DISSERTATION

Kenneth Darrell Royal

The Graduate School

University of Kentucky

2008

USING ITEM RESPONSE THEORY AND THE RASCH MEASUREMENT MODEL TO INVESTIGATE FACULTY PERCEPTIONS OF INSTRUCTIONAL GOALS

DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Education
at the University of Kentucky

By

Kenneth Darrell Royal

Lexington, Kentucky

Co-Directors: Dr. Kelly D. Bradley, Professor of Educational Policy Studies and Evaluation
and Dr. Jeffery P. Bieber, Professor of Educational Policy Studies and Evaluation
Lexington, Kentucky

UMI Number: 3315028

ProQuest LLC
789 E. Eisenhower Parkway
PO Box 1346
Ann Arbor, MI 48106-1346

USING ITEM RESPONSE THEORY AND THE RASCH MEASUREMENT MODEL TO INVESTIGATE FACULTY PERCEPTIONS OF INSTRUCTIONAL GOALS

By

Kenneth Darrell Royal

Co-Director of Dissertation

Co-Director of Dissertation

Director of Graduate Studies

4-30-2008

ABSTRACT OF DISSERTATION

Kenneth Darrell Royal

The Graduate School

University of Kentucky

2008

USING ITEM RESPONSE THEORY AND THE RASCH MEASUREMENT MODEL TO INVESTIGATE FACULTY PERCEPTIONS OF INSTRUCTIONAL GOALS

ABSTRACT OF DISSERTATION

A dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy in the
College of Education
at the University of Kentucky

By

Kenneth Darrell Royal

Lexington, Kentucky

Co-Directors: Dr. Kelly D. Bradley, Professor of Educational Policy Studies and Evaluation
and Dr. Jeffery P. Bieber, Professor of Educational Policy Studies and Evaluation
Lexington, Kentucky

ABSTRACT OF DISSERTATION

USING ITEM RESPONSE THEORY AND THE RASCH MEASUREMENT MODEL TO INVESTIGATE FACULTY PERCEPTIONS OF INSTRUCTIONAL GOALS

The purpose of the study was to provide a methodological framework for analyzing data collected via survey research techniques, especially within the realm of higher education. Further, this research sought to investigate faculty perceptions of instructional goals by academic discipline based on faculty responses to the 2001 Faculty Survey administered by UCLA's Higher Education Research Institute (HERI). This study discussed research based on Classical Test Theory (CTT) principles and revealed the inadequacies of its assumptions relating to reliable and valid measures. An argument for Item Response Theory (IRT), particularly the Rasch model, was made and supported by a discussion of how the Rasch model sufficiently meets many of the deficiencies of CTT. Data were analyzed and discussed within the framework of the one-parameter IRT, Rasch measurement model. Analyses involved testing for data-to-model fit and rating scale functioning, evaluation of potentially misfitting items, item mapping, and differential item functioning (DIF). DIF techniques generated hierarchies of academic disciplines for each of the 14 survey items. Results were extended back to higher education theory via direct comparison to three popular classification systems, including Biglan, Kolb, and Holland models/theories. Results suggest all three classification systems were useful in explaining faculty perceptions of instructional goals based on academic disciplinary affiliation. Results also suggest faculty from all academic disciplines were primarily concerned with the intellectual development of students. Sentiments regarding the other 13 instructional goals varied greatly by discipline.

KEYWORDS: Rasch Measurement, Survey Research, Higher Education, Faculty, Outcomes Assessment

Kenneth D. Royal

Student's Signature

4-30-2008

Date

This dissertation is dedicated to my parents, Kennith and Connie Royal. When I left home in January of 2001 to complete my bachelor's degree I had no idea where my life was headed. As each day passed and new clues to my future were revealed, they remained steadfast and completely supportive of all my goals and aspirations, no matter how unrealistic they may have seemed. Today, as I close another significant chapter of my life and welcome in a new one, it is with the strongest convictions that I say "thank you for everything; I couldn't have done it without you. I love you."

ACKNOWLEDGEMENTS

First and foremost, I would like to thank God. I am convinced that none of my successes would have been possible without Him. The comfort I take in knowing He is "pulling the strings" in my life has made all the difference for me.

This dissertation owes much of its content to Kelly Bradley, who first exposed me to the beauty of quantitative methods and research. Kelly has provided me with unimaginable opportunities for research and professional development, as well as support as both a mentor and friend. Her impact on my life has been immeasurable. Jeff Bieber, my co-chair, has kept me grounded since 2003 when he chaired my master's committee. His critical examination of my work has tested me to my fullest and pushed me to produce my best work possible. He has been nothing short of brilliant. The remainder of my committee, Richard Angelo, G. T. Lineberry, and Suzanne Segerstrom has each provided insights that challenged my thinking and substantially improved the finished product. I greatly appreciate their insights, encouragement and professionalism.

I received equally important assistance from family and friends. My family, Kennith and Connie Royal, Ramona and Billy Duty, and Kim and Dave Arnold, have provided immense and continuous emotional support throughout the years. Knowing my family believes in me and my abilities has fueled me with the determination and motivation necessary to succeed at any challenge. My family also taught me the value of hard work at a very early age, not only in principle but through example. Because of this lesson, I have never dismissed or lamented work, a quality which came in very handy in Graduate School.

I wish to thank my friends for not only collaborating on research with me and pushing me to become a better researcher, but also for distracting me and reminding me there is much more to life than work and academics. Specific to this dissertation, I would like to thank Dr. Jennifer Weber, Jessica Cunningham, Dr. Shannon Sampson, Alexandra Henchy, Jennifer Eli and Bill

Harris for their comments. Special thanks to the Davidson family, Rev. Ken Bolin, and the many friends and supporters that attended my dissertation defense.

Also, special thanks to Thomas Guskey and Eric Reed for allowing me the honor of teaching graduate statistics alongside them. Thanks to the entire EPE departmental faculty for providing me with opportunities to chase my dreams and for helping me grow in so many ways. I would also like to thank Dr. Connie Ray, Deb Moore, Pauline Chhooi and Lingling Ma for believing in me and offering me an opportunity to help lead UK's Office of Assessment.

Finally, I wish to thank the Office of Higher Education Research in UK's College of Education for providing funding that supported this study.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter One: Introduction

> "The lack of attention to measurement issues is one of the major deficiencies in the higher education research literature" (Smart, 2005, p. 470).

John Smart, editor of *Research in Higher Education*, reflected upon his long career in academe ranging from his experience as a doctoral student, a higher education scholar, and as an experienced editor for various scholarly publications in his "Perspectives of the Editor" article. In the article, Smart outlined what he believed to be attributes of exemplary manuscripts that employ quantitative analyses. He discussed the paramount importance of measurement in quality research and stated "Exemplary manuscripts… use measures that have established psychometric merit, and they provide evidence of the reliability and validity of those measures. Such attributes are rarely evident in the higher education research literature" (Smart, 2005, p. 470). He went on to posit that a number of higher education researchers possess strong statistical skills, but few are actually trained in measurement.

Hutchinson and Lovell (2004) offer support for Smart's comments in a meta-analysis of the methods employed in the three premiere higher education journals (*Journal of Higher Education*, *Research in Higher Education*, and the *Review of Higher Education)* "The methodologies showcased in the three journals… suggest that higher education researchers possess fairly strong methodological skills in statistical analyses, but somewhat limited training in measurement" (p. 398). The authors go on to address the types of analyses performed and provide counts and frequencies for the various techniques. Hutchinson and Lovell found nearly all quantitative analytical techniques incorporated a classical test theory (CTT) approach. This suggests a great deal of previous research may have ignored the principles of sound measurement and hastily analyzed data without great concern to measurement.

In response to both Smart and Hutchinson and Lovell's conclusions, a need surfaces to call attention to issues of measurement and expose deficiencies of training and/or practice in the current higher education arena. Conducting quantitative research without proper attention to measurement is problematic because measurement is a fundamental component of quality research. Measurement issues should be adequately addressed before any analyses are performed. Although the CTT approach has its strengths and purposes, an Item Response Theory (IRT) approach may be more appropriate for many quantitative studies, especially those that employ survey research techniques. This study exhibited the value of an IRT approach by offering a methodological tool and a more accurate depiction of the results that the data yielded. This was demonstrated by applying Rasch measurement techniques to the Higher Education Research Institute's (HERI) 2001 Faculty Survey dataset.

The HERI Faculty Survey provides a nationally representative sample and possesses a response rate of at least 50%. The survey uses many Likert-type response scales to collect data from faculty across the country, including fourteen items relating to faculty perceptions of instructional goals. One of the fundamental benefits of Rasch measurement is it overcomes the assumptions many researchers make regarding the supposed equal distance between response scale options. Erroneously, most survey researchers assume these scales are interval levels of measurement. In actuality, these scales are ordinal. That is to say when given the question "Indicate the importance to you of each of the following education goals for undergraduate students", the response options of "Essential", "Very Important", "Somewhat Important" and "Not Important" are not equidistant from another. Hypothetically, the psychometric proximity between the response options "Essential" and "Very Important" could be considerably closer to one another than the response options "Somewhat Important" and "Not Important". The proper way to treat such scales would require a calibration before any data were analyzed. The Rasch model sufficiently does this even with considerably smaller sample sizes than those typically needed for CTT models, hence making the Rasch model much less sample-dependent than CTT approaches. As a testament to the Rasch model's strength, Curtis & Keeves (1999), Peck (2001), Waugh (1999) and Wright and Masters (1981) concur the Rasch model is the only IRT model that adheres to the seven principles of true measurement (as stated below).

- Each item should function as intended;
- Each item can be positioned on a common scale;
- The scale should be an interval one;
- Each person can be located along the same common scale used for items;
- The responses should form a valid response pattern for each item;
- Estimates of precision must be available for all scale measures;
- Each item should retain its meaning and function across individuals and groups (Curtis & Keeves; Wright & Masters);

Using the HERI data and the IRT framework, faculty instructional goals were explored from the perspective of research university faculty perceptions of these goals. Because this study was largely methodological in nature, evaluation of the methods of previous research in the area of faculty instructional goals was necessary. It was not surprising to discover that the particular subset of literature on faculty instructional goals was consistent with the general higher education literature, overwhelmingly and erroneously relying on CTT principles and absent of quality measurement.

Despite all the work generated in recent years regarding faculty instructional goals, it is possible that the quality of some of higher education's quantitative research could be compromised due to erroneous methodological assumptions and techniques, largely rooted in

CTT principles. Although measurement theorists have been arguing for some time now that the solution to many of CTT's deficiencies can be alleviated by incorporating an IRT approach, particularly the Rasch measurement model (Andrich, 1978; Bond & Fox, 2001; Bradley & Sampson, 2005; Masters, 1982; Smith & Smith, 2004; Wright & Stone, 1979), the transfer and implementation of this knowledge to the higher education literature has yet to occur, at least in the mainstream higher education literature. Acknowledging Hutchinson and Lovell's (2004) findings and implementing Smart's (2005) suggestions could yield several important theoretical and methodological questions for researchers. Some of these questions include: what makes the IRT approach, specifically the Rasch model, a more valid and reliable approach to quantitative techniques? Why is this model a more powerful and precise tool for survey researchers? Bradley and Sampson (2005) have eloquently summarized the advantages of Rasch measurement stating:

> Whereas the classical model produces a descriptive summary based on statistical analysis, it is limited, if not absent, in the measurement capacity. In contrast, Rasch measurement tackles many of the deficiencies of the classical test model in that it has the capacity to incorporate missing data, produces validity and reliability measures for person measures and item calibrations, measures persons and items on the same metric, and is not dependent on the particulars of the sample. Applications of the Rasch model allow the researcher to identify where possible misinterpretation occurs and which items do not appear to measure the construct of interest, while producing information about the structure of the rating scale and the degree to which each item contributes to the construct. Thus, it provides a mathematically sound alternative to traditional approaches to survey data analysis (p. 13).

Another advantage of the Rasch model is it requires researchers to ensure model-data fit and rating scale functioning before any analyses occur. Too often data analysis is done using a "plug-and-chug" approach as researchers assume models work without paying attention to the assumptions of each model. This study answered the guiding questions with careful consideration given to the methodological approaches employed, resulting in valid and more meaningful results.

Purpose

The purpose of this study was to provide a methodological framework for analyzing data collected via survey research techniques, especially within the realm of higher education. This study discussed research based on CTT principles and revealed the inadequacies of its assumptions relating to reliable and valid measures. An argument for IRT, particularly the Rasch model, was supported by a discussion of how the Rasch model sufficiently meets many of the deficiencies of CTT. A test for model fit and rating scale functioning was then presented. Once proper calibrations were made and sufficient evidence was given for the fit of the data to the model, further analyses were performed and a presentation and discussion of its findings followed. This demonstration was performed by investigating faculty perceptions of instructional goals provided by data from the 2001 HERI Faculty Survey dataset.

Research Questions

This study was guided by four research questions. The first two questions pertained to issues specific to the Rasch model, particularly identifying the "fit" and "functioning" of the data to the model. These questions are necessary each time one uses the Rasch model, as it illustrates to what extent the model is an appropriate technique for data analysis. The subsequent questions were answered through the results of the actual Rasch model application. The Rasch model allows for the construction of item maps for both persons and items. In this study, the power of item maps was demonstrated by mapping academic disciplines and comparing their hierarchy to popular classifications systems previously found in higher education research, such as the models of Anthony Biglan (1973), David Kolb (1980), and John Holland (1966). The Rasch model also allowed analyses for Differential Item Functioning (DIF). The combination of the item and person maps and DIF supported the identification of themes amongst the academic disciplines; thus, allowing for concurrent discussions regarding classification systems in today's higher education environments.

The research questions guiding this study were:

1. How well do items from the HERI Faculty Survey measure faculty instructional goals among university faculty?
2. Do relevant items on the HERI Faculty Survey fit the expectations of the one-parameter IRT (Rasch) model by forming a unidimensional construct?
3. How does a hierarchy of academic disciplines compare to previously established higher education classification systems?
4. In what ways does presumed paradigmatic consensus influence faculty instructional goals?

Assumptions

Assumptions critical to the integrity of the study were in the form of two frameworks: 1) those specific to data collection; and 2) those specific to analytical procedures. First, it was assumed the population of study (regular series, tenure-track faculty from research universities who responded to the 2001 HERI Faculty Survey) was representative of similar faculty throughout the United States. It was further assumed that faculty completed the HERI questionnaire with honesty, leading to valid results. Second, it was assumed the theoretical construct of interest was unidimensional.

Limitations

Perhaps the most noticeable methodological limitation of this study was a byproduct of employing a measurement technique for data analysis. Although this study will be of value as it contributed to a sparse literature base, this study had limited comparability with other studies due to the uniqueness of this data analysis method. Despite the inability to compare methodologies with previous studies, it is asserted that the results and findings could be compared to previous research.

The issue of accessibility was another limitation. Although Rasch measurement is a valid and reliable science, and arguably one of the best methods for analyses of this kind, it is relatively new in comparison to traditional statistics. Although Rasch measurement is emerging and taking hold in many disciplines, especially psychology, medicine and education, measurement software is not as easily accessible (compared to some statistical programs such as SPSS or SAS) for those wishing to analyze data of their own or potentially replicate this study. Further compounding the issue of accessibility was that Rasch measurement is seldom used in higher education research. As a result of its infrequent use and visibility, this may lead to potentially unjust criticisms when reviewed by some peers in the higher education field.

Basic Terms and Definitions

Academic Discipline – A field of study that is taught and/or researched at the college or university level. Examples include: chemistry, math, sociology, education, art, law, etc.

Classical Test Theory – "CTT is based upon conceptual models in which relations among constructs are theorized from theories ground in previously published literature. Once a conceptual model of the relationships among different variables has been established, a measurement model can be constructed" (Embretson and Hershberger, 1999, p. 5). Generally, CTT is used to examine a group of individuals' responses to a test. As suggested above, a mathematical model is then applied to fit the data. "[Typically], CTT collectively considers a pool of examinees and empirically examines their success rate on an item" (Fan, 1998, p. 358).

Faculty – For this study, faculty was used in both a general and specific context. In general, faculty referred to the academic staff at a college or university whose primary role is to teach and/or conduct research. When referring to the population of this study, faculty referred to regular series, tenure track professors (of all ranks) at research universities.

Item Response Theory - "A relatively recent development in psychometric theory that overcomes deficiencies of the classical test theory with a family of models to assess model-data fit and evaluate educational and psychological tests" (Bond and Fox, 2001, p. 231).

Rasch Measurement – A measurement model derived from item response theory that converts scores (or responses) to a mathematical logarithm and measures the probability of success (or correct responses) between persons and items (in tests, surveys, etc.).

Contributions

With regard to contributions, perhaps methods are at the forefront. Related studies have largely been based on classical test theory and utilized basic descriptive statistics, regression and factor analyses. Arguably, item response theory provides a better and more comprehensive quantitative technique for studies of this nature. This is due to the one-parameter IRT model's strict adherence to the established criteria for sound measurement (see Wright and Master, 1981; and Curtis and Keeves, 1999). Further, IRT includes and investigates the survey instrument through the "validation" process, as opposed to statistical techniques that simply analyze data. Utilizing IRT will impact previous and future studies, as the possibilities for re-analyzing data from previous research will surface and a framework for future studies will be outlined.

This study also evaluated previously established higher education classification systems as the results of this study were compared to previous models. This will allow for concurrent discussions regarding classification systems in today's higher education environments.

Summary

This chapter presented a brief overview of issues pertaining to measurement in higher education research. The purpose of the study, as well as the study's assumptions, limitations and contributions were presented. In the remaining two chapters, a critical examination of relevant literature will be provided, followed by a discussion of methods to be used in the current study. Chapter Two will discuss literature central to the present study. This will include literature from the fields of measurement, educational research and higher education. Chapter Three will outline the specifics of the current study, including detailed descriptions of procedures, sampling and data analyses techniques.

Chapter Two: Review of the Literature

This review of selected literature begins by providing background on the two approaches to measurement: Classical Test Theory (CTT) and Item Response Theory (IRT). First, CTT is examined and its deficiencies revealed, and then an argument for IRT is made. This is followed by a discussion of the Rasch measurement model, which includes an outline of the model's strengths, uses with polytomous data, and a discussion of the mechanics of the model.

The next topic of discussion is measurement in higher education research. An argument is made regarding the lack of quality measurement in general quantitative higher education research. This argument includes an examination of graduate training in measurement, an examination of the methodologies employed in the three premiere higher education journals, as well as an examination of higher education studies which employ the Rasch model. The methodologies employed in the faculty instructional goals literature are then examined in regards to this specific study. This is followed by an argument discussing how previous quantitative research in this arena is biased in its analytical techniques and how that may affect both the validity and reliability of each study's findings. An appeal for a different analytical technique, particularly Rasch measurement, is then made.

Finally, previous research on faculty instructional goals is examined. This includes a discussion of Angelo and Cross's seminal work on the matter, which concluded academic disciplines are of paramount importance when attempting to understand the instructional goals of faculty. Three higher education classification systems/models are then introduced, including the Biglan model, Kolb model and Holland's theory. General findings from the faculty instructional goals literature are presented followed by findings from studies which incorporated the aforementioned models/systems.

Measurement Approach

Assumptions of Classical Test Theory.

