and the influence on students pro-environmental behavior, sustainability knowledge gain, local community environmental health, and community involvement in addressing sustainability issues are only a sample of variables that can and should be explored.

Summary

This research used a sequential mixed methods approach to develop the Disposition Toward Teaching Sustainable Systems Instrument. In the qualitative phase a phenomenological approach was used to explore experiences and beliefs of teachers concerning implementing sustainability education in practice. These transcribed interviews were then coded and analyzed formulating four latent constructs. Administrative Support, Resources, Collaboration, Professional development emerged as global themes and interrater agreement among three raters were above 80%. Content validity was then explored using two expert judges in the field of sustainability and a table of specifications. After reaching 80% agreement between the two expert judges led to investigating construct validity estimates using both a CFA and binomial index of model fit, invariance testing, discriminant validity estimates, and measurement sensitivity were conducted. All hypotheses were found to be significant in supporting the Dispositions Toward Teaching Sustainable Systems Instrument as a psychometrically sound tool. Implications concerning its application in practice were discussed as well as future suggestions for further research with consideration to the current studies limitations. One major consideration that should be noted is the value of exploring predictive validity estimates of the instrument. These results support the use of this instrument not only in evaluating dispositions in pre-service teachers towards innovative sustainability education teaching practices but also as an indicator of program effectiveness in influencing those dispositions.

LIST OF REFERENCES

- Adcock, R., & Collier, D. (2001). Measurement validity: A shared standard for qualitative and quantitative research. *American Political Science Review*, 95(3), 529..
- Atkin, J.M, & Black, P.(2007). History of science curriculum reform in the united states and the united kingdom. In S. Abell & G. Leaderman (Eds.) *Handbook of research on science education*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Bamberg, S., & Möser, G. (2007). Twenty years after hines, hungerford, and tomara: A new meta-analysis of psycho-social determinants of pro-environmental behaviour. *Journal of Environmental Psychology*, 27(1), 14.
- Barr, S. (2003). Strategies for sustainability: Citizens and responsible environmental behaviour. *Area*, 35(3), 227.
- Bogdan, R., & Biklen, S. K. (2007). *Qualitative research for education: An introduction to theories and methods*. New York, NY: Pearson Allyn & Bacon.
- Bondy, K. N., Koenigseder, L.A., Ishee, J.H, & Williams, B.G. (2001). Psychometric properties of the california critical thinking tests. *Journal of Nursing Measurement*, 9(3), 309.
- Bonnett, M. (2002). Education for sustainability as a frame of mind. *Environmental Education Research*, 8(1), 9.
- Brown, T. A. (2012). *Confirmatory factor analysis for applied research*. New York, NY: Guilford Press.
- Brown, C. A., & Cooney, T. J. (1982). Research on teacher education: A philosophical orientation. *Journal of Research and Development in Education*, 15(4), 13.
- Burden, P. R. (1980). Teachers' perceptions of the characteristics and influences on their personal and professional development. Retrieved from:
<http://files.eric.ed.gov/fulltext/ED198087.pdf>
- Brundiers, K., Wiek, A., & Redman, C. L. (2010). Real-world learning opportunities in sustainability: From classroom into the real world. *International Journal of Sustainability in Higher Education*, 11(4), 308.
- Cantor, N. (1990). From thought to behavior: "Having" and "doing" in the study of personality and cognition. *American psychologist*, 45(6), 735.
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9(2), 233.

- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3), 309.
- Cohen, B. S., & Warren, J. M. (1971). Judicial recognition of the substantive requirements of the national environmental policy act of 1969. *Boston College Law Review*, 13 (4) 685-704.
- Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (2nd ed.) New York, NY: Academic Press.
- Colucci - Gray, L., Camino, E., Barbeiro, G., & Gray, D. (2006). From scientific literacy to sustainability literacy: An ecological framework for education. *Science Education*, 90(2), 227-252.
- Council for the Accreditation of Educator Preparation (2013). Council for the Accreditation of Educator Preparation CAEP accreditation standards and evidence: Aspirations for educator preparation Recommendations from the CAEP Commission on Standards and Performance Reporting (2013) Retrieved from <http://caepnet.files.wordpress.com/2013/02/commrpt.pdf>
- Creemers, B. P. M., & K. (2006). Critical analysis of the current approaches to modelling educational effectiveness: The importance of establishing a dynamic model. *School Effectiveness and School Improvement*, 17(3), 347.
- Creswell, J. W. (1998). *Qualitative inquiry and research design : Choosing among five traditions*. Thousand Oaks, Calif: Sage Publications.
- Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334.
- Diez, M. E. (2006). Assessing dispositions: Context and questions. *New Educator*, 2(1), 57-72.
- Disinger, J.F. (2005). The purpose of environmental education: perspective of teachers, government agencies, NGOs, professional societies, and advocacy groups. In E.A. Johnson & M.J. Mappin (Eds.), *Environmental Education and Advocacy: Changing perspectives of ecology and education*. Cambridge: Cambridge University Press.
- Eisen, A., & Barlett, P. (2006). The piedmont project: Fostering faculty development toward sustainability. *Journal of Environmental Education*, 38(1), 25.
- Feinstein, N. (2009). Education for sustainable development in the United States of America. University of Wisconsin-Madison School of Education: International Alliance of Leading Education Institutes. Retrieved from www.k12.wa.us/EnvironmentSustainability/.../USAreport7-19-09.pdf