Previous research on faculty teaching objectives has relied largely on surveys. Although most researchers pilot test their instruments, few take the time to question many of the assumptions of survey research. Wright and Stone (1979), Bond and Fox (2001) and Bradley and Sampson (2005) note that many survey researchers make numerous, flawed assumptions. Similar to the rulers and instruments used in the "hard" sciences, the human sciences also need rulers (or scales) with equal distance between units of measurement. First, many researchers assume there is equal distance between response scale options. Erroneously, researchers assume these scales are interval levels of measurement; in actuality, these scales are ordinal. Take a typical five-point Likert-type scale for instance. Many assume the distance between response options 1 and 2 is the same as the distance between response options 2 and 3. The point is further illustrated when assigning meaning to the scale, such as "Strongly Agree", "Disagree", "Neither…", etc.. Although the equidistant assumption between Likert-type response scales may appear logical in nature, it is not theoretically sound. Hays (1988) writes:

> The problem of measurement, and especially of attaining interval scales, is an extremely serious one for the social and behavioral sciences. It is unfortunate that in their search for quantitative methods, researchers sometimes overlook the question of level of measurement and tend to read quite unjustified meanings into their results. …However, the core problem of level of measurement lies outside the province of mathematics and statistics (p. 71).

Another issue is the assumption that each survey item is of equal importance to the construct being measured. Bond and Fox (2001) and Sampson and Bradley (2003) point out there are a number of assumptions with this logic as well. For instance, all respondents must interpret the survey directions correctly; all items must be worded in a manner that all respondents would interpret the item the same way; and the items accurately fit the construct of measure. The researchers warn these assumptions are unstable and problematic in survey research. Alarmingly, these assumptions may have a significant impact upon validity and reliability measures as well.

Applications of Classical Test Theory.

Classical Test Theory (CTT) was introduced in 1904 by Charles Spearman. "CTT is based upon conceptual models in which relations among constructs are theorized… from theories ground in previously published literature. Once a conceptual model of the relationships among different variables has been established, a measurement model can be constructed" (Embretson and Hershberger, 1999, p. 5). Generally, CTT is used to examine a group of individuals' responses to a test. As suggested above, a mathematical model is then applied to fit the data.

Some sixty years later, however, scholars began to re-examine CTT and its assumptions and began to develop new models with stronger theoretical underpinnings. "CTT does not invoke a complex theoretical model to relate an examinee's ability to succeed on a particular item. Instead CTT collectively considers a pool of examinees and empirically examines their success rate on an item" (Fan, 1998, p. 358). The more theoretically sound solution came to be known as Item Response Theory (IRT). Bond and Fox (2001) define IRT as "a relatively recent development in psychometric theory that overcomes deficiencies of the classical test theory with a family of models to assess model-data fit and evaluate educational and psychological tests" (p. 231).

CTT is often criticized for several important reasons: First, all measurement units are considered equivalent (Becker, 2001). Second, error across measurement units is independent and uncorrelated (Becker). Other cited disadvantages of CTT include the argument that it is sample dependent and requires larger samples and/or test items (Bond & Fox, 2001; Bunderson, 2000; Hambleton, Swaminathan, & Rogers, 1991), as well as its use of test-retest reliability. Richter, Werne, Heerlein, Kraus, and Sauer (1998) argue test-retest reliability is problematic due to timing issues, meaning there is too much time between initial and follow-up administration which might lead to an underestimation of measures. Likewise, too little time between initial and follow-up administration might lead to an overestimation of measures.

Argument for Item Response Theory.

The IRT family of models attempts to resolve the deficiencies of CTT largely by providing a theoretical justification that CTT generally lacks. IRT models use item characteristic curves (ICC) which display the relationship(s) between a trait and a response. In other words, IRT models can predict the likelihood of one's response and the extent to which respondents possess the single trait being measured. Also, because CTT models are sample dependent, they have unconditional standard errors of measurement. With IRT models, standard error is calculated for both persons and items. As previously mentioned, CTT models depend largely on sample size; IRT models (particularly, one-parameter models) are not as dependent, which can be a strong asset to many researchers working with limited resources and datasets (Embretson and Hershberger, 1999).

There are a number of differences between IRT models and CTT. According to Fan (1998), CTT focuses on test level information whereas IRT focuses on item-level information. In other words, IRT focuses on the interactions between individual persons and items, as the model suggests, each affects the other. Fan suggests IRT models assume a single trait is responsible for the subject's response to a particular item. Another major difference is CTT assumes test-takers have both observed and true scores, where the observed score is an estimate of the true score plus or minus measurement error (Crocker & Algina, 1986; Hambleton & Swaminathan, 1985). IRT, on the other hand, assumes the person's ability is independent of the content of a test, and the relationship between the probability of choosing the correct answer and the ability of the person can be modeled differently depending on the content of the test (Hambleton, et al., 1991). This explains why IRT models generally assume unidimensionality, or the notion that test items measure a single trait.

Overview of the Rasch Model.

Similar to the rulers and instruments used in the "hard" sciences, the human sciences also need rulers (or scales) with equal distance between units of measurement. Edward Thorndike introduced this notion in 1926, but it was not until the 1960s that scholars such as Loevinger, Gulliksen, Angoff, and Michell furthered Thorndike's initial ideas. In the 1960s, Georg Rasch created a logistic model that appeared to help meet the need for precision in measurement. This work led to the creation of a family of IRT models, including the Rasch model (Bond and Fox, 2001). The Rasch model allows constructs to be measured as if one were using a ruler to measure them. Essentially, the Rasch model manipulates data to create a ruler. Bond and Fox say the Rasch model "[helps] construct a measure of a construct, and then [one can] interpret each person estimate as a measure of the person's revelation of the latent trait as indicated on the assessment instrument" (p. 73). Bond and Fox go on to say the Rasch model "provides us with useful approximations of measures that help us understand the reason why people and items behave in a particular way. These approximations help us to solve problems that cannot be solved currently with any other model" (p. 8). Further, "Unlike other probabilistic measurement models, the Rasch model is the only one that provides the necessary objectivity for the construction of a scale that is separable from the distribution of the attitude in the persons it measures." (p. 7).

There are a number of IRT models, some of which include one-, two-, and three-parameter models. The one-parameter model focuses on difficulty. The two-parameter model focuses on item difficulty and item discrimination. The three-parameter model takes into account both difficulty and discrimination, but also controls for guessing. The Rasch model (Rasch, 1960) is another member of the IRT family. The Rasch model is the mathematical equivalent of the one-parameter IRT model, but differs conceptually. Originally, the Rasch model was designed for dichotomous data, but has since been extended to polytomous (formerly called polychromous) and Likert-type data (Andrich, 1978). Two particular examples include the Rating Scale Model (Andrich) and the Partial Credit Model (Masters, 1982), which will be discussed later in this section.

Unlike IRT and CTT models where the model is designed to fit the data, the Rasch model works if and only if the data fit the model. This means the model is fixed and therefore the data must adhere to the model's prescription. With the Rasch model, Wright and Stone (1979) suggest it is the researcher's responsibility to control for discrimination and guessing. If the data do not fit the model, the Rasch approach is of little utility. However, if the data fit the model, Rasch measurement provides a great means for interpreting and understanding the relationships between individual responses and selected items.

According to Andrich (1978), the Rasch model was originally intended to deal with only dichotomous data (e.g. correct versus incorrect responses). This changed as scholars began expanding the principles of Rasch measurement to apply to polytomous data as well, particularly data from Likert-type scales. Historically speaking, Andrich (1999) warned that the general use of

Likert-type scales may be problematic because the scales are theoretically weak. He argues researchers assume equal distance between responses (e.g., Strongly Agree, Agree, Disagree, and Strongly Disagree). Respondents, on the other hand, are likely to interpret the rating scales differently. Because of this, traditional statistics lack a great deal of validity and reliability.

Andrich (1999) believes this rating scale problem can be corrected by using the Rasch model. He argues the Rasch approach takes the distances between responses (also called thresholds) into consideration and holds them equidistant for all persons and items, hence providing more valid measures than traditional statistics. Bond and Fox (2001) agree as they say "the Rasch model treatment of Likert scale data is intuitively more satisfactory and mathematically more justifiable than the traditional 'allocate 1 2 3 4 5 and add them up' approach [of traditional statistics]" (p. 71). This is largely a result of treating data as interval, as opposed to ordinal. Arguably, this method will provide more stable results than traditional statistics while at the same time maintaining high measures of validity and reliability.

Rating Scale Model.

With regard to polytomous data, there are two predominant models: the Rating Scale Model and the Partial Credit Model. Bond and Fox (2001) define the Rating Scale Model as "a version of the Rasch model routinely used for the sort of polytomous data generated by Likert scales" (p. 233). They go on to say "[the rating scale model] requires that every item in a test have the same number of response options, and applies the one set of threshold values to all items on the test" (p. 233). Bond and Fox define the Partial Credit Model as "a Rasch model for polytomous data… which allows the number of ordered item categories and/or their threshold to vary from item to item" (p. 232). To summarize the major differences and state in another way, Wright (1999) says "the rating scale model specifies that a set of items share the same rating scale structure… The partial credit model specifies that each item has its own rating scale structure" (p. 641). Linacre (2005) argues there is little difference between the two models, as each have different formulas but generally arrive at very similar results. However, despite the similarities between the two models, this research will focus on the Rating Scale Model as all items on the survey will be pulled from the same section which utilizes the same response scale.

According to Linacre (2006), the standard Rasch model equation for dichotomous data is:

$$\log\left(P_{ni} / (1 - P_{ni})\right) = B_n - D_i$$

where, P_{ni} is the probability that person $_n$ will succeed on item $_i$, where person $_n$ has ability B_n and item $_i$ has difficulty D_i.

According to Masters (1982), the Rating Scale Model equation is:

$$\log\left(P_{nik} / P_{ni(k-1)}\right) = B_n - D_i - F_{gk}$$

where, P_{nik} is the probability of observing category k for person n encountering item i.

$P_{ni(k-1)}$ is the probability of observing category $k-1$ and F_{gk} is the difficulty of being observed in category k relative to category $k-1$ for an item in group g.

Other Issues Related to Rasch Measurement.

It is important to note that the concepts of Rasch measurement are quite different from traditional statistics. Linacre (2006) reports that statisticians may find Rasch measurement difficult to interpret as the methodologies have opposite positions. With statistics, the belief is data points are key and it is the researcher's responsibility to find models to explain them. With Rasch measurement, however, "the latent variable is the truth, and when that latent variable is expressed in linear terms, it is the Rasch model that is necessary and sufficient to describe it" (Linacre, 2006, 12). Linacre further asserts that data points that do not fit the model provide a "distorted picture of the latent variable" (p. 12).

Exactly how does the Rasch model work? Via the mathematical model shown above, the Rasch model converts a raw score to its natural logarithm. The conversion transforms the measure from an ordinal scale to an interval scale, which is key in addressing the issue of equidistant scales. A log odds scale prevents the scale from being biased towards scores in the middle and from persons responding at the extremes (Bond and Fox, 2001).

Bond and Fox (2001) give an example of how to convert a raw score to its natural logarithm. They suggest considering a raw score of 64%. The odds ratio then becomes 64/36. By using a calculator, the logarithm can be determined by entering 64/36 and pushing the log function. The result will be +0.58. By plotting individual logarithms along an interval scale, one can infer much more meaningful information than descriptive statistics alone can provide, particularly questions of 'how much?'

Why is the Rasch model a good measurement tool and why is it appropriate for many uses? Wright and Masters (1981) suggest there are seven criteria of true measurement. Curtis and Keeves (1999) outline the following criteria of true measurement:

- Each item should function as intended;
- Each item can be positioned on a common scale;
- The scale should be an interval one;
- Each person can be located along the same common scale used for items;
- The responses should form a valid response pattern for each item;
- Estimates of precision must be available for all scale measures;
- Each item should retain its meaning and function across individuals and groups;

Peck (2001), Curtis and Keeves (1999), and Waugh (1999) believe the Rasch model is the only IRT model that adheres to the above criteria, hence making it an excellent technique for fundamental measurement. Curtis and Keeves argue it is difficult to determine if all items contribute to a common scale with other forms of measurement, as some methods intending to ensure equal contribution lack sensitivity. They go on to say the seven criteria rule has not always been strictly enforced. Given the outlined measurement criteria and Rasch measurement's adherence to the criteria, many consider this infrequently used technique of Rasch analysis in the human sciences an excellent method for calibrating survey instruments and scales, and analyzing survey data.

Measurement in Higher Education Research

Despite the praise bestowed by many measurement theorists on the Rasch measurement model, the dissemination of this powerful technique to other academic fields has been relatively slow. Historically, educational psychology has been at the forefront for the use of Rasch measurement, as the theory originated from psychometrics. Increasingly, the use of the Rasch model is becoming more and more popular in health-related disciplines, market research and education. The question remains to what extent is Rasch measurement used in the higher education research arena. As discussed in the introduction of this study, the majority of quantitative research in the higher education arena lacks sound measurement. Why is that? Exactly how much is the field lacking as it relates to quality measurement? What can be done to alleviate this problem? These questions and more will be explored in the following section.

Measurement and Graduate Training.

As discussed in the introduction of this study, the majority of quantitative research in the higher education arena lacks sound measurement. Interestingly, however, there is an abundance of researchers skilled in statistical techniques (Smart, 2005; Hutchinson and Lovell, 2004). Hutchinson and Lovell (2004), Lovell and Hutchinson (2003), Lovell, Hutchinson and Fairweather (1999), and Aiken, West, Sechrest and Reno (1990) argue the problem with measurement has largely to do with many higher education graduate programs' exclusion of measurement courses from the curriculum. Hutchinson and Lovell state:

> In the field of higher education, the inattention to measurement likely reflects a lack of appropriate measurement training as suggested by a survey of research requirements among higher education doctoral programs conducted by Lovell et al. (1999) and Lovell and Hutchinson (2003). Of the higher education programs responding to the survey, few required measurement courses, and most tended to require only introductory level, statistically focused courses (p. 398).

The authors go on to conclude a persistent link exists between the attention measurement issues are given in doctoral training programs and that of measurement issues discussed in the premiere higher education journals. Hutchinson and Lovell say "the lack of awareness about measurement issues in the three journals reviewed in the current study seems to mirror the general inattention to measurement in many doctoral training programs" (p. 398).

Methodologies Used in the Higher Education literature.

Inspired by Hutchinson and Lovell's analysis of higher education journals, it is useful to conduct a meta-analysis to determine the frequency with which studies in the top higher education journals incorporated either a CTT or an IRT approach. Understanding the frequency of these approaches would allow one to more closely examine the quality of measurement taking place in higher education research.

Similar to Hutchinson and Lovell's 2004 study, the meta-analysis was begun by choosing the three journals considered to be the most prestigious in higher education; the *Journal of Higher Education*, the *Review of Higher Education* and *Research in Higher Education*. A timeframe of five years was arbitrarily chosen, and articles which spanned from 2003 to the present (summer of 2007) were analyzed. Each article was examined in detail and the analysis techniques employed were cited, as well as relevant information regarding the authors and the journal. Once the lists were generated, a code was provided for each technique according to whether it falls under the criteria of a CTT or an IRT approach. Counts and frequencies were then generated. The results were astounding. See Table 2.1 below.

Table 2.1

Frequency of CTT and IRT Applications in Higher Education's Top Journals

	Journal of Higher Education		*Review of Higher Education*		*Research in Higher Education*	
	Count	Percent	Count	Percent	Count	Percent
Classical Test Theory approach	61	96.8	41	97.6	149	97.4
Item Response Theory approach	2	3.2	1	2.4	4	2.6

Based on the results of this meta-analysis, only two to four percent of the quantitative research published in the past five years in these journals incorporated a methodological approach based on some form of item response theory.

Taking this meta-analysis a step further, the number of instances in which Rasch measurement was employed in published higher education literature was investigated. Performing a search in multiple databases spanning approximately 4,700 academic journals, conference papers, etc., I entered the terms "higher education", the connector "AND", and "Rasch measurement" in open search fields with no limitations. Results yielded only 21 records. Revising the terms to produce maximum hits, the phrases "higher education" AND "Rasch" AND "measurement" were entered into the search. Only 67 hits were recorded. Of those 67 articles, the vast majority were published in educational measurement journals. Exclusively searching the three premiere higher education journals, the word "Rasch" was entered to detect the most hits possible. Results revealed a total of three articles published in 1993, 1994, and 2000, respectively. Collectively, the results of my meta-analysis suggest there is little doubt there is a significant lack of research rooted in measurement theory in the higher education literature. Now, the focus will shift to the faculty instructional goals literature where the methodologies employed in the literature relevant to this study were evaluated.

Methodologies Used in the Faculty Instructional Goals literature.

Much like the general higher education literature, the use of IRT techniques in the faculty teaching goals literature is incredibly sparse. The table below highlighs key studies and their respective methodologies.

Table 2.2

Methodologies Employed in Relevant Studies Regarding Faculty Teaching Goals

Author(s)	Descriptive Statistics	t-tests	ANOVA/ MANOVA	Correlations	Regression	Cluster/Discriminant Analyses	Path Analyses	Factor Analyses
Angelo and Cross (1993)	X							
Barnes et al (1998)			X					
Barnes et al. (2001)[2]					X			
Blank (1976)					X		X	
Blank (1978)					X		X	
Braxton et al (1998)								
Brown (2003)		X	X					
Colbeck et al (2002)								X
Fox (1997)		X	X	X				
Gaff and Wilson (1971)[1]						X		
Hardy (2002)			X					
Hativa (1997)			X					
Hativa and Birenbaum (2000)[2]								X
Johnson (1997)			X		X			X
Kaya et al. (2005)				X				
NCPI (1999)	X							
Schwarze (1996)		X	X					
Shavelson et al. (2005)				X				
Smart (1982)			X					X
Smeby (1996)	X			X				
Stark and Morstain (1978)[1]						X		X
Swenson (1995)			X	X				
Trice and Dey (1997)					X			

[1] Article appeared in the *Journal of Higher Education*
[2] Article appeared in *Research in Higher Education*

As evidenced by the above table, none of the studies employed an IRT approach. In a paper entitled "Measurement for Social Science and Education: The History of Social Science Measurement", Benjamin Wright reminds us what early measurement pioneer Edward Thorndike discovered over a century ago… raw scores are not measures (Wright, 1997). Researchers should not take abstract concepts and treat them as concrete. Wright goes on to discuss raw score bias and how "The bias is just as severe for partial credits, rating scales and… the infamous Likert Scale, the misuse of which pushed Thurstone's seminal 1920's work on how to transform concrete raw scores into abstract linear measures out of use" (Wright, Psychometrics ¶6). Wright states "Any statistical method like linear regression, analysis of variance, generalizability, or factor analysis that uses raw scores or Likert scales as though they were linear measures will have its output hopelessly distorted by this bias" (Wright, Psychometrics ¶8). Wright says:

> Many social scientists still believe that misusing raw scores as measures does no harm. They are unaware of the consequences for their work of the raw score bias against extreme scores. Some believe that they can construct measures by decomposing raw score matrices with some kind of factor analysis. There is a similarity between measurement construction and factor analysis in the way that they expose multi-dimensionality (Smith, 1996). But factor analysis does not construct measures (Wright, 1996). All results from raw score analyses are spoiled by their non-linearity, their extreme score bias and their sample dependence (Wright, 1997).

Nearly every study highlighted in the above table erroneously treats raw scores as measures. This can have significant implications upon the validity of the results of these studies. In an attempt to steer researchers away from making this mistake Wright suggests researchers heed to the following measurement law:

> *Before applying linear statistical methods to concrete raw data, one must first use a measurement model to construct, from the observed raw data, abstract sample and test free linear measures* (Wright & Linacre, 1997; Linacre & Wright, 1997).

Wright asserts that this suggested law for model-controlled linearization has two benefits. The first benefit pertains to statistical validity, as each measure and calibration possesses a realistic estimate of precision. Second, when the measures are ready to be plotted and linear statistics applied, researchers now have linear measures with which they know their numerical precision and validity.

It is clear based on the evidence above that the results of the faculty instructional goals literature were clearly biased with regard to validity measures. As Hutchinson and Lovell (2004) remind us "The lack of attention to measurement quality is disturbing given that even the most sophisticated statistical technique provides meaningless results if not performed on valid and reliable data" (p. 397). Although there is no guarantee Rasch measurement will provide results different from that of previous research, what is important methodologically is that analyses are conducted in a careful, theoretically sound manner which further minimizes assumptions. With regard to content, if results vary considerably from previous research, then it would serve useful to revisit existing studies and reanalyze data with this measurement technique. The following section will examine results from previous research.

Previous Research on Faculty Instructional Goals

The seminal work on faculty teaching objectives was published in 1993 by Angelo and Cross. The authors concluded from the study, "what you teach has a good deal to do with how you teach—or at least what your teaching priorities are and how you perceive your primary role as a teacher" (p. 369). The authors identified academic discipline as the main factor in explaining differences amongst college faculty. They state:

> Faculty teaching priorities are related more to academic discipline than to any other factor. Teachers of a given discipline—whether male or female, full-time or part-time, experienced or inexperienced, teaching in a public community college or a private four-year college—share a value system with respect to teaching goals that is distinctly discipline-related and significantly different from that of colleagues in different disciplines (p. 366).

Numerous studies have corroborated the significant influence of academic disciplines when understanding faculty attitudes and behaviors (see Alpert, 1985; Becher, 1987; Clark, 1980; Ladd and Lipsett, 1975; Lee, 2004; Smart, Feldman and Ethington, 2000). A number of scholars have attempted to use various classification systems and/or models to serve as a framework for understanding the similarities and differences. Some of the more popular classification systems/models include Biglan and Kolb's models and Holland's theory. Anthony Biglan's model classifies disciplines according to 'hardness' ("soft" versus "hard" sciences), whether the field is pure versus applied in nature, and whether it pertains to "life" versus "non-life" subjects (Biglan, 1973a; 1973b). David Kolb's research on learning styles and experiential learning (1980) added to Biglan's model by including two additional dimensions: "active" versus "reflective", and "abstract" versus "concrete".

Another classification system is Holland's theory of person and environment fit (Holland, 1966; Smart, Feldman, and Ethington, 2000). This theory was borrowed from the psychology literature and essentially classifies both person and academic discipline according to six measures, the RIASEC classification system, otherwise known as "Holland types". The types include: Realistic, Artistic, Investigative, Social, Enterprising and Conventional. Holland's theory suggests if an individual and environment share the same RAISEC code, the individual will likely persist and find satisfaction within that environment. The converse is expected when a lack of congruence exists between the person and the environment.

General Findings Relating to Faculty Instructional Goals.

Donald (1990), Fox (1997), Franklin and Theall (1992), Neumann, Parry, and Becher (2002), and Swenson (1997) all found faculty to have different instructional goals depending on their disciplinary affiliation. The one exception to this general trend is that regardless of disciplinary affiliation, most research has suggested faculty are primarily concerned with the intellectual growth of students (see Jervis and Congdon, 1958; Lawrence, Hart, Mackie, Muniz, & Dickmann, 1990; Liebert and Bayer, 1975; Platt, Parsons, & Kirshstein, 1976; Royal, Eli, & Bradley, 2005; and Wilson, Gaff, Dienst, Wood, & Bavry, 1975). Lawrence et al. found this primary focus on intellectual growth was consistent across all college types (i.e., community colleges, four-year colleges, universities, etc.).

Liebert and Bayer (1975) found goals pertaining to students' moral and personal development were generally considered less important when compared to the intellectual growth of students by faculty at four-year colleges and universities. Jervis and Congdon (1958) asked faculty to rank four major outcomes in order of importance and found faculty ranked "intellectual growth" first, "self-fulfillment" second, "self-understanding" third and "social growth" last. Royal, Eli, and Bradley (2005) found community college faculty as a whole were overwhelmingly concerned with the intellectual growth of students, followed moderately by emotional, social, and cultural growth outcomes.

Stark and Morstain (1978) found natural science and faculty from professional fields were more concerned with "preparation for life and work" than faculty from the social sciences and humanities. Conversely, however, social science and humanities faculty tend to be more concerned with the "pursuit of ideas" than faculty from the natural sciences and professional fields. In extant research, Braxton and Nordvall (1985), Gaff and Wilson (1971), Lattuca and Stark (1994), and Smart and Ethington (1995) found faculty in natural and physical sciences were more likely to require memorization and application, whereas faculty in the social and behavioral sciences and humanities were more likely to address critical thinking. Also taking disciplines into account, Royal et al. (2005) found community college faculty who consider themselves "strong/moderate 'hard' scientist" were concerned with non-cognitive outcomes (social, emotional and cultural growth) at a significantly lower degree than faculty who aligned themselves more with the social/behavioral sciences and humanities.

In extended research, Leverenz and Lewis (1981) found faculty often have different instructional goals depending on whether one's educational background is consistent with the current teaching appointment. The researchers found faculty with an educational background consistent with their current teaching appointment were concerned primarily with "discipline-oriented goals". Faculty whose background was inconsistent with their current teaching appointment were largely concerned with teaching students life skills.

Findings Based on Relevant Models.

Research on faculty instructional goals based on previously established higher education classification systems has found interesting results as well. Smart (1982) used Holland's theory as a classification system and found faculty from Realistic, Conventional, and Enterprising disciplines were more concerned with vocational development than faculty from Artistic, Social and Investigative fields. Smart also found faculty from Social and Artistic disciplines were more likely to be concerned with issues of personal development and character-building than faculty from other Holland environments.

Research using Biglan's model has also found interesting results. In Biglan's 1973(b) study, faculty from the "hard" sciences were more concerned with research and less concerned with teaching than faculty from the "soft" sciences. Additionally, applied disciplines appeared to be more service-oriented than pure disciplines, and nonlife-systems faculty appeared to possess a greater sense of commitment to teaching than faculty from life-systems disciplines. Smart and Elton's (1975) researched echoed much of Biglan's as they found (using the Biglan model) faculty from the "hard" disciplines were more concerned with research and student development than faculty from "soft" disciplines. Smart and Elton also found that faculty from applied disciplines shared a greater sense of commitment to service and were more concerned with student development than faculty in the pure disciplines. Further, Smart and Elton found faculty from the life-systems disciplines were more concerned with service than faculty from nonlife-system disciplines.

Conclusion

This chapter has provided an overview of the competing approaches to measurement used in quantitative research, Classical Test Theory (CTT) and Item Response Theory (CTT). One IRT model in particular, the Rasch model, was introduced and its mechanics discussed. Discussion then turned to quantitative research in higher education and an argument was made that the literature generally provides a lack of quality measurement. A meta-analysis of the methodologies employed in higher education literature and literature specific to faculty instructional goals was performed. An argument for bias was presented with regard to the validity of much of the established quantitative research in the field and a call for a different methodological approach, particularly the Rasch model, was made. This chapter closes with an examination of research related to faculty instructional goals, as well as an introduction to three popular classification systems/models typically employed in higher education literature. The next chapter outlines the methods and procedures for this study.

Chapter Three: Methodology

The purpose of this study was to provide a methodological framework for analyzing data collected via survey research techniques, especially within the realm of higher education. In doing so, this study discussed research based on CTT principles and revealed the inadequacies of its assumptions relating to reliable and valid measures. An argument for IRT, particularly the Rasch model, was presented and supported by a discussion of how the Rasch model sufficiently addresses many of the deficiencies of CTT. This study utilized data collected from the Higher Education Research Institute's (HERI) nationally-administered 2001 Faculty Survey. Data were analyzed and discussed within the framework of the one-parameter IRT, Rasch measurement, model. Using the 2001 HERI Faculty Survey data and the IRT framework, research university faculty perceptions of instructional goals were explored. The methods and procedures used in this investigation are outlined in this chapter.

Instrumentation

This study utilized the UCLA Higher Education Research Institute's (HERI) 2001 Faculty Survey. The HERI Faculty Survey is administered triennially, with the most recent survey administered in the 2004-2005 academic year. Since its inception in 1989, over 300,000 faculty at more than 1,100 higher education institutions have participated in the survey. The survey covers the following issues:

- Teaching practices and research activities
- Interactions with students and colleagues
- Professional activities
- Faculty attitudes and values
- Perceptions of the institutional climate
- Job satisfaction (HERI Faculty Survey, 2006)

Response Frame

The complete HERI Faculty Survey dataset contains over 20,000 records of faculty from all institutional types. This study investigated only regular series, tenure-track faculty at both public and private research universities, resulting in a reduced-data set. The rationale for this exclusion included issues of direct relevance and simplicity in reporting. Because this study intended to make direct comparisons to existing higher education theories and hierarchies, only faculty from relevant academic disciplines were included in the sample. The final data set for this study included 7,356 responses.

Variables

Researchers may request any number of variables from the HERI database[1]. For purposes of anonymity, the HERI masks data so individual persons and institutions cannot be identified. Requested data are distributed in aggregate form. In this study, data were requested for relevant demographic items and item #19 of the Faculty Survey, which asked faculty to "Indicate the importance to you of each of the following education goals for undergraduate students:" and provided a 4-point scale with response options: 4 = Essential; 3 = Very Important; 2 = Somewhat Important; and 1 = Not Important.

(see Table 3.1 below).

Table 3.1

Item 19 of the HERI Faculty Survey

Statement
Develop ability to think critically
Prepare students for employment after college
Prepare students for graduate or advanced education
Develop moral character
Provide for students' emotional development
Prepare students for family living
Teach students the classics of Western civilization
Help students develop personal values
Enhance the out-of-class experience of students
Enhance students' self-understanding
Instill in students a commitment to community service
Prepare students for responsible citizenship
Enhance students' knowledge of and appreciation for other racial/ethnic groups
Study a foreign language

Characteristics of Respondents

The sample for this study consisted of 7,356 regular series, tenure-track faculty from both public and private research universities throughout the United States. The term research university, in this case, refers to institutions which award doctoral level degrees in at least five different disciplines. The sample consisted of more male (66%) than female (34%) respondents.

[1] See the Appendix for a copy of the 2001 HERI Faculty Survey.

With regard to age, the sample was normally distributed with most respondents (approximately 85%) reporting between the ages of 35-64. Few faculty (about 15%) reported being less than 34 or older than 65 years of age. The sample consisted of 89.5% White/Caucasian, 4.2% Asian American/Asian, and 2.1% African American/Black respondents. American Indians, Mexican American/Chicanos, Puerto Rican Americans and other Latinos comprised the remaining 4.2% of the sample. For complete personal demographic results for sex, age, and race, see Table A1 in the Appendix.

Professional demographic characteristics for the sample include the variables: Employment Status, Principle Activity, Academic Rank, Tenure Status, Primary Interest, Type of Degree Earned, and Political Views. Ninety-five percent of the respondents in the sample were employed full-time, with 85% of the sample reporting teaching as their principle activity, and only 11.3% reporting research. With regard to academic rank, 34.7% reported holding full professor status, 27.8% as associate professor, and 24.2% as assistant professor. Faculty reporting the rank of Instructor, Lecturer, and "Other" account for the remaining 13.5%. When asked about tenure status, 58.5% of the sample reported holding tenure, while 41.5% did not. It should be noted that approximately six percent (n = 438) of the sample did not answer the question about tenure status, explaining the discrepancy in percentage based on those reporting at least associate professor rank. Over half of the respondents (56.8%) reported a primary interest in teaching, as opposed to 43.2% who reported a primary interest in research. Approximately 80% of the sample reported holding a doctorate degree and 14.8% reported holding a masters degree as the highest degree earned. Finally, with regard to political views, 17.7% reported being Conservative, 32.8% Middle of the Road, and 49.5% Liberal. For complete professional demographic results, see Table A2 in the Appendix.

As noted in Chapter 1, this study defined *academic discipline* as a field of study that is taught and/or researched at the college or university level. On the 2001 Faculty Survey, faculty were asked to select the discipline/field in which they received their highest degree. The item was titled "Major of Degree Earned" and a list of 99 academic disciplines was provided. Because only academic disciplines relevant to the theories and hierarchies discussed in this study were included in the sample, 59 disciplines were used in this study. The rationale for excluding the remaining 40 disciplines was due to a lack of direct comparison between the list of disciplines provided for each theoretical model and the list of disciplines provided on the HERI survey. Some of the disciplines that were removed from this sample include technical, interdisciplinary and broad, general fields such as: Secretarial Studies, Higher Education, Health Technology, and a host of disciplines labeled "General, Other [Foreign Languages]" in the HERI directory of disciplines.

For a complete and detailed list of the 59 academic disciplines contained in this study, including counts and frequencies, see Table A3 in the Appendix.

Instrumentation and Rasch Measurement

As discussed in Chapter 2, the Rasch model intends to create a ruler based on measures. Although the HERI Faculty Survey was not intended to measure how faculty perceive various instructional goals based on one's disciplinary affiliation, the Rasch model can manipulate items to construct a meaningful, accurate measure. To provide some background, traditional psychometrics analyzes raw scores, a process which is purely descriptive in nature. With traditional psychometrics, a snapshot is taken of the research situation. The snapshot reveals how a specific sample of people is responding to a particular item, or set of items. When this happens, all the elements are bound together haphazardly. When fellow social scientists replicate the study, they take an additional snapshot. Ultimately, snapshots are compared, although the samples therein are not directly comparable. Information produced from each snapshot simply provides a description of what is happening in each research situation at the time of survey (or test) administration. Recall that raw scores are not measures and are not linear. Rasch measurement, however, can convert raw scores to measures and untangle all the haphazardly bound elements to form a straight line. Once data are in linear form, they can be calibrated to provide a ruler which can be used to take measurements. Unlike the use of raw scores and its resulting fuzzy descriptions, results from Rasch analysis provides measures that are both precise and stable across samples and time (Meaningful Measurement, 2008). For the purposes of this study, the HERI Faculty Survey dataset was used to construct a measure for the purposes outlined above.

Data Analysis

The analysis began with cleaning the dataset, followed by an exercise in recoding data. An index of academic disciplines was coded on a scale of 1-99 to represent the 99 degree fields outlined in the HERI Faculty Survey. As stated previously, only data for academic disciplines with direct relevance to the theories and models employed in this study were retained, resulting in 59 disciplines. Data were then exported to SPSS (version 15.0) where descriptive statistics were generated for each of the demographic variables. These descriptive statistics included counts, frequencies, and percentages, which intended to depict the characteristics of the study's sample.

Next, fit of the data to the Rasch model was examined. Winsteps software (version 3.51) was used to complete the Rasch analysis. The Rating Scale Model was utilized, as Bond and Fox (2001) define the model as "a version of the Rasch model… routinely used for the sort of polytomous data generated by Likert scales" (p. 233). In other words, the Rating Scale Model assumes every item on a survey has the same number of response categories for all questions.

Fit of the data to the model was determined by assessing summary and fit statistics, separation measures, and reliability statistics. Although many researchers use the commonly accepted range of 0 to 2.0 to assess fit (Linacre, 2004a), Wright and Linacre (1994), as cited in Bond and Fox (2001), recommend a mean square cutoff criteria ranging from 0.6 to 1.4 for both infit and outfit statistics for Likert-type surveys utilizing a rating scale (p. 179). Items with mean-square values greater than 1.4 or less than 0.6 were identified as potentially misfitting and highlighted for further review. Separation measures assess whether the items discriminate levels of importance for faculty educational goals. Due to the large sample size with only 14 items, separation measures amongst items were expected to be larger, and smaller amongst persons. Reliability estimates were expected to be at least moderately high, as indicated by an estimation of at least .70 (Bruning and Kintz, 1997). All of these analyses were performed using the Summary Statistics function in Winsteps.

Next, to determine the "functioning" of the rating scale and to test for unidimensionality, the rating scale structure was analyzed. In measurement research, functioning refers to how well the rating scale captures data. Particularly, does the rating scale provide an appropriate number of possible responses to each item? Does the rating scale force respondents to provide answers consistent with the construct being measured? Does the rating scale force respondents to use the same set of possible response options? (Low, 1988). Response options were coded (4 = Essential; 3 = Very Important; 2 = Somewhat Important; and 1 = Not Important) and graphically displayed via the Probability Curves function in Winsteps. When evaluating these curves, it is important to notice the proximity of response options to one another. Ideally, response patterns are somewhat evenly dispersed and each response is clearly separated from the others. A proper distribution along the probability curve provides evidence that the four-category rating scale is functioning properly.

Another technique to test for functioning is to investigate rating scale diagnostics. Rating scale diagnostics were used to determine how well the four response options created an interpretable measure. This analysis was performed by using the Category Function technique in Winsteps. By examining the shape of the observed count/frequency distribution one can infer whether the data falls along a normal distribution curve. Counts and frequencies were provided for each of the response options (Essential, Very Important, Somewhat Important, and Not Important) and the shape of the distribution was evaluated. Normally distributed data indicates respondents fully utilized all response options on the scale and the response options were sufficient in both breadth and appropriateness.

Once stability of the measure was determined, item maps were constructed to present a graphical illustration that would visually display any potential relationships amongst item responses. These maps display person and item distributions along a hierarchy, usually according to highest/lowest average rankings. In this study, however, items were arranged according to its level of difficulty to endorse. The items most difficult to endorse fell at the top of the map and the easiest to endorse fell at the bottom. The left side of the item map provided a scale which include means and standard deviations from the mean, indicated by "M" (indicating the mean), "S" (indicating one standard deviation), and "T" (indicating two standard deviations). The proximity of items in relation to the others offered visual evidence of the relationships amongst each of the items.

Differential Item Functioning (DIF) techniques were carried out to determine how items were functioning amongst various subgroups. Rasch measurement assumes that test takers, or in this case the individuals responding to the survey, with similar knowledge and abilities or opinions, will respond alike regardless of sex, race, etc. DIF allows data to be examined by subgroup to detect differences amongst their responses on a given variable. In this study, DIF was used to detect differences among faculty perceptions of instructional goals based on one's academic discipline.

In order to interpret DIF, several issues should be discussed, including how to ascertain if DIF is present. Throughout the literature, DIF is presented in multiple contexts, utilizing statistical and measurement models. According to Thissen, Steinberg, and Wainer (1993), as cited in Roever (2005), "IRT techniques are the 'gold standard' of DIF detection" (p. 5). A number of other researchers agree (see Lord, 1980; Ironson, 1982; Shepard, Camilli and Williams, 1984; and Marascuilo and Slaughter, 1981). Here, DIF was detected through the one-parameter Rasch model. According to Zwick & Thayer (1996), "average observation" measures are the primary indicator of DIF, so these values were given focal attention in this study. These values can range anywhere from 0.0 – 3.0. A value of 3.0 would indicate perfect agreement amongst the subset or group being compared with the item of interest. Conversely, a value of 0.0 would indicate complete disagreement amongst the persons as it relates to the item of interest. An example might include asking faculty to rate how important they perceive the educational goal "Prepare students for family living", and then comparing faculty responses from each academic discipline for the item. Faculty from the area of Home Economics might have a DIF observation measure of 2.0, indicating high agreement from the collective faculty from this discipline, and faculty from Marine Science might have a DIF observation measure of 0.2, indicating very low agreement among the collective faculty from this discipline about the importance of this particular outcome.

Anchoring the DIF value range at the mean for each set of scores provides a more meaningful scale to determine how various disciplines score in relation to one another. Disciplines with a score higher than the mean would indicate those particular disciplines are more likely to find agreement with the item of interest. Conversely, disciplines with a score lower than the mean would indicate those disciplines are more likely to find disagreement with the item of interest.