- Ferketich, S. (1991). Focus on psychometrics aspects of item analysis. *Research in Nursing & Health*, 14(2), 165.
- Firestone, W. A., & Pennell, J. R. Teacher commitment, working conditions, and differential incentive policies. *Review of Educational Research*, 63(4), 489.
- Fox, R. J. (1983). Confirmatory factor analysis. *Wiley international encyclopedia of marketing* John Wiley & Sons, Ltd.
- Fraas, J. W., & Newman, I. (1994). A binomial test of model fit. *Structural Equation Modeling*, 1(3), 268-273,273.
- Freeman, L. (2007). An overview of dispositions in teacher education. In M. Diez & J. Rath (Eds.), *Dispositions in teacher education*, Information Age Publishing Inc. 3-29.
- Hassan M.,(2001). Transition to sustainability in the twenty first century: The contribution of science and technology- report of the world conference of scientific academies held in Tokyo, Japan, 15-18 May 2000, *International Journal of Sustainability in Higher education*, 2(1), 70-78.
- Helen K., & Frans M., (2014) Education for sustainable development (ESD): Exploring theoretical and practical challenges, *International Journal of Sustainability in Higher Education*, 15(2), 188-207
- Heselink, F., & Goldstein, W. (2000). The role of IUCN—The World Conservation Union—in shaping education for sustainability. In K. A. Wheeler & A. P. Bijur (Eds.), *Education for a sustainable future: A paradigm of hope for the 21st century* (pp. 123–142). New York, NY: Kluwer.
- Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, 6(1).
- Husserl, E. (1970). *The crisis of european sciences and transcendental phenomenology; an introduction to phenomenological philosophy*. Evanston: Northwestern University Press.
- Interstate New Teacher Assessment and Support Consortium. (2002). Model standards for beginning teacher licensing and development. Washington, DC.

- Johnstone, C. J., Bottsford - Miller, Nicole A., & Thompson, S. J. (2006). *Using the think aloud method (cognitive labs) to evaluate test design for students with disabilities and english language learners. technical report 44* National Center on Educational Outcomes. University of Minnesota, 350 Elliott Hall, 75 East River Road, Minneapolis, MN 55455. Tel: 612-626-1530; Fax: 612-624-0879; e-mail: nceo@umn.edu; Web site: <http://education.umn.edu/NCEO>.
- Jöreskog, K. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, 34(2), 183.
- Kaiser, F. G., & Shimoda, T. A. (1999). Responsibility as a predictor of ecological behavior. *Journal of Environmental Psychology*, 19(3), 243-253.
- Katz, L. G., & Raths, J. D. (1985). Dispositions as goals for teacher education. *Teaching and Teacher Education*, 1(4), 301-307.
- Kelly, F. J., & Veldman, D. J. (1964). Delinquency and school dropout behavior as a function of impulsivity and nondominant values. *Journal of Abnormal and Social Psychology*, 69(2), 190.
- Kenny, D.A. & McCoach, D.B. (2003), Effect of the number of variables of measures of fit in structural equation modeling. *Structural Equation Modeling*, 10(3), 333-351.
- Kline, P. (2014). *An easy guide to factor analysis*. London Routledge.
- Kopnina, H., & Meijers, F. (2014). Education for sustainable development (ESD): Exploring theoretical and practical challenges. *International Journal of Sustainability in Higher Education*, 15(2), 188-207.
- Lester, S. (1999). An introduction to phenomenological research. Retrieved from http://www.psyking.net/HTMLobj3825/Introduction_to_Phenomenological_Research-Lester.pdf
- Lincoln, Y. S. & Guba, E.G. (1985). *Naturalistic inquiry*. Thousand Oaks, CA: Sage Publications
- Little, T. D., Preacher, K. J., Selig, J. P., & Card, N. A. (2007). New developments in latent variable panel analyses of longitudinal data. *International Journal of Behavioral Development*, 31(4), 357-365.
- Lozano, R., Lukman R. Lozano F., Huisingh, D., & Lambrechts, W. (2013). Declarations for sustainability in higher education: Becoming better leaders, through addressing the university system. *Journal of Cleaner Production*, 48, 10-19.

- Luyt, R. (2012). A framework for mixing methods in quantitative measurement development, validation, and revision: A case study. *Journal of Mixed Methods Research*, (4), 294-316.
- Merriam, S. B. (2002). *Qualitative research in practice: Examples for discussion and analysis*. San Francisco, CA: Jossey-Bass.
- McKeown, R., Hopkins, C. A., Rizi, R., & Chrystalbridge, M. (2002). *Education for sustainable development toolkit* Energy, Environment and Resources Center, University of Tennessee Knoxville.
- McKeown, R. & Hopkins, C. (2003). EE is not equal to ESD: Defusing the worry. *Environmental Education Research*, 9(1), 117- 128.
- McNeil, K., Newman, I., & Kelly, F.J. (1996). *Testing research hypotheses with the general linear model*. Carbondale, IL: Southern Illinois University Press.
- Morris, D. & Martin, S. (2011) Complexity, systems thinking, and practice. In A. Stibbe (Ed.), *The handbook of sustainability literacy: Skills for a changing world* (pp. 156-164).
- National Council for Accreditation of Teacher Education (NCATE). (2002). Professional standards for the accreditation of schools, colleges, and departments of education (2006 ed.). Washington, DC.
- National Research Council. (1997) Preparing for the 21st century: The environment and the human future. Washington, DC: The National Academies Press.
- Nespor, J. (1987). The role of beliefs in the practice of teaching. *Journal of Curriculum Studies*, 19(4), 317-328. Retrieved from <http://ezproxy.fiu.edu/login?url=http://search.proquest.com/docview/63280609?accountid=10901>
- Newman, I. (1988). There is no such thing as multivariate analysis: All analyses are univariate! *Mid-Western Educational Researcher* (Presidential Address).
- Newman, I., Fraas, J., & Norfolk, T. Binomial index of model fit: An elaboration. *Structural Equation Modeling: A Multidisciplinary Journal*, 2(2), 155.
- Newman, I., Lim, J., & Pineda, F. (2013). Content Validity Using a Mixed Methods Approach Its Application and Development Through the Use of a Table of Specifications Methodology. *Journal of Mixed Methods Research*, 7(3), 243-260.
- Newman, I., Mcneil, K., & Fraas, J. (2004). Two methods of estimating a study's replicability. *Mid-Western Educational Researcher*, 17(2), 36.

- Newman, I., Ridenour, C. S., Newman, C., Smith, S., & Brown, R. C. (2013). Detecting low incidents effects: The value of mixed methods research designs in low-N studies. *Mid-Western Educational Researcher*, (4), 31.
- Newman, I., Newman, C., Brown, R. & McNeeley, S. (2006). *Conceptual statistics for beginners* (3rd Ed.). Lanham, MD: University Press of America, Inc. 146-155
- Nolet, V. (2009). Preparing sustainability-literate teachers. *Teachers College Record*, 111(2), 409-442. Retrieved from <http://ezproxy.fiu.edu/login?url=http://search.proquest.com/docview/61898470?accountid=10901>
- Nunnally, J. C. (1967). *Psychometric Theory*. New York, NY: McGraw-Hill.
- Nunnally, J. C., Bernstein I. H., (1994). *Psychometric Theory*. New York, NY: McGraw-Hill.
- O'Farrell, E. M. (2010). The effects of participation of school children as mediators in contrast to non-mediators in a mentored mediation program as related to academic achievement, developmental disposition, and conflict orientation. (Ed.D., California State University, Fresno). *ProQuest Dissertations and Theses*, . (756466781).
- Orr, D. W. (1992). *Ecological literacy: Education and the transition to a postmodern world*. Albany NY: State University of New York Press.
- Orr, D. W. (2004). *Earth in mind, on education, environment, and the human prospect*. Washington D.C. Island Pr.
- Osman, A., Gutierrez, P. M., Barrios, F., Wong, J. L., Freedenthal, S., & Lozano, G. (2010). Development and initial psychometric properties of the university of texas at san antonio future disposition inventory. *Journal of Clinical Psychology*, 66(4), 410.
- Patton, M. Q., & Patton, M. Q. (2002). *Qualitative research and evaluation methods*. Thousand Oaks, CA: Sage.
- Polit, D.F., and Beck, C.T. (2006). The content validity index: are you sure you know what's being reported? Critique and recommendations. *Research in Nursing & Health*, 29,489-497.
- Plichta, S. B., & Kelvin, E. A. (2012). *Munro's statistical methods for health care research* (6th ed.). Philadelphia, PA: Wolters Kluwer Health/Lippincott Williams & Wilkins.