How does one identify meaningful differences regarding DIF? Because some researchers argue DIF can result by accident, the best test for detecting meaningful differences in DIF is by asking the question "does it replicate?" (Du, 1995). Using large and diverse samples, like the one utilized in this study, is one way to provide strong evidence for the real presence of DIF. Replicating studies using different samples, whether random or systematic, is the other primary method for determining meaningful differences.

Finally, results of the DIF analyses were arranged in hierarchical form ranging from the highest average observation score to the lowest for each of the 14 items. This resulted in a hierarchy of academic disciplines for each of the items. A hierarchy was then generated for each dimension of the Biglan, Kolb, and Holland models, and then a table for each model was formed. These tables served as checklists, which would contain an "X" in the appropriate column if the particular dimension of the model was present in the upper bound of the hierarchy, and an "O" if the particular dimension of the model was present in the lower bound of the hierarchy, for each of the 14 items. An anchor was set in each DIF range at the mean. This criteria was used to determine what constituted the upper and lower bounds for each hierarchy. Through a careful, thematic comparison, each item's hierarchy was evaluated to determine if any of the aforementioned models' themes were present amongst the academic disciplines. This involved direct comparisons and matching of academic disciplines from the DIF item hierarchies to those from each dimension of the higher education models. The results were plotted with the intent to provide a visual display of both the frequency and relevance of each model's dimensions to the 14 instructional goals.

Summary

This study utilized Rasch measurement to investigate faculty perceptions of instructional goals. This chapter outlined the methodology employed, including instrumentation, sampling, and the variables under investigation. A description of the sample's characteristics was exhibited and a detailed explanation of data analysis procedures was presented. This study provided a systematic procedure for assessing item fit and functioning. Specifically, relevant items from the HERI Faculty Survey were evaluated to determine how well the items measured faculty instructional goals, and fit statistics were investigated to determine how well the data fit the

Rasch model. Procedures for identifying and reviewing possible misfitting items were then described, and explanations were presented for the use of item maps and DIF techniques. Finally, procedures were discussed for tying the study's results back to theory by comparing results to previously established higher education models. The next chapter will focus on data analysis and results.

Chapter Four: Analysis and Results

This chapter presents the results of the Rasch analysis conducted to investigate faculty perceptions of instructional goals. Results include the analyses of data-to-model fit and survey item functioning, the investigation and review of potentially misfitting items, item map and DIF results, and the extension of results back to higher education theory via direct comparison to popular classification systems. The following questions guided the analyses:

1. How well do items from the HERI Faculty Survey measure faculty instructional goals among university faculty?
2. Do relevant items on the HERI Faculty Survey fit the expectations of the one-parameter IRT (Rasch) model by forming a unidimensional construct?
3. How does a hierarchy of academic disciplines compare to previously established higher education classification systems?
4. In what ways does presumed paradigmatic consensus influence faculty instructional goals?

The first two questions pertain to issues specific to the Rasch model, particularly identifying the "fit" and "functioning" of the data to the model. These questions are necessary each time one uses the Rasch model, as they illustrate to what extent the model is an appropriate technique for data analysis. The subsequent questions will be answered through the results of the actual Rasch model application. The Rasch model allows for the construction of item maps for both persons and items. In this study, the utility of item maps will be demonstrated by mapping academic disciplines and comparing their hierarchy to popular classifications systems previously found in higher education research, such as the models of Anthony Biglan (1973), David Kolb (1980), and John Holland (1966). The Rasch model will also allow analysis for DIF. The combination of the item and person maps and DIF will support identification of themes amongst the academic disciplines.

Fit of the Data to the Model

The initial analyses sought to answer the first two research questions by determining data/model fit and rating scale functioning. This procedure involved calculating "Summary Statistics" including the means, standard deviations, separation and reliability estimates for both persons and items. Fit statistics were computed to address to what extent the data fit the model. Table 4.1 displays the summary statistics.

Table 4.1

Summary Statistics

	Measure	Model Error	Infit Mean Square	Outfit Mean Square
Faculty				
Mean	.19	.40	1.01	1.03
S.D.	1.07	.05	.55	.66
Item				
Mean	.00	.02	1.01	1.03
S.D.	1.26	.00	.28	.31

Item infit and outfit statistics identify the extent to which items fit the Rasch model. Mean square measures are typically considered the most important measure in this category. According to Linacre (2004), mean squares display the "size of the randomness". Values of 1.0 are ideal because values higher than 1.0 (overfit) represent unpredictability, and values lower than 1.0 (underfit) represent observations that are too predictable, which ultimately may inflate reliability statistics.

A mean squared value range cutoff is usually determined by the type of test. In the case of a survey, Wright and Linacre (1994), as cited in Bond and Fox (2001), recommend a mean square cutoff criteria ranging from 0.6 to 1.4. The authors acknowledge there are no specific rules for determining an appropriate cutoff range, but research has shown the aforementioned cutoff to be reasonable given its type of test (p. 179). Adhering to Wright and Linacre's recommendations, items with logit measures higher than 1.4 may overfit and items less than 0.6 may underfit, thus making the measures potentially misfitting. Usually, misfitting items suggest the instrument is failing to measure what it is intending to measure, or respondents are interpreting items differently than the researcher(s) intended when the questionnaire was generated. Table 4.2 outlines mean square measures for both infit and outfit item statistics.

Table 4.2

Fit Statistics for Each of the Items

Item	INFIT Mean Square	OUTFIT Mean Square
Develop ability to think clearly	1.09	1.35
Prepare students for employment after college	1.32	1.40
Prepare students for graduate or advanced education	1.02	1.11
Develop moral character	.86	.85
Provide for students' emotional development	.68	.68
Prepare students for family living	.94	.86
Teach students the classics of Western civilization	1.55	1.56
Help students develop personal values	.69	.68
Enhance the out-of-class experience of students	.98	.99
Enhance students' self-understanding	.89	.88
Instill in students a commitment to community service	.71	.70
Prepare students for responsible citizenship	.79	.79
Enhance students' knowledge of and appreciation for other racial/ethnic groups	1.00	.99
Study a foreign language	1.56	1.57

Evaluation of Misfitting Items.

Following the suggested cutoff range mentioned above, misfitting items were evaluated. Linacre (2004) says items above the 1.4 threshold suggest "off-variable noise". Depending on how far the scores extend past the range could determine how useful the measures are. Scores closer to the range may not affect measurement (positively or negatively), however scores with greater deviations may indicate the presence of non-useful items. Scores with the greatest deviations from the suggested ranges are considered "noisy" items. Items below the 0.6 threshold may be considered "overly predictable" and might lead the researcher to believe his or her measures are better than they are in actuality. Same as before, items closer to the range may not affect measurement, but scores with greater deviations may indicate the presence of non-useful and noisy items. If noisy items are present, it typically means extreme categories are being overused. To remedy noisy items, Linacre suggests researchers check for poorly worded items or consider collapsing response categories. Again, items between the 0.6 and 1.4 thresholds are ideal for productive measurement.

According to the above table, only two items were potentially misfitting in relation to the suggested cutoff range of 0.6 – 1.4 for Likert-type surveys which utilize a rating scale. The two items in question were: *Teach students the classics of Western civilization* and *Study a Foreign Language*. Because both sets of values (1.55 and 1.56; and 1.56 and 1.57, respectively) fall outside the suggested cutoff range of 0.6 – 1.4, further inspection of these items was necessary. After investigating Item Characteristic Curves (ICCs) and the distribution of both misfitting items and persons, it appeared the items were sound, but the Rasch model may have expected faculty to respond somewhat differently than they actually did based on individual response patterns per person. One possibility could be the item "teach students the classics of Western civilization" may have been a bit unclear, or at least unspecific for some respondents. Because these two items demonstrated fit statistics close to the cutoff range and because their ICCs demonstrated sufficient deviation between responses, there was sufficient evidence to retain these items and include them in subsequent analyses.

Reliability.

Winsteps software reports both person and item reliability measures. Person reliability is approximately equivalent to the traditional test reliability, as low estimates suggests a small range of person measures or items and high estimates suggest a sufficient range of person or item measures (Linacre, 2004). Item reliability has no statistical equivalent in classical test tradition. Generally, low item reliability is a sign of small sample size or a small range of item measures, which consequently affects the stability of measures. High item reliability measures indicate sufficient sampling for stable measures.

Winsteps software reports reliability in two forms: 1) Real; and 2) Model. Real reliability pertains to the lower bound estimate and reports reliability values at their worst. Model reliability pertains to the upper bound estimate and reports reliability at its best. True reliability estimates fall somewhere in between the two measures (Linacre, 1997). In this case, Real person reliability is .83 and Model person reliability is .86, therefore the true person reliability estimate would be somewhere between the .83 and .86 range. Extending to traditional test reliability, Cronbach's Alpha estimates were analyzed via SPSS software. This estimate was .85. According to Linacre (1997), Cronbach's Alpha overestimates reliability and Rasch measurement underestimates it. For item reliability, both Real and Model estimates are 1.0. Because reliability estimates rarely equal a perfect score of 1.0, it is likely this statistic is inflated as there were over 7,500 responses to just

14 items. Table 4.3 below highlights reliability estimates according to both Winsteps and SPSS software programs.

Table 4.3

Comparison of Reliability Estimates

	Real Reliability	Model Reliability	Cronbach's Alpha
Item	.83	.86	.85
Person (Faculty)	1.00	1.00	* (no equivalent)

Evaluating Rating Scale Function

Rating Scale Structure.

Probability curves were calculated to evaluate the quality of the rating scale structure. Response options are coded (4 = Essential; 3 = Very Important; 2 = Somewhat Important; and 1 = Not Important) and graphically displayed in Figure 4.1.

```
INPUT: 7356 FAC, 14 RSPS MEASURED: 7269 FAC, 14 RSPS, 4 CATS v3.62.1
-----------------------------------------------------------------------

      CATEGORY PROBABILITIES: MODES-Structure measures at intersections
P      ++---------+---------+---------+---------+---------+---------++
R  1.0 +                                                               +
O      |                                                               |
B      |                                                               |
A      |                                                               |
B   .8 +                                                              4+
I      |111                                                       444 |
L      |   11                                                   44    |
I      |     11                                               44      |
T   .6 +       11                                           44        +
Y      |         11       222222222                       44          |
    .5 +           11 222         222        3333333333  4            +
O      |             22*             222333            3**            |
F   .4 +           22   11             3322           44  333         +
       |       222        11        333    22        4       33       |
R      |     22             11     33        22   44           333    |
E      |  222                 11 33            2*4                333 |
S   .2 +2                     33*1            444 22                 33+
P      |                   333    111       44      222               |
O      |               3333          111*444           2222           |
N      |        3333333         444444 111111              2222222    |
S   .0 +******4444444444444444             1111111111111111111*****+
E      ++---------+---------+---------+---------+---------+---------++
       -3        -2        -1         0         1         2         3
           FAC [MINUS] RSP MEASURE
```

Figure 4.1: Probably Curve

Probability curves should demonstrate separation amongst responses; that is, response patterns should be independent and separate from others. Responses should not flat-line, or closely follow another similar pattern, as it would suggest respondents could not truly distinguish the difference between some of the response options. When probability curves are evenly dispersed and the responses separated (as in the figure above), this provides evidence that the four-category rating scale functioned as intended.

In Table 4.4, rating scale diagnostics were used to determine how well the four response options created an interpretable measure. By examining the shape of the observed count distribution it appears the data fall along a normal distribution curve. Based on the probability curve and rating scale diagnostic data, there is sufficient evidence to suggest faculty who responded to the HERI Faculty Survey fully utilized the range of response options on the scale. Further, this provides evidence the response options for this particular item were both sufficient in breadth and appropriateness. Table 4.4 presents both counts and frequencies for response options utilized by faculty, as well as fit statistics for each response option.

Table 4.4

Summary of Rating Scale Diagnostics

Category	Observed Count (%)	INFIT Mean Square	OUTFIT Mean Square
Essential	15495 (15%)	.93	.95
Very Important	35100 (35%)	.97	.96
Somewhat Important	30621 (30%)	.98	1.06
Not Important	19540 (19%)	1.10	1.13

Note. Category, observed count, and percentage indicate the numbers of respondents who chose a particular response category, summed for each category across all 14 items.

Summary of Fit and Functioning

Based on evidence resulting from the evaluation of summary statistics for the model and fit statistics for each item, and the investigation of potentially misfitting items, it was clear the data formed a unidimensional construct and adequately addressed the assumptions of the one-parameter Rasch model. Based on evidence ascertained from examining the rating scale structure and rating scale diagnostics, the rating scale functioned appropriately by adequately measuring what it intended to measure. It was concluded that Rasch analyses was an appropriate technique for analyzing these data. The next section will report the results of additional Rasch analyses and will make inferences about the data.

Faculty Instructional Goals Results

This section will report the results of the Rasch analysis as it pertains to faculty instructional goals. First, counts and percents will be reported for each of the 14 items. Next, item maps will be presented. Item maps will visually display response patterns of HERI Faculty Survey respondents. This will be followed by a report of Differential Item Functioning (DIF) results. DIF will indicate how faculty from various academic disciplines responded to individual items on the survey. Finally, existing classification systems and models in the current higher education literature will be compared to item map and DIF results to determine if any of the existing models resemble the hierarchies formed from this analysis, and to what extent. These analyses will answer this study's third and fourth research questions.

Counts and Percentages.

Counts and percentages of responses are reported for each of the 14 items. Table 4.5 below summarizes these results.

Table 4.5

Counts and Percents of Responses for Each of the 14 Items

	Essential	Very Important	Somewhat Important	Not Important
Think critically	6384	833	37	2
	(88%)	(11%)	(1%)	(0%)
Employment after college	1557	3109	2230	342
	(22%)	(43%)	(31%)	(5%)
Prepare for graduate education	766	3460	2837	181
	(11%)	(48%)	(39%)	(2%)
Moral character	1710	2392	2461	660
	(24%)	(33%)	(34%)	(9%)
Emotional development	588	1882	3442	1307
	(8%)	(26%)	(48%)	(18%)
Family living	254	756	2751	3434
	(4%)	(11%)	(38%)	(48%)
Classic works of Western civilization	598	1631	2758	2220
	(8%)	(23%)	(38%)	(31%)
Personal values	1361	2754	2464	626
	(19%)	(38%)	(34%)	(9%)
Enhance out-of-class experiences	670	2024	3095	1411
	(9%)	(28%)	(43%)	(20%)
Self Understanding	1535	2734	2302	634
	(21%)	(38%)	(32%)	(9%)
Community service	557	1935	3315	1400
	(8%)	(27%)	(46%)	(19%)
Responsible citizenship	1377	2866	2368	572
	(19%)	(40%)	(33%)	(8%)
Racial/ethnic appreciation	1639	2463	2337	772
	(23%)	(34%)	(32%)	(11%)
Foreign language	774	1785	2703	1940
	(11%)	(25%)	(38%)	(27%)

Based on the responses indicated in the table above, it appeared faculty overwhelmingly agreed the item "Think critically" was most essential (88%). The least essential item, as reported by overall faculty, was the item "Family Living". Only 4% of faculty rated this item essential, whereas 48% rated the item as "Not Important".

Item Maps.

Item-person maps are useful for identifying meaning constructs, as these graphical illustrations visually display any potential relationships amongst item responses. These maps display person and item distributions along a hierarchy, usually according to highest/lowest p-values or highest/lowest average rankings. Here, the numbers along the left column indicate logit measure. On this map, these logits descend according to difficulty, meaning the hardest item to endorse will fall at the top of the map and the easiest item to endorse will fall at the bottom of the map. "M" markers along the map indicate the location of the mean measure. Likewise, a marker of "S" indicates one standard deviation from the mean and "T" indicates two standard deviations from the mean, as shown by Figure 4.2.

```
Item Map of Responses
                <more>|<rare>
    5               .  +
                       |
                    .  |
                       |
                       |
    4               .  +
                    .  |
                    .  |
                    .  |
                    .  |
    3               .  +
                    .  |
                    .  |T
                   .# T|
                   .#  |
    2              .#  +  Family living
                   .#  |
                  .##  |
                 .###  |
             .####### S|S
    1           .####  +  Classic works of West civ
                .####  |  Community service           Foreign language
          ###########  |  Emotional development       Enhance out-of-class exp
               .######  |
        .############ M|
    0          .######  +M
               .######  |  Moral character             Personal values
                          Prepare for graduate ed      Racial/ethnic appreciation
                          Responsible citizenship      Self-understanding
         .###########  |  Employment after college
               .#####  |
                .#### S|
   -1          .######  +
                  .##  |S
                  .##  |
                   .#  |
                   .#  |
   -2              .# T+
                    .  |
                    .  |
                    .  |T
                    .  |
   -3               .  +
                    .  |
                       |
                    .  |
                    .  |  Think critically
   -4                  +
                       |
                       |
                       |
                       |
   -5                  +
                       |
                       |
                       |
                       |
   -6               .  +
                <less>|<frequ>
 EACH '#' IS 64.
```

Figure 4.2: Hierarchy Map of Persons and Items

The above item map reveals respondents to the HERI Faculty Survey found very little difficulty in endorsing the item *Think Critically*. The second easiest item to endorse was *Employment after College*. Items *Moral Character, Personal Values, Prepare for Graduate Education, Racial/Ethnic Appreciation, Responsible Citizenship,* and *Self-Understanding* followed very closely with virtually identical endorsability measures. Slightly more difficult items to endorse included *Emotional Development* and *Enhancing Out-of-Class Experiences*. The hierarchy continues upward until it reaches the most difficult item to endorse, *Family Living*.

Differential Item Functioning.

To review, Differential Item Functioning (DIF) assumes individuals responding to the survey, with similar knowledge and abilities, or opinions, will perform alike regardless of sex, race, etc. DIF allows data to be examined by subgroup to detect differences between their responses on a given variable. Here, DIF was used to detect differences among faculty instructional goals based on one's academic discipline. Below, relevant DIF results and mean DIF scores for each item will be presented. See Tables A4 - A17 the Appendix for complete DIF results by item.

Relevant DIF Results

In this section, relevant DIF results will be presented for the 14 items of interest on the HERI Faculty Survey. Each of these items will be examined individually.

Item 1 - Develop Ability to Think Clearly.

Based on the results of the Item Map, faculty found this item the easiest to endorse. DIF analysis confirmed there was little variance between faculty responses from all disciplines. To illustrate the point, a total range of only .26 resulted from the highest and lowest Average Observation scores of the 59 disciplines represented. In other words, there was no clear evidence to suggest faculty from any particular disciplines were more or less concerned with this particular construct. The mean DIF score for this item was 2.87.

Item 2 - Prepare students for employment after college.

The highest scoring disciplines for this item included disciplines from applied fields such as Education and Engineering. The single highest measure, however, was Nursing. The lowest measures included several disciplines within the humanties, such as Religion, Philosophy, Foreign Languages, History, Art and English. A couple of the social science disciplines also scored very low, these included Political Science and Anthropology. The mean DIF score for this item was 1.92.

Item 3 - Prepare students for graduate or advanced education.

The disciplines with the highest average observations scores for this item included disciplines primarily from the "hard" sciences, such as Medicine, Nursing, Physiology, Biochemistry, Zoology, Biology, Chemistry, Botany, Physics and a few fields from Engineering. There were a few "softer" fields falling sporadically near the top of the distribution, including Music, Social Work and Spanish. Disciplines reporting the lowest average observations scores included Law Enforcement, Marketing, Journalism, Management, History and Agriculture. The mean DIF score for this item was 1.67.

Item 4 - Develop moral character.

The four fields with distinguishing preference for developing moral character include: Medicine, Law Enforcement, Elementary Education and Nursing. Disciplines indicating the least concern for moral character included a number of social science disciplines such as Economics, Psychology, Political Science, Anthropology and Sociology. Astronomy reported the lowest measure, with Math and Statistics also near the bottom of the distribution. The mean DIF score for this item was 1.77.

Item 5 - Provide for students' emotional development.

With regard to providing for students' emotional development, applied disciplines such as Medicine, Nursing and various disciplines within Education (including Elementary Education, Educational Psychology and Counseling, Home Economics, Education Administration, and Physical and Health Education) reported the highest measures. With the exception of several fields from the Engineering disciplines, the majority of disciplines with the lowest measures came from disciplines in the "pure" sciences. These included: Economics, Astronomy, Political Science, Biochemistry, History, Math and Statistics, Botany, Microbiology, Zoology and Physics. The mean DIF score for this item was 1.27.

Item 6 - Prepare students for family living.

Home Economics scored considerably higher on this measure that other disciplines. In fact, a range of .65 exists between the next closest discipline, Elementary Education. Other educational fields such as Education Administration, Educational Psychology and Counseling, Secondary Education and Physical and Health Education were the next closest measures, with Medicine and Law Enforcement also near the top of the distribution. The lowest measures came from various engineering disciplines (e.g., Civil, Chemical, Electrical, Mechanical and Nuclear), and from Physiology, Astronomy and Architecture. The lowest measure, however, came from Marine Science with International Business just above it. The mean DIF score for this item was .73.

Item 7 - Teach students the classics of Western civilization.

A number of disciplines from the humanities reported the greatest concern for teaching students the classics of Western civilization. Philosophy, Music, Religion, Dramatics and Speech, English and several foreign languages (including French, German and Spanish) all reported measures at the top of the distribution. The lowest scores on the distribution were dominated by various disciplines within Business and Engineering. In particular, Business-related disciplines such as General Business, International Business, Finance, Marketing, Accounting and Management, and Engineering-related disciplines such as Mechanical, Civil, Electrical and Chemical Engineering reported the lowest measures for this construct. The mean DIF score for this item was .99.

Item 8 - Help students develop personal values.

The disciplines reporting the highest concern for helping students develop personal values included numerous Education fields, such as Home Economics, Elementary Education, Education Administration, Educational Psychology and Counseling and Special Education. Also included near the top of the distribution were Nursing, Medicine, Religion and Law Enforcement. The disciplines reporting the lowest measures included several disciplines from the "hard" sciences, such as Astronomy, Marine Science, Math and Statistics, Physics and Zoology. The mean DIF score for this item was 1.70.

Item 9 - Enhance the out-of-class experience of students.

Disciplines mostly concerned with enhancing students' out-of-class experience include several from the field of Education (Education Administration, Educational Psychology and Counseling, Home Economics, Physical and Health Education, and Elementary Education) and other applied fields such as Law Enforcement, Social Work, Agriculture, Nursing, Art and Journalism. Disciplines least concerned with enhancing students' out-of-class experience include a number of the "pure" sciences. In particular, Math and Statistics, Economics, History, Political Science, Chemistry, Philosophy, Religion and Physics. The mean DIF score for this item was 1.34.

Item 10 - Enhance students' self-understanding.

With regard to enhancing students' self-understanding, several Education disciplines (e.g., Elementary Education, Educational Psychology and Counseling, Education Administration, Special Education, Music and Art Education, Physical and Health Education and Home Economics) as well as humanities disciplines reported the highest average observation scores. The humanities disciplines included Art, Dramatics and Speech, Religion, English and Literature, Music and Philosophy. The lowest average observation scores were reported from several engineering fields (Electrical, Chemical, Nuclear and Mechanical) and other "hard" science disciplines such as Math and Statistics, Biochemistry, Physics, Chemistry, Astronomy, Geography, Zoology and Computer Science. A couple "softer" disciplines, Economics and International Business, were also found at the bottom of the distribution. The mean DIF score for this item was 1.69.

Item 11 - Instill in students a commitment to community service.

Disciplines with the highest average observation scores for this item were predominantly from the "life" sciences. These include: Social Work, Law Enforcement, Nursing, Medicine and several fields from Education (Elementary Education, Education Administration, Educational Psychology and Counseling, Special Education, Secondary Education, Home Economics and Health and Physical Education). Disciplines with the lowest scores were from "non-life" fields such as Astronomy, Economics, Math and Statistics, Computer Science, Chemistry, Physics, Finance and various Engineering disciplines (e.g., Mechanical, Electrical, Nuclear, Chemical and Aeronautical/Astronautical). The mean DIF score for this item was 1.28.

Item 12 - Prepare students for responsible citizenship.

"Life" sciences comprised the majority of disciplines with the highest average observation scores for preparing students for responsible citizenship. These disciplines included various disciplines from Education (Elementary Education, Education Administration, Educational Psychology and Counseling, Secondary Education, Physical and Health Education and Home Economics), as well as, Social Work, Law Enforcement, Medicine and Nursing. Conversely, "non-life" sciences comprised the majority of disciplines with the lowest average observation scores. These disciplines included Math and Statistics, Electrical and Mechanical Engineering, Finance, Astronomy, Computer Science, Physics and Chemistry. The mean DIF score for this item was 1.73.

Item 13 - Enhance students' knowledge of and appreciation for other racial/ethnic groups.

The disciplines with the highest average observations scores for this item included disciplines primarily from the "life" sciences, such as Social Work, Anthropology, Nursing, Law Enforcement, and a plethora of disciplines from the Education arena (e.g., Elementary Education, Educational Psychology and Counseling, Secondary Education, Music and Art Education, Education Administration and Special Education). The disciplines with the lowest average observation scores included disciplines primarily from the "hard" and "non-life" sciences, such as Computer Science, Chemistry, Physics, Astronomy, Math and Statistics, Finance, and several fields from the Engineering arena (e.g., Electrical, Mechanical, Chemical and Civil). The mean DIF score for this item was 1.68.

Item 14 - Study a Foreign Language.

As expected, disciplines specializing in foreign languages (German, French, Spanish and Foreign Language and Literature) were at the top of this distribution. Disciplines with the least average observation scores included several from Engineering (Civil, Electrical and Mechanical), Finance, Medicine, Computer Science, Biochemistry and Pharmacy. The mean DIF score for this item was 1.15.

Extension to Higher Education Classification Systems

As noted previously, for nearly half a century the higher education literature has explored several classification systems and/or models in an effort to explain various phenomena relating to academic disciplines. Some of the more popular classification systems/models include Biglan and Kolb's models and Holland's theory. Anthony Biglan's model classifies disciplines according to 'hardness' ("soft" versus "hard" sciences), whether the field is pure versus applied in nature, and whether it pertains to life versus non-life subjects (Biglan, 1973a; 1973b). David Kolb's research on learning styles and experiential learning (1980) added to Biglan's model by including two additional dimensions: "active" versus "reflective", and "abstract" versus "concrete". Holland's theory offers six personality/environment "types" (Holland, 1966; Smart, Feldman, and Ethington, 2000). For a comprehensive list of academic disciplines that comprise each of these models/systems, see Tables A18 – A20 in the Appendix.

Upon review of the DIF analysis, the results were investigated according to the various dimensions of each of the aforementioned classifications systems/models. As noted in Chapter 3, results of the DIF analyses were arranged in hierarchical form ranging from the highest average observation score to the lowest for each of the 14 items. This resulted in a hierarchy of academic disciplines for each of the items. A hierarchy was then generated for each dimension of the Biglan, Kolb, and Holland models, and then a table for each model was formed. These tables served as checklists, which would contain an "X" in the appropriate column if the particular dimension of the model was present in the upper bound of the hierarchy, and an "O" if the particular dimension of the model was present in the lower bound of the hierarchy, for each of the 14 items. The ordering of the letters "X" and "O" reveal the location of both the first and second construct in each dimension. For example, if one were examining the 'hard' v. 'soft' dimension of Biglan's model, a mark of O/X would indicate faculty from 'hard' science disciplines could be found at the lower bound of the hierarchy, whereas faculty from 'soft' science disciplines could be found in the upper bound. An anchor was set in each DIF range at the mean. This criteria was used to determine what constituted the upper and lower bounds for each hierarchy. Through a careful, thematic comparison, each item's hierarchy was evaluated to determine if any of the aforementioned models' themes were present amongst the DIF results of academic disciplines. This involved direct comparisons and matching of academic disciplines from the DIF item hierarchies to those from each dimension of the higher education models. The results were plotted to provide a visual display of both the frequency and relevance of each model's dimensions to the 14 instructional goals. The results of these comparisons are presented in Tables 4.6, 4.7, and 4.8.

Table 4.6

Hierarchy According to Biglan Dimensions

	Hard v. Soft dimensions	Pure v. Applied dimensions	Life v. Nonlife dimensions
Develop ability to think critically			
Prepare students for employment after college		O/X	
Prepare students for graduate or advanced education	X		
Develop moral character	O/X	O/X	
Provide for students' emotional development	O/X	O/X	
Prepare students for family living	O/X	O/X	X/O
Teach students the classics of Western civilization	O/X	X/O	
Help students develop personal values	O/X	O/X	X/O
Enhance the out-of-class experience of students	O/X		X/O
Enhance students' self-understanding	O/X		
Instill in students a commitment to community service	O/X	O/X	X/O
Prepare students for responsible citizenship	O/X		X/O
Enhance students' knowledge of and appreciation for other racial/ethnic groups	O/X		X
Study a foreign language	O/X	X/O	

Table 4.7

Hierarchy According to Kolb Dimensions

	Active v. Reflective dimensions	Concrete v. Abstract dimensions
Develop ability to think critically		
Prepare students for employment after college	X/O	
Prepare students for graduate or advanced education		X/O
Develop moral character	X/O	O/X
Provide for students' emotional development	X/O	O/X
Prepare students for family living	X/O	O/X
Teach students the classics of Western civilization	O/X	
Help students develop personal values	X/O	O/X
Enhance the out-of-class experience of students	X	O/X
Enhance students' self-understanding	X/O	O/X
Instill in students a commitment to community service	X/O	O
Prepare students for responsible citizenship	X/O	O/X
Enhance students' knowledge of and appreciation for other racial/ethnic groups		O/X
Study a foreign language		O/X

Table 4.8

Hierarchy According to Holland Dimensions

	Social	Artistic	Realistic	Investigative	Conventional	Enterprising
Develop ability to think critically						
Prepare students for employment after college		O				
Prepare students for graduate or advanced education					O	O
Develop moral character		O		O		O
Provide for students' emotional development	X		O	O		O
Prepare students for family living	X		O		X	O
Teach students the classics of Western civilization		X		O		
Help students develop personal values	X	X	O	O	X	O
Enhance the out-of-class experience of students	X		O	O	O	O
Enhance students' self-understanding	X	X	O	O	O	O
Instill in students a commitment to community service	X		O	O		O
Prepare students for responsible citizenship	X		O			
Enhance students' knowledge of and appreciation for other racial/ethnic groups	X	X	O			
Study a foreign language		X	O	O		

Summary

This chapter presented the results of the Rasch analysis. First, it was determined that the data fit the model. Second, it was determined that the survey's rating scale functioned properly, which further validated the instrument. Testing for fit and functioning resulted in the determination that the Rasch model was an appropriate technique for data analysis. Third, item maps revealed the order in which faculty from various disciplines found agreeability among the 14 survey items. Next, via DIF techniques, a hierarchy was formed for the 14 survey items and the results were presented for each item. Finally, the results of the DIF analyses were compared to three popular higher education classification systems/models and the results were presented for each model. The final chapter will present a summary of the results and findings, provide a discussion of the results, and address implications and avenues for future research.

Chapter Five: Summary, Major Findings, Conclusions, and Recommendations for Future Study

The purpose of the study was to provide a methodological framework for analyzing data collected via survey research techniques, especially within the realm of higher education. Further, this research sought to investigate faculty perceptions of instructional goals based on faculty responses to the 2001 Faculty Survey administered by UCLA's Higher Education Research Institute. This study discussed research based on CTT principles and revealed the inadequacies of its assumptions relating to reliable and valid measures. An argument for IRT, particularly the Rasch model, was made and supported by a discussion of how the Rasch model sufficiently meets many of the deficiencies of CTT. A test for model fit and rating scale functioning was then presented. Once proper calibrations were made and sufficient evidence was given for the fit of the data to the model, the data were further analyzed. Results of the analyses were presented in the previous chapter. This chapter will summarize the study and present relevant findings and conclusions, as well as a discussion of the results. Implications of this research and avenues for future research will also be included. The chapter will close with a summary of the major contributions of the current study and recommendations for future study.

Research Questions and Findings

Question 1: How well do items from the HERI Faculty Survey measure faculty instructional goals among university faculty?

Via tests for "functioning", particularly by examining rating scale structure (see Figure 4.1) and investigating rating scale diagnostics (see Table 4.4), ample evidence suggested that the HERI Faculty Survey is both a valid and reliable instrument. These analyses revealed evidence indicating that the survey items were written clearly and all respondents interpreted the items similarly. Additionally, response options provided on the HERI Faculty Survey were determined to be both appropriate and sufficient.

Question 2: Do relevant items on the HERI Faculty Survey fit the expectations of the one-parameter IRT (Rasch) model by forming a unidimensional construct?

Testing for "fit", Rasch analyses determined the HERI Faculty Survey fit the expectations of the one-parameter IRT (Rasch) model by forming a unidimensional construct. Evidence of fit was provided by the examination of the summary statistics for the overall model (see Table 4.1), and the assessment of fit statistics for each item (see Table 4.2). An evaluation of potentially misfitting items revealed only two of 14 items were questionable. Upon investigating Item Characteristic Curves (ICCs) and the distribution of both misfitting items and persons, it appeared the items were sound, therefore the items were retained and included in subsequent analyses.

Question 3: How does a hierarchy of academic disciplines compare to previously established higher education classification systems?

Via DIF analyses, a hierarchy of academic disciplines was formed for each of the 14 items of interest (see Tables A4 – A17 for complete DIF results for each item). "A direct comparison was made between the DIF results and the Biglan, Kolb and Holland models. It was evident that various dimensions of the aforementioned models could be used to explain phenomena occurring between faculty from various academic disciplines and their perceptions of undergraduate instructional goals. Tables 4.6, 4.7, and 4.8 display the hierarchies formed for the three models as it relates to each of the 14 items on the survey.

Question 4: In what ways does presumed paradigmatic consensus influence faculty instructional goals?

The results of the DIF analyses provided clear evidence that faculty from various academic disciplines tend to exhibit varying degrees of concern for the 14 items of interest (see Tables A4 – A17 for complete details). In many instances, faculty from disciplines with shared characteristics exhibited similar preferences for undergraduate instructional goals. In particular, Biglan's 'hard' v. 'soft' dimensions could explain phenomena on 12 of 14 items, the pure v. applied dimensions could explain eight of 14 items, and the life v. nonlife dimension could explain six of 14 items. As for Kolb's model, which is essentially an extension of Biglan's, his active v. reflective dimensions could explain phenomena on 10 of 14 items, and the concrete v. abstract dimension could explain 11 of 14 items. With regard to Holland's model, all six dimensions were moderately helpful. The counts include: Social – 8 of 14; Artistic – 7 of 14; Realistic – 9 of 14; Investigative – 8 of 14; Conventional – 5 of 14; and Enterprising – 8 of 14.

Findings Relevant to Previous Research

First, this study found faculty from nearly every discipline were primarily concerned with the intellectual growth of students. This finding yields additional support to such studies as: Jervis and Congdon, 1958; Lawrence, Hart, Mackie, Muniz, and Dickmann, 1990; Liebert and Bayer, 1975; Platt, Parsons, & Kirshstein, 1976; Royal, Eli, and Bradley, 2005; and Wilson, Gaff, Dienst, Wood, and Bavry, 1975. A byproduct of this finding reveals additional support for both Liebert and Bayer's and Royal et. al's research, which found faculty were generally less concerned with moral and personal development than the intellectual growth of students.

Stark and Morstain (1978) found faculty from the natural sciences and professional fields were more concerned with "preparation for life and work" than faculty from other disciplines. While measuring a related item (Preparing students for employment after college), this study found a significant number of faculty from active and applied fields expressed strong interest in such employment preparation, thus providing some distant support for Stark and Morstain's findings. An additional caveat resulting from findings of this study, however, was that Education fields tend to be very concerned with this construct as well, emphasizing interest in preparing student for post-college employment.

Consistent with Smart's (1982) research on Holland's theory, this study also found that faculty from the Social and Artistic environments were more likely to be concerned with issues of personal development and character building. Also, consistent with Biglan (1973b) and Smart and Elton's (1975) research, this study found a good bit of evidence to support the notion that applied fields appear to be more service-oriented than pure fields, and life fields tend to be more service-oriented than non-life fields. Specific to this research, when investigating the item "Prepare students for responsible citizenship," a clear hierarchy was revealed for fields ranging from life to non-life and active to reflective disciplines, as illustrated by hierarchical location where most life and active disciplines appeared at the top of the hierarchy while few non-life and reflective disciplines appeared at the bottom. Faculty at the very top of the hierarchy also fit Holland's "social" type.

Other General Findings

A number of additional findings resulted from this study as well. These findings were based on a nationally representative sample of 7,356 faculty. This sample's responses to various demographic questions are presented below:

- 66% were male; 34% were female.
- 89.5% classified themselves as White/Caucasian.
- 84% considered teaching their principle activity.
- 34.7% held the rank of Professor; 27.8% held the rank of Associate Professor; and 24.2% held the rank of Assistant Professor.
- 58.5% of faculty respondents were tenured; 41.5% were untenured.
- 20.5% of faculty reported a primary interest in "heavily teaching"; 36.2% reported a primary interest "toward teaching". The remaining 43.3% reported a primary interest in research.
- 79.9% reported holding some form of a doctorate degree.
- With regard to political views, 17.7% reported being Conservative; 32.8% in the Middle of the Road; and 49.5% reported being Liberal.

Based on the results of the item map generated in Figure 4.2, overall faculty responses to the 14 survey items can be clustered into four separate regions based on a pattern of general consensus for undergraduate instructional goals. First, faculty were primarily concerned with the undergraduate instructional goal to "Develop ability to think critically". Second, faculty were also largely concerned with the following instructional goals: Employment after college; Responsible citizenship; Self-understanding; Preparing students for graduate education; Racial/ethnic appreciation; Moral character; and Personal values. Third, faculty were moderately concerned with the following: Emotional development; Enhancing students' out-of-class experiences; Community service; Foreign language; and Classic works of Western Civilization. Finally, faculty were least concerned with the instructional goal of "Family living".

With regard to hierarchies formed during the DIF analyses, a hierarchy consisting of 'soft' science disciplines appeared at the top for all 12 of the items which could be explained through Biglan's 'hard' v. 'soft' dimension. Biglan's applied disciplines were located near the top of the hierarchy on six of the eight relevant items for this dimension (pure v. applied), and life fields were located on the upper bound of the hierarchy on all six relevant items (see Table 4.6). In Kolb's model, active fields were found in the upper boundary on nine of 10 items, and abstract fields were found in the upper boundary of nine of 11 relevant items in these dimensions (see Table 4.7). Holland's model identified a number of interesting discipline placements on the hierarchy as well. All eight social disciplines were located in the upper boundary of the hierarchy, and five of seven artistic disciplines and two of five conventional fields fell in the upper boundary as well. Interestingly, all nine realistic fields, all eight investigative fields, and all eight enterprising fields fell along the lower boundary of its respective hierarchy (see Table 4.8).