- Richardson, V. P. P.(2001). Teacher change. In V. Richardson (Ed.),. *Handbook of research on teaching*, American Educational Research Association 905-947.
- Richardson, V. (2003). Preservice teachers' beliefs. *Teacher beliefs and classroom performance: The impact of teacher education*, 6, 1-22.
- Rokeach, M. (1968). *Beliefs, attitudes, and values: A theory of organization and change* (Ed.1). San Francisco, CA: Jossey-Bass.
- Rosefsky, S. & Opfer, V. (2012). Learning 21st century skills requires 21st century teaching. *Phi Delta Kappan*, 94(2), 8-13.
- Rubin, H. J., & Rubin, I. (2012). *Qualitative interviewing: The art of hearing data*. Thousand Oaks, CA: Sage Publications.
- Saylan, C., & Blumstein, D. T. (2011). *The failure of environmental education (and how we can fix it)*. Berkeley, CA: University of California Press.
- Schussler, D. (2006) Defining dispositions: Wading through murky waters. *The Teacher Educator*, 41(4), 251.
- Schumacker, R. E., Lomax, R. G., & NetLibrary, I. (2010). *A beginner's guide to structural equation modeling*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research*, 99(6), 323-338.
- Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information*, 22(2), 63.
- Shephard, K. (2008). Higher education for sustainability: Seeking affective learning outcomes. *International Journal of Sustainability in Higher Education*, 9(1), 87-98.
- Shephard, K., Smith, N., Deaker, L., Harraway, J., Broughton-Ansin, F., & Mann, S. (2011). Comparing different measures of affective attributes relating to sustainability. *Environmental Education Research*, 17(3), 329. Retrieved from <http://ezproxy.fiu.edu/login?url=http://search.proquest.com/docview/871575662?accountid=10901>
- Sterling, S. (2001). *Sustainable education*. UK: Green Books.
- Stibbe, A. (2009). *The handbook of sustainability literacy, skills for a changing world*. Devon: Green Books Ltd.

- Shur, D. D. (2006). Exploratory or confirmatory factor analysis?. *Statistics and Data Analysis*, 31 Retrieved from <http://www2.sas.com/proceedings/sugi31/200131.pdf>
- Taylor, M. (2009). What is sensitivity analysis. Consortium YHE: University of York, 1-8.
- Taylor, R. L., & Wasicsko, M. M. (2000, November). The dispositions to teach. In annual meeting of the South Regional Association of Teacher Education, Lexington, KY.
- Tabachnick, B.G., & Fidell, L.S. (2007), *Using Multivariate Statistics* (5th ed.). New York, NY: Allyn and Bacon.
- Tabachnick, B. R., & Zeichner, K. M. (1984). The impact of the student teaching experience on the development of teacher perspectives. *Journal of teacher education*, 35(6), 28-36.
- Tashakkori, A., & Newman, I. (2010). Mixed methods: Integrating quantitative and qualitative approaches to research. In B. McGaw, E. Baker, & P. P. Peterson (Eds.), *International encyclopedia of education* (3rd ed., pp. 514-520). Oxford, UK:
- Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. Washington, DC: American Psychological Association.
- United Nations. (1992). Agenda 21, Chapter 36: Promoting education, public awareness and training: Report of the nations conference on environment and development. New York, NY: Author. Retrieved from <https://sustainabledevelopment.un.org/content/documents/Agenda21.pdf>
- United Nations, Department of Economic and Social Affairs, Commission on Sustainable Development. (1998). Measuring changes in consumption and production patterns. New York, NY: Author. Retrieved from <http://www.un.org/documents/ecosoc/cn17/1998/background/ecn171998-mccpp59text.htm>
- United Nations, Department of Economic and Social Affairs, Commission on Sustainable Development. (2002). Report of the world summit on sustainable development. New York, NY: Author Retrieved from http://www.un.org/jsummit/html/documents/summit_docs/131302_wssd_report_r_eissued.pdf
- United Nations Education, Scientific, and Cultural Organization (UNESCO). (2004). United Nations decade of education for sustainable development: Draft international implementation scheme. New York, NY: Author Retrieved from <http://unesdoc.unesco.org/images/0013/001399/139937e.pdf>