Discussion

Discussion of the results focuses on issues pertaining to sample, underlying problems of survey research and classification systems. First, with regard to the sample, some key demographic variables may initially appear inflated, such as the two-thirds majority male respondents, 89.5% of faculty being of White/Caucasian ethnicity, and so on. As alarming as these statistics may seem, the resulting demographics are rather typical, as they are nationally representative of the university professorate. Given that the data utilized in this study came from a large, national dataset, one can infer the demographic results were valid and reliable for this particular sample.

It is also important to note that 56.6% of faculty respondents reported a primary interest in teaching as opposed to research, and 84% reported teaching as their principle activity. Although this study analyzed faculty perceptions of instructional goals by academic discipline, it is important to recognize the potential for multiple factors confounding any explanation of the results. For instance, previous research has suggested faculty from various disciplines often incorporate different instructional techniques in their courses based on the norms of the field. Braxton and Nordvall (1985), Gaff and Wilson (1971), Lattuca and Stark (1994), and Smart and Ethington (1995) found faculty in natural and physical sciences were more likely to require memorization and application, whereas faculty in the social and behavioral sciences and humanities were more likely to address critical thinking. These differences in instructional preferences may lead to different expectations for students, which may in turn have some bearing on the importance of various instructional goals that faculty set for their students. Other possible factors may include: the extent to which teaching is valued and rewarded at various institutions; the extent to which faculty take their teaching seriously; the amount of effort faculty exerted in investigating best practices and teaching literature; and the extent to which faculty are provided, and take advantage of, professional development opportunities.

Additional issues that could potentially cloud any results include factors such as class size, course level, and specifics amongst demographic items. Class size could have a profound impact on the way faculty perceive various instructional goals. Faculty in smaller classes may have more opportunities to reach students in deeper, more meaningful ways than faculty who are limited to lecturing large groups of students. Similarly, course level may have some bearing on these results as well. Faculty who teach introductory level courses may face a number of different dynamics and instructional issues than faculty who teach intermediate and advanced level courses within a discipline. Because this study sought to investigate faculty perceptions of instructional goals on a macro level, any microanalyses of data would have been overwhelming for this type of project. Therefore, the aforementioned factors were not controlled for in the analysis of these data. Further, this study did not isolate subsets and samples of demographic characteristics (e.g., gender, age, rank) according to disciplinary affiliation and compare responses about various instructional goals. However, future research should certainly investigate such issues and questions.

Another issue that would likely go undetected despite favorable response scale diagnostics and functioning pertains to the response scale used in the survey. In this study, the large sample of 7,356 respondents may mask a potentially critical fundamental survey research flaw. The response options "Essential", "Very Important", "Somewhat Important" and "Not Important" do not illustrate a balanced scale. The first three options tend to represent positive sentiments, whereas the last option represents a negative option. A better scale would provide two clearly positive statements and two clearly negative statements, ranging from extreme high to extreme low.

When reviewing the results of the hierarchical comparisons, every dimension of each model was useful to varying degrees in explaining phenomena. As a result, there were a number of instances in which items possessed some overlap, as more than one dimension or model was able to explain phenomena. Perhaps the most obvious explanation for this overlap is that some dimensions may be comprised of many of the same academic disciplines. This might suggest that although these constructs represent very different phenomena, some intertwining between constructs is inevitable due to the nature of academic disciplines. A closer examination of the academic disciplines that comprise each dimension of the various classification models might reveal additional similarities between dimensions. Having a better understanding of the relationships among the various dimensions of each model could shed some additional light on the utility of each model, as well as possibly open doors for new theories, dimensions, and classification schemas and models.

Overall, Kolb's model possessed the best explanatory power as it relates to faculty sentiments regarding the importance of various undergraduate instructional goals. This was largely due to the comprehensive nature of Kolb's model, as Kolb's model is essentially an extension of Biglan's model with two additional dimensions. Holland's model was able to explain the least of the three models examined in this study, however its contribution was still moderate. Despite these findings, research in the area of higher education classification systems warrants additional attention, particularly in the area of quality measurement. Due to the introduction of new disciplines and the evolution of existing disciplines over the last several decades, investigation of the changing landscape of disciplines and re-evaluation of how well present-day disciplines fit the criteria outlined by each model would serve very useful.

Contribution of the Study

With regard to contributions, perhaps methods are at the forefront. Related studies have largely been based on Classical Test Theory (CTT) and utilized basic descriptive statistics, regression and factor analyses. Arguably, Item Response Theory (IRT) provides a better and more comprehensive quantitative technique for studies of this nature. This is due to one-parameter IRT model's strict adherence to the established criteria for sound measurement (see Wright and Master, 1981; and Curtis and Keeves, 1999). Furthermore, IRT includes and investigates the items on the survey instrument through the "functioning" process, as opposed to statistical techniques that simply analyze data. Utilizing IRT can impact future as well as previous studies, as the possibilities for re-analyzing data from previous research may surface. This research should serve as a useful framework for outlining similar survey-related studies in the future.

Finally, this study also evaluated previously established higher education classification systems, as results were compared to previous models. The evaluation could open doors for concurrent discussions regarding classification systems in today's higher education environments, particularly as it relates to understanding differences between persons and phenomena from various academic disciplines.

Implications and Future Research

As Smart (2005) noted, "The lack of attention to measurement issues is one of the major deficiencies in the higher education research literature" (p. 470). This study addressed the problem of measurement in higher education research and discusses competing methodological approaches, particularly CTT versus IRT. This study argued that an IRT approach may be more appropriate and the results more precise than a CTT approach. This study then offered a demonstration of how to use an IRT technique, namely Rasch measurement, to analyze data. As a result of the demonstration, this study serves as a model for related and future studies which utilize and evaluate the quality of survey research.

Another implication of this study is to challenge other researchers to further explore issues of measurement within their own research. Rasch measurement is not intended to take the place of statistics, but rather to complement the use of statistics. Utilizing a theoretically-sound and mathematically-just approach like Rasch measurement eliminates many assumptions researchers often make regarding methodological issues. Therefore, once proper measurement takes place, statistical analyses can then be applied and the results will become more precise, and possibly more meaningful.

This study could also benefit Offices of Assessment and Institutional Effectiveness. Accrediting agencies require institutions to capture data regarding nearly every facet of an institution, especially student learning outcomes. Understanding how faculty from various disciplines perceive a wide range of instructional goals could help deans and department chairs better strategize discipline-specific learning outcomes, in addition to those outcomes that are more general in nature. Further, understanding such differences in perceptions could help key administrators/leaders predict how difficult each goal will be to accomplish for their given departments.

Further, this research brings to light the importance of understanding the differences between academic disciplines, as a "one size fits all" approach is not necessarily the best approach in practice. This is especially true when comparisons are made around the board, as is often the case with Teacher Course Evaluations (TCEs). As fundamental measurement suggests that items have varying degrees of difficulty, importance and relevance, we must also remember that these item differences will only become more problematic when applied to academic disciplines. Understanding how faculty perceive various instructional goals could assist in tailoring TCEs specific to disciplines, thus generating a more valid and meaningful evaluation.

Finally, this research provides implications regarding higher education classification systems. First, new research on established models is needed. Much of the research on higher education classification systems was conducted in the 1970s and 1980s. With the changes that have occurred in higher education since this time, it is important that researchers revisit this literature and investigate to what extent previously established models, theories and classification systems still hold true in today's higher education landscape. Further, with the advancement in various methodological techniques and inquiries, it is important that researchers develop new and innovative ways to challenge and test what we already know and that which we do not. With the addition of IRT techniques to a researcher's methodological skill set, the possibilities for new and innovative studies are virtually unlimited.

267253

2001 Faculty Survey

Higher Education Research Institute, UCLA

MARKING DIRECTIONS

Your responses will be read by an optical mark reader. Please,

- Use a pencil or black or blue pen.
- Fill the oval completely.
- Erase cleanly any marks you wish to change or "X" out mark if in pen.

CORRECT MARK ● INCORRECT MARKS

1. What is your principal activity in your current position at this institution? (Mark one)
- Administration
- Teaching
- Research
- Services to clients and patients
- Other

2. Are you considered a full-time employee of your institution for at least nine months of the current academic year? (Mark one)
- Yes
- No

3. Your sex:
- Male
- Female

4. What is your present academic rank?
- Professor
- Associate Professor
- Assistant Professor
- Lecturer
- Instructor
- Other

5. What is your administrative title?
- Not applicable
- Director, coordinator, or administrator of an institute, center, lab, or specially-funded program
- Department Chair
- Dean
- Associate or Assistant Dean
- Vice-President, Provost, Vice-Chancellor
- President, Chancellor
- Other

6. Are you currently: (Mark one)
- Married
- Unmarried, living with partner
- Single

7. Have you ever been: (Mark all that apply)
- Divorced
- Widowed
- Separated

8. If you were to begin your career again, would you still want to be a college professor?
- Definitely yes
- Probably yes
- Not sure
- Probably no
- Definitely no

9. Racial/Ethnic group: (Mark all that apply)
- White/Caucasian
- African American/Black
- American Indian
- Asian American/Asian
- Mexican American/Chicano
- Puerto Rican American
- Other Latino
- Other

10. How many children do you have in the following age ranges?

0 - 4 years old	0 1 2 3 4+
5 - 12 years old	0 1 2 3 4+
13 - 17 years old	0 1 2 3 4+
18 - 23 years old	0 1 2 3 4+
24 years old or older	0 1 2 3 4+

11. ~~Do your interests lie primarily in teaching~~ or research?
- Very heavily in research
- In both, but leaning toward research
- In both, but leaning toward teaching
- Very heavily in teaching

12. On the following list, please mark one in each column:

	Highest Degree Earned	Degree Currently Working On
Bachelor's (B.A., B.S., etc.)	○	○
Master's (M.A., M.S., etc.)	○	○
LL.B., J.D.	○	○
M.D., D.D.S. (or equivalent)	○	○
Other first professional degree beyond B.A. (e.g., D.D., D.V.M.)	○	○
Ed.D.	○	○
Ph.D.	○	○
Other degree	○	○
None	○	○

13. During the past two years, have you engaged in any of the following activities?

(Mark one for each item)

	Yes	No
Taught an honors course	Y	N
Taught an interdisciplinary course	Y	N
Taught an ethnic studies course	Y	N
Taught a women's studies course	Y	N
Team-taught a course	Y	N
Taught a service learning course	Y	N
Worked with undergraduates on a research project	Y	N
Used intra- or extramural funds for research	Y	N
Participated in a teaching enhancement workshop	Y	N
Placed or collected assignments for a course on the Internet	Y	N
Taught a course exclusively through the Internet	Y	N

14. In the two sets of ovals shown below, please mark the most appropriate code from the fields listed on the back of the accompanying letter. (Please see example on back of accompanying letter.)

- Major of highest degree held
- Department of current faculty appointment

15. In the set of ovals to the right, please mark the dollar value of your base institutional salary, rounded to the nearest $1,000. (Note: Amounts above $199,000 should be marked "199")

The above salary is based on:
- 9/10 months
- 11/12 months

16. In the four sets of ovals below, please mark the last two digits of the year of each of the following:

- Year of birth
- Year of highest degree now held
- Year of appointment at present institution
- Year tenure was awarded

Are you tenured?
- Yes
- No

NOTE: If you are between terms, on leave, or in an interim term, please answer questions 17 and 18 as they apply to the full term most recently completed at this institution.

17. During the present term, how many hours per week on the average do you actually spend on each of the following activities?

(Mark one for each activity)

	Hours Per Week								
	None	1-4	5-8	9-12	13-16	17-20	21-34	35-44	45+
Scheduled teaching (give actual, not credit hours)	○	○	○	○	○	○	○	○	○
Preparing for teaching (including reading student papers and grading)	○	○	○	○	○	○	○	○	○
Advising and counseling of students	○	○	○	○	○	○	○	○	○
Committee work and meetings	○	○	○	○	○	○	○	○	○
Other administration	○	○	○	○	○	○	○	○	○
Research and scholarly writing	○	○	○	○	○	○	○	○	○
Other creative products/performances	○	○	○	○	○	○	○	○	○
Consultation with clients/patients	○	○	○	○	○	○	○	○	○
Community or public service	○	○	○	○	○	○	○	○	○
Outside consulting/freelance work	○	○	○	○	○	○	○	○	○
Household/childcare duties	○	○	○	○	○	○	○	○	○

18. How many of the following courses are you teaching this term?

(Mark one for each activity)

General education courses	0	1	2	3	4	5+
Other BA or BS undergraduate credit courses	0	1	2	3	4	5+
Non-BA credit courses (developmental/remedial)	0	1	2	3	4	5+
Graduate courses	0	1	2	3	4	5+

19. Indicate the importance to you of each of the following education goals for undergraduate students:

(Mark one for each item)

	Essential	Very Important	Somewhat Important	Not Important
Develop ability to think clearly	E	V	S	N
Prepare students for employment after college	E	V	S	N
Prepare students for graduate or advanced education	E	V	S	N
Develop moral character	E	V	S	N
Provide for students' emotional development	E	V	S	N
Prepare students for family living	E	V	S	N
Teach students the classic works of Western civilization	E	V	S	N
Help students develop personal values	E	V	S	N
Enhance the out-of-class experience of students	E	V	S	N
Enhance students' self-understanding	E	V	S	N
Instill in students a commitment to community service	E	V	S	N
Prepare students for responsible citizenship	E	V	S	N
Enhance students' knowledge of and appreciation for other racial/ethnic groups	E	V	S	N
Study a foreign language	E	V	S	N

20. How influential were the following people in your decision to pursue an academic career?

(Mark one for each item)

	Very Influential	Somewhat Influential	Not Influential
Father	V	S	N
Mother	V	S	N
Other relatives	V	S	N
Undergraduate faculty or advisor	V	S	N
Graduate faculty or advisor	V	S	N

For questions 21-23, mark only one response for each question.

	None	1-2	3-4	5-10	11-20	21-50	51+
21. How many of the following have you published?							
Articles in academic or professional journals	○	○	○	○	○	○	○
Chapters in edited volumes	○	○	○	○	○	○	○
Books, manuals, or monographs	○	○	○	○	○	○	○
22. How many exhibitions or performances in the fine or applied arts have you presented?	○	○	○	○	○	○	○
23. How many of your professional writings have been published or accepted for publication in the last two years?	○	○	○	○	○	○	○

24. For each of the following items, please mark either Yes or No:

	Yes	No
Have you ever held an academic administrative post?	Y	N
Have you ever received an award for outstanding teaching?	Y	N
Do you commute a long distance to work?	Y	N
Has any of your research or writing focused on women?	Y	N
Does your spouse/partner work in the same city?	Y	N
Is your spouse/partner an academic?	Y	N
Has any of your research or writing focused on racial or ethnic minorities?	Y	N
Were you born in the USA?	Y	N
Are you a U.S. citizen?	Y	N
Have you ever interrupted your professional career for more than one year for family reasons?	Y	N
Have you been sexually harassed at this institution?	Y	N
Do you plan on working beyond age 70?	Y	N
Are you a member of a faculty union?	Y	N
Is (or was) your father an academic?	Y	N
Is (or was) your mother an academic?	Y	N
During the Last Two Years, Have You:		
Received at least one firm job offer?	Y	N
Developed a new course?	Y	N
Considered early retirement?	Y	N
Considered leaving academe for another job?	Y	N
Taught courses at more than one institution during the same term?	Y	N
Served as a paid consultant?	Y	N
Requested/sought an early promotion?	Y	N

25. How important were each of the following in your decision to work at this college or university?

(Mark one for each item)

	Very Important	Somewhat Important	Not Important	A Deterrent
Institutional emphasis on teaching	V	S	N	D
Institutional emphasis on research	V	S	N	D
Prestige of institution	V	S	N	D
Prestige of department	V	S	N	D
Salary/benefits	V	S	N	D
Research facilities	V	S	N	D
Academic rank offered	V	S	N	D
Colleagues	V	S	N	D
Geographic location	V	S	N	D
Job opportunities for spouse	V	S	N	D
Other personal/family considerations	V	S	N	D

26. Indicate how important you believe each priority listed below is at your college or university:

(Mark one for each activity)

	Highest Priority	High Priority	Medium Priority	Low Priority
To promote the intellectual development of students	4	3	2	1
To help students examine and understand their personal values	4	3	2	1
To develop a sense of community among students and faculty	4	3	2	1
To develop leadership ability among students	4	3	2	1
To facilitate student involvement in community service	4	3	2	1
To help students learn how to bring about change in American society	4	3	2	1
To increase or maintain institutional prestige	4	3	2	1
To hire faculty "stars"	4	3	2	1
To recruit more minority students	4	3	2	1
To enhance the institution's national image	4	3	2	1
To create a diverse multi-cultural campus environment	4	3	2	1
To promote the religious/spiritual development of students	4	3	2	1
To mentor new faculty	4	3	2	1

27. Below are some statements about your college or university. Indicate the extent to which you agree or disagree with each of the following:

(Mark one for each item)

	Agree Strongly	Agree Somewhat	Disagree Somewhat	Disagree Strongly
Faculty are interested in students' personal problems	4	3	2	1
Racial and ethnic diversity should be more strongly reflected in the curriculum	4	3	2	1
Faculty feel that most students are well-prepared academically	4	3	2	1
This institution should hire more faculty of color	4	3	2	1
Student Affairs staff have the support and respect of faculty	4	3	2	1
Faculty are committed to the welfare of this institution	4	3	2	1
Faculty here are strongly interested in the academic problems of undergraduates	4	3	2	1
There is a lot of campus racial conflict here	4	3	2	1
Many courses include feminist perspectives	4	3	2	1
Faculty of color are treated fairly here	4	3	2	1
Women faculty are treated fairly here	4	3	2	1
Many courses involve students in community service	4	3	2	1
This institution should hire more women faculty	4	3	2	1
Most students are strongly committed to community service	4	3	2	1
Gay and lesbian faculty are treated fairly here	4	3	2	1
My research is valued by faculty in my department	4	3	2	1
My teaching is valued by faculty in my department	4	3	2	1

28. During the past two years, how involved have you been in efforts to reform the following at your institution?

(Mark one for each item)

	Very Involved	Minimally Involved	Not Involved
Overall mission, purpose	V	M	N
General education	V	M	N
Faculty roles/rewards	V	M	N
Governance	V	M	N
Curriculum	V	M	N

29. How important are each of the following in your decision to pursue an academic career?

(Mark one for each item)

	Very Important	Somewhat Important	Not Important
Autonomy	V	S	N
Flexible schedule	V	S	N
Intellectual challenge	V	S	N
Intellectual freedom	V	S	N
Freedom to pursue my scholarly/teaching interests	V	S	N
Opportunities for teaching	V	S	N
Opportunities for research	V	S	N
Occupational prestige/professional status	V	S	N
Opportunity to influence social change	V	S	N

30. Please indicate the extent to which each of the following has been a source of stress for you during the last two years:

(Mark one for each item)

	Extensive	Somewhat	Not At All
Managing household responsibilities	E	S	N
Child care	E	S	N
Care of elderly parent	E	S	N
My physical health	E	S	N
Review/promotion process	E	S	N
Subtle discrimination (e.g., prejudice, racism, sexism)	E	S	N
Personal finances	E	S	N
Committee work	E	S	N
Faculty meetings	E	S	N
Colleagues	E	S	N
Students	E	S	N
Research or publishing demands	E	S	N
Institutional procedures and "red tape"	E	S	N
Teaching load	E	S	N
Children's problems	E	S	N
Marital friction	E	S	N
Time pressures	E	S	N
Lack of personal time	E	S	N
Keeping up with information technology	E	S	N

31. How satisfied are you with the following aspects of your job?

(Mark one for each item)

	Very Satisfied	Satisfied	Marginally Satisfied	Not Satisfied	Not Applicable
Salary and fringe benefits	V	S	M	N	O
Opportunity for scholarly pursuits	V	S	M	N	O
Teaching load	V	S	M	N	O
Quality of students	V	S	M	N	O
Office/lab space	V	S	M	N	O
Autonomy and independence	V	S	M	N	O
Professional relationships with other faculty	V	S	M	N	O
Social relationships with other faculty	V	S	M	N	O
Competency of colleagues	V	S	M	N	O
Visibility for jobs at other institutions/organizations	V	S	M	N	O
Job security	V	S	M	N	O
Relationships with administration	V	S	M	N	O
Overall job satisfaction	V	S	M	N	O
Opportunity to develop new ideas	V	S	M	N	O
Availability of child care at this institution	V	S	M	N	O

32. Indicate how well each of the following describes your college or university:

(Mark one for each item)

	Very Descriptive	Somewhat Descriptive	Not Descriptive
It is easy for students to see faculty outside of regular office hours	V	S	N
There is a great deal of conformity among the students	V	S	N
The faculty are typically at odds with campus administrators	V	S	N
Faculty here respect each other	V	S	N
Most students are treated like "numbers in a book"	V	S	N
Social activities are overemphasized	V	S	N
Students here do not usually socialize with one another	V	S	N
Faculty are rewarded for being good teachers	V	S	N

33. In how many of the undergraduate courses that you teach do you use each of the following?

(Mark one for each item)

	All	Most	Some	None
Evaluation Methods:				
Multiple-choice mid-term and/or final exams	A	M	S	N
Essay mid-term and/or final exams	A	M	S	N
Short-answer mid-term and/or final exams	A	M	S	N
Quizzes	A	M	S	N
Weekly essay assignments	A	M	S	N
Student presentations	A	M	S	N
Term/research papers	A	M	S	N
Student evaluations of each others' work	A	M	S	N
Grading on a curve	A	M	S	N
Competency-based grading	A	M	S	N
Instructional Techniques/Methods:				
Class discussions	A	M	S	N
Computer or machine-aided instruction	A	M	S	N
Cooperative learning (small groups)	A	M	S	N
Experiential learning/Field studies	A	M	S	N
Teaching assistants	A	M	S	N
Recitals/Demonstrations	A	M	S	N
Group projects	A	M	S	N
Independent projects	A	M	S	N
Extensive lecturing	A	M	S	N
Multiple drafts of written work	A	M	S	N
Readings on racial and ethnic issues	A	M	S	N
Readings on women and gender issues	A	M	S	N
Student-developed activities (assignments, exams, etc.)	A	M	S	N
Student-selected topics for course content	A	M	S	N
Community service as part of coursework	A	M	S	N

34. What is the highest level of education reached by your spouse/partner and your parents?

(Mark one in each column)

	Spouse/Partner	Father	Mother
8th grade or less	S	F	M
Some high school	S	F	M
Completed high school	S	F	M
Some college	S	F	M
Graduated from college	S	F	M
Attended graduate or professional school	S	F	M
Attained advanced degree	S	F	M
Does not apply (No spouse or partner)	S		

DO NOT MARK IN THIS AREA

35. Please indicate your agreement with each of the following statements:

(Mark one for each item)

	Agree Strongly	Agree Somewhat	Disagree Somewhat	Disagree Strongly
Western civilization and culture should be the foundation of the undergraduate curriculum	4	3	2	1
College officials have the right to ban persons with extreme views from speaking on campus	4	3	2	1
The chief benefit of a college education is that it increases one's earning power	4	3	2	1
Promoting diversity leads to the admission of too many underprepared students	4	3	2	1
Colleges should be actively involved in solving social problems	4	3	2	1
Tenure is an outmoded concept	4	3	2	1
Colleges should encourage students to be involved in community service activities	4	3	2	1
Community service should be given weight in college admissions decisions	4	3	2	1
Tenure is essential to attract the best minds to academe	4	3	2	1
A racially/ethnically diverse student body enhances the educational experience of all students	4	3	2	1
External pressures often prevent researchers from being completely objective in the conduct of their work	4	3	2	1

36. How would you characterize your political views? (Mark one)

- ◯ Far Left
- ◯ Liberal
- ◯ Middle-of-the-road
- ◯ Conservative
- ◯ Far Right

37. Indicate the importance to you personally of each of the following:

(Mark one for each item)

	Essential	Very Important	Somewhat Important	Not Important
Becoming an authority in my field	E	V	S	N
Influencing the political structure	E	V	S	N
Influencing social values	E	V	S	N
Raising a family	E	V	S	N
Being very well-off financially	E	V	S	N
Helping others who are in difficulty	E	V	S	N
Becoming involved in programs to clean up the environment	E	V	S	N
Developing a meaningful philosophy of life	E	V	S	N
Helping to promote racial understanding	E	V	S	N
Obtaining recognition from my colleagues for contributions to my special field	E	V	S	N
Integrating spirituality into my life	E	V	S	N
Being a good colleague	E	V	S	N
Being a good teacher	E	V	S	N
Achieving congruence between my own values and institutional values	E	V	S	N

ADDITIONAL QUESTIONS: If you received additional questions, mark answers below:

38. A B C D E — 45. A B C D E — 52. A B C D E
39. A B C D E — 46. A B C D E — 53. A B C D E
40. A B C D E — 47. A B C D E — 54. A B C D E
41. A B C D E — 48. A B C D E — 55. A B C D E
42. A B C D E — 49. A B C D E — 56. A B C D E
43. A B C D E — 50. A B C D E — 57. A B C D E
44. A B C D E — 51. A B C D E — 58. A B C D E

Please return your completed questionnaire in the postage-paid envelope to:
Higher Education Research Institute
2905 West Service Road, Eagan, MN 55121

THANK YOU!

R17430-HERI/QDS/10464C-03-54321

Table A1

Personal Demographic Characteristics of Respondents (N = 7356)

	Characteristic	n	%
Sex			
	Male	4854	66.0
	Female	2502	34.0
Age			
	<30	126	1.7
	30-34	515	7.0
	35-39	800	10.9
	40-44	924	12.6
	45-49	1121	15.3
	50-54	1233	16.9
	55-59	1185	16.2
	60-64	922	12.6
	65-69	329	4.5
	70+	151	2.1
Race			
	White/Caucasian	6581	89.5
	African American/Black	156	2.1
	American Indian	87	1.2
	Asian American/Asian	312	4.2
	Mexican American/Chicano	70	1.0
	Puerto Rican American	21	0.3
	Other Latino	125	1.7

Table A2

Professional Demographic Characteristics of Respondents (n = 7356)

Characteristic	n	%
Employment Status		
Full-Time	6938	94.5
Part-Time	400	5.5
Principle Activity		
Administration	260	3.6
Teaching	6084	84.0
Research	821	11.3
Client/Patients	55	0.8
Other	27	0.4
Academic Rank		
Professor	2546	34.7
Associate Professor	2036	27.8
Assistant Professor	1774	24.2
Lecturer	400	5.5
Instructor	429	5.8
Other	151	2.1
Tenure Status		
Tenured	4050	58.5
Untenured	2868	41.5
Primary Interest		
Heavily Teaching	1501	20.6
Toward Teaching	2635	36.2
Toward Research	2672	36.7
Heavily Research	480	6.6
Type of Degree Earned		
Bachelor's (B.A., B.S., etc.)	70	1.0
Master's (M.A., M.S., etc.)	1088	14.8
LL.B., J.D.	60	0.8
M.D., D.D.S., (or equivalent)	71	1.0
Other first professional degree	64	0.9

beyond B.A. (e.g., D.D., D.V.M.)		
Ed.D.	138	1.9
Ph.D.	5520	75.3
Other degree	261	3.6
None	59	0.8
Political Views		
Far Right	20	0.3
Conservative	1224	17.4
Middle of Road	2315	32.8
Liberal	3103	44.0
Far Left	388	5.5

Table A3

Frequency of Academic Disciplines Reported by the Selected Sample (N = 7356)

Major of Degree Earned	n	%
Agriculture	170	2.3
Architecture	93	1.3
Bacteriology/Microbiology	98	1.3
Biochemistry	62	0.8
Botany	72	1.0
Physiology	76	1.0
Zoology	105	1.4
Biology	142	1.9
Accounting	120	1.6
Finance	72	1.0
International Business	5	0.1
Marketing	112	1.5
Management	148	2.0
Business	61	0.8
Computer Science	109	1.5
Elementary Education	59	0.8
Education Administration	84	1.1
Educational Psychology and Counseling	72	1.0
Music/Art Education	53	0.7
Physical and Health Education	132	1.8
Secondary Education	77	1.0
Special Education	58	0.8
Aeronautical/Astronautical Engineering	21	0.3
Chemical Engineering	57	0.8
Civil Engineering	97	1.3
Electrical Engineering	118	1.6
Mechanical Engineering	104	1.4
Nuclear Engineering	6	0.1

Art	211	2.9
Speech	173	2.4
Music	319	4.3
Geography	78	1.1
Medicine	61	0.8
Nursing	184	2.5
Pharmacy	84	1.1
Home Economics	44	0.6
English	541	7.4
Foreign Languages	122	1.7
French	86	1.2
German	61	0.8
Spanish	120	1.6
History	387	5.3
Philosophy	197	2.7
Religion	154	2.1
Journalism	77	1.0
Law	73	1.0
Law Enforcement/Administration of Justice	8	0.1
Library Science	40	0.5
Math and/or Statistics	458	6.2
Astronomy	17	0.2
Chemistry	233	3.2
Marine Science	17	0.2
Physics	180	2.4
Psychology	78	1.1
Anthropology	151	2.1
Economics	233	3.2
Political Science	268	3.6
Sociology	259	3.5
Social Work	59	0.8

Table A4

DIF for "Develop Ability to Think Clearly".

Academic Discipline	Count	Average Observations
International Business	5	3.00
Elementary Education	55	2.95
Chemical Engineering	57	2.95
Philosophy	194	2.95
Nursing	181	2.93
Spanish	120	2.93
History	385	2.93
Architecture	93	2.92
English and Literature	532	2.92
Religion and Theology	153	2.92
Library Science	39	2.92
Chemistry	226	2.92
Biology	140	2.91
Journalism	76	2.91
Botany	72	2.90
Secondary Education	77	2.90
Aeronautical/Astronauticaul Engineering	21	2.90
French	86	2.90
Bacteriology/Microbiology	97	2.89
Physiology	75	2.89
Dramatics and Speech	169	2.89
Political Science	265	2.89
Foreign Languages and Literature	120	2.88
Law Enforcement/Administration of Justice	8	2.88
Astronomy	17	2.88
Marine Science	17	2.88
Psychology	76	2.88
Economics	232	2.88
Zoology	104	2.87
Accounting	115	2.87
Music and Art Education	53	2.87
Art	210	2.87
Pharmacy	84	2.87
German	60	2.87
Anthropology	151	2.87
Educational Psychology & Counseling	70	2.86
Music	314	2.86
Home Economics	42	2.86
Law	72	2.86
Math and Statistics	451	2.86
Physics	177	2.86
Sociology	256	2.86

Electrical Engineering	117	2.85
Management	147	2.84
Marketing	108	2.83
Computer Science	109	2.83
Civil Eng	95	2.83
Nuclear Engineering	6	2.83
Social Work	59	2.83
Education Administration	83	2.82
Geography	78	2.81
Mechanical Engineering	104	2.80
Finance	72	2.79
Agriculture	167	2.77
Special Education	58	2.76
Medicine	54	2.76
Business	59	2.75
Biochemistry	62	2.74
Physical and Health Education	131	2.74

Table A5

DIF for "Prepare students for employment after college".

Academic Discipline	Count	Average Observations
Nursing	180	2.51
Special Education	58	2.47
Elementary Education	55	2.36
Secondary Education	77	2.36
Nuclear Engineering	6	2.33
Journalism	76	2.33
Chemical Engineering	57	2.30
Civil Engineering	95	2.29
Home Economics	42	2.29
Accounting	115	2.28
Electrical Engineering	117	2.26
Pharmacy	84	2.26
Mechanical Engineering	104	2.25
Agriculture	167	2.24
Education Administration	82	2.24
Aeronautical/Astronautical Engineering	21	2.24
Physical and Health Education	131	2.21
Business	59	2.20
Finance	72	2.17
Marketing	109	2.17
Social Work	59	2.15
Educational Psychology & Counseling	69	2.14
Management	147	2.13
Music and Art Education	53	2.09
Library Science	39	2.08
Computer Science	108	2.02
Medicine	54	2.02
International Business	5	2.00
Music	315	1.93
Architecture	93	1.92
Law Enforcement	8	1.88
Chemistry	227	1.86
Dramatics and Speech	168	1.85
Geography	78	1.85
Bacteriology/Microbiology	96	1.84
Physiology	75	1.84
Physics	176	1.80
Biology	140	1.79
Math and Statistics	448	1.77
Spanish	118	1.75
Economics	232	1.75
Law	72	1.71

Astronomy	17	1.71
Marine Science	17	1.71
Biochemistry	61	1.69
Botany	72	1.67
Psychology	75	1.65
Sociology	254	1.57
French	85	1.54
Zoology	103	1.53
Art	209	1.49
English and Literature	532	1.48
Political Science	265	1.46
Anthropology	151	1.42
History	383	1.39
Foreign Language and Literature	120	1.38
German	60	1.38
Philosophy	195	1.29
Religion and Theology	152	1.16

Table A6

DIF for "Prepare students for graduate or advanced education".

Academic Discipline	Count	Average Observations
Medicine	55	2.02
Music	314	2.00
Nursing	181	1.96
Physiology	75	1.95
Biochemistry	62	1.94
Social Work	59	1.93
Zoology	104	1.92
Bacteriology/Microbiology	97	1.90
Electrical Engineering	117	1.84
Biology	140	1.84
Nuclear Engineering	6	1.83
Spanish	119	1.82
Psychology	76	1.82
Chemistry	228	1.81
Botany	72	1.81
Aeronautical/Astronautical Engineering	21	1.81
Physics	177	1.80
Educational Psychology & Counseling	69	1.80
Art	209	1.80
Civil Engineering	95	1.78
Elementary Education	55	1.76
Anthropology	151	1.73
Physical and Health Education	130	1.70
German	60	1.70
Chemical Engineering	57	1.70
Dramatics and Speech	168	1.69
Mechanical Engineering	104	1.68
Math and Statistics	448	1.68
Foreign Language and Literature	120	1.67
Music and Art Education	53	1.64
Geography	77	1.64
Computer Science	108	1.64
Accounting	115	1.64
French	86	1.63
Architecture	93	1.63
Pharmacy	84	1.62
International Business	5	1.60
Finance	72	1.60
Special Education	58	1.59
Secondary Education	76	1.59
Astronomy	17	1.59
Education Administration	83	1.58

Sociology	255	1.56
Political Science	265	1.55
Home Economics	42	1.55
Philosophy	195	1.54
Library Science	39	1.54
Marine Science	17	1.53
Economics	231	1.53
Business	60	1.53
Religion and Theology	153	1.52
English and Literature	530	1.50
Agriculture	166	1.46
History	383	1.45
Management	147	1.43
Journalism	76	1.38
Marketing	109	1.35
Law	72	1.32
Law Enforcement	8	1.25

Table A7

DIF for "Develop moral character".

Academic Discipline	Count	Average Observations
Medicine	54	2.39
Law Enforcement	8	2.38
Elementary Education	55	2.36
Nursing	181	2.30
Education Administration	83	2.19
Home Economics	42	2.19
Educational Psychology & Counseling	69	2.12
Religion and Theology	153	2.07
Physical and Health Education	130	2.04
Secondary Education	76	2.00
Journalism	76	1.99
Pharmacy	84	1.98
Agriculture	166	1.95
Accounting	115	1.95
Management	147	1.93
Dramatics and Speech	169	1.93
Music	312	1.93
Special Education	58	1.91
Chemical Engineering	57	1.89
Art	210	1.87
Architecture	93	1.86
Civil Engineering	95	1.85
Spanish	119	1.85
Business	59	1.83
Law	72	1.83
Finance	72	1.81
Philosophy	194	1.79
Library Science	39	1.79
Social Work	59	1.75
Aeronautical/Astronautical Engineering	19	1.74
Music and Art Education	52	1.73
Physiology	75	1.72
Marketing	108	1.72
Electrical Engineering	117	1.68
Nuclear Engineering	6	1.67
Mechanical Engineering	104	1.65
Bacteriology/Microbiology	97	1.63
English and Literature	527	1.62
Computer Science	107	1.61
International Business	5	1.60
Foreign Language and Literature	121	1.60
Chemistry	228	1.59

Marine Science	17	1.59
Biochemistry	62	1.58
Biology	139	1.56
Geography	78	1.56
Botany	72	1.54
Zoology	104	1.54
French	83	1.52
History	383	1.50
Physics	175	1.50
German	60	1.48
Sociology	255	1.45
Anthropology	148	1.44
Political Science	263	1.44
Math and Statistics	447	1.43
Psychology	76	1.42
Economics	231	1.28
Astronomy	17	1.24

Table A8

DIF for "Provide for students' emotional development".

Academic Discipline	Count	Average Observations
Medicine	53	1.87
Nursing	180	1.81
Elementary Education	55	1.78
Educational Psychology & Counseling	70	1.76
Home Economics	42	1.71
Education Administration	83	1.69
Physical and Health Education	131	1.69
Social Work	59	1.69
Music	312	1.66
Secondary Education	76	1.64
Dramatics and Speech	169	1.63
Music and Art Education	53	1.57
Art	210	1.57
Special Education	58	1.53
Religion and Theology	153	1.45
Spanish	118	1.42
Law Enforcement	8	1.38
Library Science	39	1.38
Nuclear Engineering	6	1.33
Journalism	75	1.33
English and Literature	527	1.31
Agriculture	167	1.29
Architecture	92	1.29
Pharmacy	84	1.29
Management	147	1.28
Law	72	1.28
Physiology	75	1.24
Foreign Language and Literature	120	1.24
German	60	1.22
Philosophy	192	1.21
Biology	140	1.20
International Business	5	1.20
Accounting	114	1.18
Psychology	76	1.17
Anthropology	149	1.17
Finance	72	1.15
French	83	1.14
Business	60	1.13
Chemistry	228	1.13
Chemical Engineering	57	1.11
Physics	175	1.10
Zoology	104	1.09

Marine Science	17	1.06
Aeronautical/Astronautical Engineering	19	1.05
Geography	78	1.05
Bacteriology/Microbiology	96	1.04
Civil Engineering	95	1.04
Botany	72	1.03
Computer Science	107	1.03
Mechanical Engineering	104	1.03
Sociology	254	1.03
Marketing	107	1.00
History	384	0.98
Math and Statistics	447	0.98
Biochemistry	61	0.97
Electrical Engineering	117	0.97
Political Science	264	0.89
Astronomy	17	0.82
Economics	231	0.81

Table A9

*DIF for "*Prepare students for family living*".*

Academic Discipline	Count	Average Observations
Home Economics	42	1.90
Elementary Education	55	1.25
Education Administration	83	1.23
Educational Psychology & Counseling	70	1.21
Medicine	54	1.17
Law Enforcement	8	1.13
Physical and Health Education	131	1.10
Secondary Education	76	1.08
Agriculture	165	1.02
Nursing	180	1.02
Special Education	58	1.00
Social Work	59	0.92
Religion and Theology	152	0.90
Psychology	75	0.85
Library Science	38	0.84
Philosophy	192	0.81
Accounting	115	0.80
Music and Art Education	53	0.79
Aeronautical/Astronautical Engineering	19	0.79
Music	313	0.77
Sociology	254	0.77
Management	146	0.75
Spanish	119	0.74
Journalism	76	0.74
Bacteriology/Microbiology	96	0.73
Dramatics and Speech	168	0.71
Pharmacy	83	0.71
English and Literature	525	0.68
Law	72	0.68
Geography	78	0.67
Biology	138	0.64
Business	59	0.63
Foreign Language and Literature	121	0.62
Biochemistry	61	0.61
Botany	71	0.61
Finance	72	0.61
Computer Science	107	0.60
Anthropology	150	0.59
Chemistry	226	0.58
Math and Statistics	443	0.57
Zoology	104	0.56
Physics	172	0.56

Art	209	0.55
French	85	0.55
German	60	0.55
History	378	0.54
Mechanical Engineering	104	0.53
Marketing	107	0.52
Civil Engineering	95	0.52
Economics	229	0.52
Chemical Engineering	57	0.51
Electrical Engineering	116	0.50
Nuclear Engineering	6	0.50
Physiology	75	0.49
Architecture	93	0.46
Political Science	263	0.45
Astronomy	17	0.41
International Business	5	0.40
Marine Science	17	0.29

Table A10

DIF for "Teach students the classics of Western civilization".