- United Nations Education, Scientific, and Cultural Organization (UNESCO). (2015). Education for sustainable development. Retrieved from <http://www.unesco.org/new/en/education/themes/leading-the-international-agenda/education-for-sustainable-development/education-for-sustainable-development/>
- Van der Akker, J. (2003) The science curriculum: Between ideals and outcomes. In B.J Fraser and K.G. Tobin (Eds.) *International handbook of science education, part one* (pp. 421-444). Kluwer Academic Publishers, Great Britain.
- Van Driel, J. H., Beijaard, D., & Verloop, N. (2001). Professional development and reform in science education: The role of teachers' practical knowledge. *Journal of Research in Science Teaching*, 38(2), 137.
- Villegas, A. M. (2007). Dispositions in teacher education: A look at social justice. *Journal of Teacher Education*, 58(5), 370.
- Wheaton, B., Muthen, B., Alwin, D. F., & Summers, G. F. Assessing reliability and stability in panel models. *Sociological Methodology*, 8, 84.
- Wilkerson, J. (2006). Measuring teacher dispositions: Standards-based or morality-based?. *The Teachers College Record*. Retrieved from <http://www.tcrecord.org/Content.asp?ContentID=12493>
- Zwickle, A., Koontz, T. M., Slagle, K. M., & Bruskotter, J. T. (2014). Assessing sustainability knowledge of a student population. *Int J of Sus in Higher Ed*, 15(4), 375-389. doi:10.1108/IJSHE-01-2013-0008

VITA

JENNIFER MORALES

- 1981 Born, Miami, Florida
- 1999-2006 B.S., Curriculum and Instruction
Florida International University
Miami, Florida
- 2006-2011 M.S., Curriculum and Instruction
Florida International University
Miami, Florida
- 2011 -2015 Doctoral Candidate
Florida International University
Miami, Florida
- Teaching Assistant
Florida International University
Miami, Florida

PUBLICATIONS AND PRESENTATIONS

- Mills (Morales), J., O'Brien, G., & Newman, I. (2013, February). "A systemic and holistic approach to improve ecological affinity and environmental awareness: Implications for educators." Paper presented at the Eastern Educational Research Association (EERA) Annual Conference in Sarasota, FL, February 20 - 23.
- Morales, J. (2015, April) "Using structural equation modeling to facilitate the validity and reliability estimations of an instrument to measure dispositions toward sustainable systems instruction: A GLM approach." Paper presented at the American Educational Research Association (AERA) Annual Conference, Chicago, Il., April 16 -20
- Morales, J. (2015, in press). Using structural equation modeling to facilitate the validity and reliability estimations of an instrument to measure dispositions towards sustainable systems instruction: A GLM approach. *General Linear Model Journal* Formerly titled *Multiple Linear Regression Viewpoints*
- Morales, J., Newman I., Newman D. (2013, June) "Suggestions for methodological considerations in research: A critical examination of environmental orientation in the educational setting." Poster Presentation at South Florida Education Research Conference(SFERC) Miami, Fl., June 1

- Morales, J., & O'Brien, G. E. (2014, October). "Phenomenological study of sustainability: NAAEE workshop participants' experiences and beliefs." Poster presented at the Annual NAAEE Conference, Ottawa, Canada, October 9-11
- Newman, D., Morales, J. (2014, February) "Mixed methods in the medical setting." Paper Presentation at the Eastern Educational Research Association (EERA) Annual Conference in Jacksonville FL, February 20-23
- Newman, D., Newman, I., Ridenour, C., & Morales, J. (2014). Value added methods: Moving from univariate to multivariate criteria. *Journal of Research in Education*, 24(1), 99-111.
- O'Brien, G. E., Sparrow, K., Morales, J., & Clayborn, J. (2014, October). "Preparing sustainability literate K-5 teachers: Utilizing active engagement experiences." Poster presented at the Annual NAAEE Conference, Ottawa, Canada, October 9-11.
- O'Brien, G., Sparrow, K., Morales, J., and Clayborn, J. (2015). Re-orientating a Science Methods Course to Prepare Sustainability Literate K-6 Preservice Teachers: A Mixed Methods Investigation, in Stratton, S., Hagevik, R., Feldman, A., and Bloom, M. (Eds.). *Educating Science Teachers for Sustainability*. ASTE Series in Science Education, Heidelberg: Springer, 205-234.
- Sparrow, K., O'Brien, G. E., Morales, J., & Clayborn, J. (2015, January). "The inclusion of sustainability education K-6 science teacher education: Exploring the importance of targeted and teacher mentored school-based experiences." Paper presented at the ASTE International Conference, Portland, Oregon, January 7-10.