Academic Discipline	Count	Average Observations
Philosophy	195	1.87
Music	312	1.81
French	84	1.75
German	60	1.72
Religion and Theology	153	1.66
Foreign Language and Literature	120	1.63
Dramatics and Speech	169	1.62
Spanish	118	1.54
English and Literature	531	1.53
Art	210	1.38
History	385	1.35
Law Enforcement	8	1.25
Architecture	93	1.24
Astronomy	17	1.24
Library Science	39	1.21
Political Science	265	1.17
Music and Art Education	52	1.15
Journalism	76	1.08
Education Administration	83	1.07
Educational Psychology & Counseling	69	1.07
Geography	78	0.94
Marine Science	17	0.94
Physical and Health Education	129	0.93
Math and Statistics	441	0.93
Physics	172	0.93
Social Work	59	0.93
Law	71	0.92
Chemistry	224	0.92
Psychology	76	0.89
Sociology	255	0.89
Secondary Education	76	0.88
Economics	232	0.86
Elementary Education	55	0.85
Anthropology	151	0.85
Aeronautical/Astronautical Engineering	19	0.84
Nuclear Engineering	6	0.83
Zoology	104	0.82
Home Economics	42	0.81
Botany	71	0.80
Special Education	58	0.76
Medicine	52	0.75
Bacteriology/Microbiology	97	0.74

Nursing	179	0.74
Biology	139	0.73
Computer Science	108	0.73
Physiology	74	0.72
Accounting	114	0.71
Management	147	0.70
Chemical Engineering	56	0.70
Pharmacy	83	0.66
Marketing	109	0.65
Electrical Engineering	117	0.65
Biochemistry	61	0.64
Agriculture	164	0.60
Finance	72	0.60
Civil Engineering	94	0.56
Mechanical Engineering	103	0.53
Business	58	0.52
International Business	5	0.40

Table A11

DIF for "Help students develop personal values".

Academic Discipline	Count	Average Observations
Home Economics	42	2.33
Elementary Education	55	2.20
Education Administration	83	2.19
Nursing	179	2.19
Educational Psychology & Counseling	70	2.16
Religion and Theology	153	2.05
Special Education	58	2.03
Medicine	53	2.02
Law Enforcement	8	2.00
Dramatics and Speech	168	1.99
Music	313	1.98
Physical and Health Education	131	1.97
Secondary Education	76	1.93
Social Work	59	1.92
Journalism	74	1.91
Music and Art Education	53	1.89
Spanish	119	1.88
Management	145	1.86
Art	210	1.86
Nuclear Engineering	6	1.83
Philosophy	194	1.82
Foreign Language and Literature	121	1.79
English and Literature	529	1.78
Library Science	39	1.77
Architecture	93	1.73
Agriculture	164	1.72
Accounting	115	1.72
German	60	1.72
Chemical Engineering	57	1.70
Law	72	1.69
Aeronautical/Astronautical Engineering	18	1.67
Pharmacy	83	1.65
Marketing	108	1.63
History	384	1.63
Physiology	75	1.61
Business	59	1.61
French	82	1.60
Finance	72	1.57
Civil Engineering	95	1.57
Psychology	75	1.56
Anthropology	151	1.55
Mechanical Engineering	103	1.53

Geography	78	1.53
Bacteriology/Microbiology	96	1.47
Computer Science	108	1.47
Botany	72	1.44
Biology	139	1.44
Chemistry	226	1.44
Political Science	264	1.44
Sociology	254	1.43
Electrical Engineering	116	1.42
Biochemistry	61	1.41
International Business	5	1.40
Zoology	104	1.37
Physics	171	1.36
Math and Statistics	442	1.28
Economics	231	1.27
Marine Science	17	1.24
Astronomy	17	1.00

Table A12

DIF for "Enhance the out-of-class experience of students".

Academic Discipline	Count	Average Observations
Law Enforcement	8	2.00
Education Administration	82	1.80
Social Work	59	1.75
Home Economics	42	1.74
Educational Psychology & Counseling	70	1.73
Dramatics and Speech	166	1.73
Special Education	58	1.72
Physical and Health Education	130	1.69
Elementary Education	55	1.67
Agriculture	166	1.64
Nursing	178	1.63
Art	207	1.61
Journalism	76	1.61
Pharmacy	84	1.56
Secondary Education	76	1.55
Music	314	1.50
Management	144	1.44
Botany	72	1.42
Marketing	109	1.42
Music and Art Education	53	1.42
Spanish	118	1.42
Architecture	93	1.38
Medicine	53	1.38
Geography	77	1.34
Biology	139	1.33
Psychology	76	1.33
Zoology	104	1.32
Anthropology	151	1.30
Bacteriology/Microbiology	96	1.29
Business	59	1.29
Marine Science	17	1.29
Accounting	114	1.28
Aeronautical/Astronautical Engineering	19	1.26
Library Science	39	1.26
Physiology	75	1.23
Chemical Engineering	57	1.23
Civil Engineering	95	1.23
Foreign Language and Literature	119	1.23
German	60	1.22
French	84	1.20
Nuclear Engineering	6	1.17
Mechanical Engineering	104	1.16

Electrical Engineering	116	1.15
English and Literature	528	1.14
Computer Science	107	1.13
Sociology	254	1.13
Astronomy	17	1.12
Finance	72	1.11
Law	72	1.11
Physics	172	1.10
Religion and Theology	151	1.09
Philosophy	192	1.08
Biochemistry	61	1.07
Chemistry	226	1.05
Political Science	265	1.05
International Business	5	1.00
History	383	0.97
Economics	232	0.91
Math and Statistics	443	0.90

Table A13

DIF for "Enhance students' self-understanding".

Academic Discipline	Count	Average Observations
Elementary Education	55	2.27
Social Work	59	2.25
Nursing	177	2.24
Art	209	2.23
Dramatics and Speech	169	2.23
Educational Psychology & Counseling	70	2.17
Education Administration	83	2.11
Religion and Theology	153	2.10
Special Education	58	2.07
English and Literature	529	2.04
Music and Art Education	53	2.02
Physical and Health Education	129	2.02
Music	314	2.01
Philosophy	194	1.99
Home Economics	42	1.98
Spanish	119	1.92
German	59	1.90
Secondary Education	75	1.88
Law Enforcement	8	1.88
Anthropology	150	1.86
Management	147	1.84
Foreign Language and Literature	120	1.84
French	85	1.84
Medicine	53	1.81
Psychology	75	1.80
History	382	1.75
Architecture	93	1.73
Journalism	76	1.71
Pharmacy	84	1.68
Sociology	256	1.68
Biology	138	1.67
Political Science	263	1.63
Agriculture	167	1.60
Law	72	1.60
Physiology	75	1.59
Library Science	39	1.59
Botany	72	1.56
Marketing	108	1.56
Bacteriology/Microbiology	97	1.48
Aeronautical/Astronautical Engineering	19	1.47
Marine Science	17	1.47
Business	59	1.46

Accounting	114	1.42
Finance	72	1.40
Computer Science	108	1.39
Zoology	104	1.38
Geography	78	1.36
Nuclear Engineering	6	1.33
Astronomy	16	1.31
Chemistry	224	1.31
Mechanical Engineering	104	1.28
Physics	172	1.28
Civil Engineering	95	1.27
Biochemistry	61	1.26
Chemical Engineering	57	1.26
Math and Statistics	440	1.26
Economics	231	1.25
International Business	5	1.20
Electrical Engineering	116	1.16

Table A14

DIF for "Instill in students a commitment to community service".

Academic Discipline	Count	Average Observations
Social Work	59	2.12
Law Enforcement	8	2.00
Elementary Education	55	1.98
Nursing	180	1.96
Education Administration	83	1.84
Educational Psychology & Counseling	70	1.79
Medicine	53	1.75
Special Education	58	1.69
Secondary Education	76	1.68
Home Economics	42	1.64
Physical and Health Education	131	1.62
Architecture	93	1.61
Music and Art Education	51	1.55
Pharmacy	84	1.54
Journalism	76	1.45
Library Science	39	1.44
Agriculture	166	1.39
Spanish	119	1.39
Religion and Theology	153	1.39
Sociology	255	1.38
Law	72	1.36
Dramatics and Speech	169	1.32
Geography	78	1.31
Anthropology	151	1.30
Art	209	1.29
Political Science	264	1.27
Botany	72	1.25
Accounting	114	1.25
Music	313	1.19
English and Literature	529	1.19
Philosophy	192	1.19
Marine Science	17	1.18
Psychology	76	1.18
Business	59	1.17
Civil Engineering	94	1.17
Foreign Language and Literature	120	1.16
Management	147	1.14
History	383	1.13
Physiology	75	1.11
Bacteriology/Microbiology	96	1.10
Marketing	109	1.10
Biology	138	1.06

French	84	1.02
Aeronautical/Astronautical Engineering	19	1.00
Chemical Engineering	57	1.00
Nuclear Engineering	6	1.00
German	59	1.00
Zoology	103	0.99
Electrical Engineering	116	0.97
Finance	72	0.96
Mechanical Engineering	103	0.95
Physics	171	0.94
Biochemistry	61	0.93
Chemistry	226	0.91
Computer Science	107	0.90
Math and Statistics	442	0.85
Economics	231	0.85
International Business	5	0.80
Astronomy	17	0.76

Table A15

DIF for "Prepare students for responsible citizenship".

Academic Discipline	Count	Average Observations
Elementary Education	54	2.28
Education Administration	81	2.17
Educational Psychology & Counseling	69	2.14
Social Work	59	2.14
Secondary Education	75	2.13
Law Enforcement	8	2.13
Physical and Health Education	130	2.05
Medicine	53	2.04
Home Economics	42	2.02
Nursing	179	2.01
International Business	5	2.00
Political Science	264	2.00
Journalism	75	1.93
Special Education	58	1.91
History	382	1.91
Music and Art Education	52	1.88
Library Science	39	1.87
Spanish	118	1.85
Architecture	93	1.83
Geography	78	1.83
Law	71	1.83
Sociology	254	1.81
Pharmacy	84	1.80
English and Literature	528	1.79
Botany	72	1.76
Dramatics and Speech	168	1.76
Anthropology	150	1.76
Philosophy	193	1.75
Agriculture	166	1.72
Religion and Theology	152	1.72
Foreign Language and Literature	119	1.71
Marine Science	17	1.71
Aeronautical/Astronautical Engineering	19	1.68
Psychology	76	1.68
Accounting	114	1.67
Nuclear Engineering	6	1.67
Biology	138	1.66
Management	147	1.65
Civil Engineering	94	1.62
German	59	1.61
Bacteriology/Microbiology	94	1.60
Music	313	1.60

French	85	1.58
Economics	229	1.58
Marketing	109	1.56
Chemical Engineering	56	1.55
Zoology	100	1.54
Physiology	74	1.53
Art	209	1.52
Business	59	1.49
Biochemistry	61	1.43
Chemistry	224	1.43
Physics	173	1.39
Computer Science	108	1.37
Mechanical Engineering	104	1.35
Astronomy	17	1.35
Finance	71	1.34
Electrical Engineering	115	1.33
Math and Statistics	441	1.33

Table A16

DIF for "Enhance students' knowledge of and appreciation for other racial/ethnic groups".

Academic Discipline	Count	Average Observations
Elementary Education	55	2.53
Social Work	59	2.47
Spanish	119	2.46
Anthropology	151	2.46
Educational Psychology & Counseling	69	2.38
Nursing	179	2.32
French	85	2.26
Law Enforcement	8	2.25
Secondary Education	76	2.22
Music and Art Education	52	2.15
Foreign Language and Literature	121	2.13
Education Administration	83	2.12
Special Education	58	2.10
Dramatics and Speech	169	2.08
History	383	2.07
Journalism	76	2.07
Home Economics	42	2.05
English and Literature	533	2.05
Sociology	255	2.04
German	59	1.98
Religion and Theology	153	1.96
Library Science	39	1.95
Physical and Health Education	130	1.91
Art	209	1.88
Psychology	76	1.76
Political Science	265	1.74
Music	313	1.73
Medicine	53	1.72
Geography	78	1.69
Architecture	93	1.61
Pharmacy	84	1.57
Law	72	1.51
Nuclear Engineering	6	1.50
Management	147	1.49
Marketing	108	1.48
Philosophy	192	1.48
Physiology	75	1.45
Business	60	1.45
Agriculture	167	1.41
International Business	5	1.40
Biology	139	1.38
Botany	71	1.37

Aeronautical/Astronautical Engineering	19	1.37
Bacteriology/Microbiology	96	1.32
Accounting	114	1.30
Marine Science	17	1.29
Zoology	103	1.27
Biochemistry	61	1.21
Finance	72	1.19
Economics	229	1.16
Civil Engineering	94	1.15
Math and Statistics	441	1.12
Astronomy	17	1.12
Physics	171	1.10
Chemical Engineering	56	1.09
Chemistry	226	1.09
Computer Science	108	1.06
Mechanical Engineering	104	1.04
Electrical Engineering	116	0.89

Table A17

DIF for "Study a Foreign Language".

Academic Discipline	Count	Average Observations
German	59	2.78
French	86	2.77
Spanish	119	2.71
Foreign Language and Literature	122	2.55
Anthropology	151	1.69
Library Science	39	1.62
History	382	1.59
Religion and Theology	153	1.56
Music	314	1.49
English and Literature	531	1.47
Political Science	262	1.30
Philosophy	193	1.24
Elementary Education	55	1.22
Educational Psychology & Counseling	69	1.22
Dramatics and Speech	169	1.22
Journalism	76	1.22
International Business	5	1.20
Geography	78	1.18
Social Work	59	1.17
Secondary Education	76	1.16
Architecture	93	1.15
Law Enforcement	8	1.13
Art	208	1.12
Zoology	103	1.11
Psychology	76	1.11
Home Economics	41	1.07
Music and Art Education	51	1.04
Marketing	108	1.02
Math and Statistics	440	1.02
Sociology	254	1.02
Physics	173	1.01
Nuclear Engineering	6	1.00
Marine Science	17	1.00
Botany	72	0.97
Nursing	179	0.97
Education Administration	83	0.96
Law	72	0.94
Astronomy	17	0.94
Chemistry	225	0.94
Agriculture	166	0.89
Physical and Health Education	130	0.88
Economics	229	0.87

Business	60	0.85
Aeronautical/Astronautical Engineering	19	0.84
Chemical Engineering	56	0.84
Physiology	75	0.83
Biology	139	0.83
Bacteriology/Microbiology	94	0.81
Management	145	0.81
Accounting	113	0.80
Pharmacy	83	0.77
Special Education	58	0.76
Biochemistry	61	0.75
Computer Science	108	0.75
Medicine	53	0.74
Finance	72	0.72
Mechanical Engineering	104	0.70
Electrical Engineering	116	0.69
Civil Engineering	94	0.67

Table A18

Academic Disciplines Comprising Various Dimensions of Biglan's Model

"Hard" disciplines	"Soft" disciplines	"Pure" disciplines	"Applied" disciplines	"Life" disciplines	"NonLife" disciplines
Nuclear Engineering	Agricultural Economics	Philosophy	Communication	Entomology	Philosophy
Physics	Accounting	Russian	Computer Science	Horticulture	Geology
Chemistry	Anthropology	German	Agronomy	Botany	Communication
Mechanical	Psychology	English	Horticulture	Dairy Science	Chemistry
Engineering	Finance	History	Dairy Science	Agronomy	Astronomy
Ceram Engineering	Economics	Anthropology	Agricultural Economics	Zoology	Physics
Civil Engineering	Philosophy	Botany	Nuclear Engineering	Microbiology	Accounting
Astronomy	Communication	Math	Civil Engineering	Special Ed.	Finance
Geology	Sociology	Zoology	Ceram Engineering	Agricultural	English
Agronomy	Political Science	Physiology	Mechanical	Economics	Civil Engineering
Microbiology	Vocational Technology	Geology	Engineering	Physiology	Nuclear Engineering
Agriculture	Special Education	Entomology	Finance	Secondary Ed.	Mechanical
Botany	Russian	Chemistry	Accounting	Sociology	Engineering
Horticulture	English	Physics	Special Ed.	Ed. Admin.	Ceram Engineering
Zoology	German	Psychology	Secondary Ed.	Anthropology	German
Entomology	History	Political Science	Ed. Admin.	Political Science	Russian
Dairy Science	Ed. Admin.	Microbiology	Vocational Technology	Vocational Tech.	Economics
Physiology	Secondary Ed.	Sociology		Psychology	Computer Science
Computer Science		Astronomy		History	Math
Math		Economics			

Table A19

Academic Disciplines Comprising Various Dimensions of Kolbs's Model

"Abstract" disciplines	"Concrete" disciplines	"Active" disciplines	"Reflective" disciplines
Physics	Physical Ed.	Ed Admin	Economics
Electrical Engineering	Medicine	Civil Engineering	Anthropology
Mechanical Engineering	Education	Chemical Engineering	Sociology
Chemical Engineering	Psychology	Architecture	Journalism
Math	Anthropology	Social Work	Physiology
Civil Engineering	Ed. Admin.	Ed Psych	Speech
Chemistry	Home Economics	Mechanical Engineering	Music
Ecology	Political Science	Education	Physics
Biochemistry	Architecture	Law	Biochemistry
Bacteriology	Sociology	Medicine	Bacteriology
Economics	Educational Psychology	Psychology	Agriculture/Forestry
Physiology	Nursing	Electrical Engineering	Political Science
Agriculture/Forestry	Law	Business	Art
Botany	Social Work	Ecology	Geography
Business	Philosophy		Botany
Zoology	Speech		Chemistry
Geography	Library Science		Home Economics
	Journalism		Zoology
	Music		Math
	Art		Library Science
	English		Nursing
	History		History
	German		Physical Education
	Spanish		English
	French		German
	Dramatic Arts		Spanish
			French
			Philosophy

Table A20

Academic Disciplines Comprising Various Dimensions of Holland's Model

Realistic	Artistic	Enterprising
Architecture	Architectural Environmental Design	Art History
Drafting/Design	Art	Commercial Music
Electrical Engineering	Drama	Communication
Electrical Engineering Technology	English	Finance
Manufacturing Engineering	Foreign Languages	History
Technology	Journalism	Hospitality and Resort Management
Marine Science	Music	International Business
Mechanical	Music/Art Education	Logistics/Marketing
Military Science	Philosophy	Logistics/Supply Chain Management
	Speech	Marketing Management
	Theatre Drama	Music Industry
		Political Science
		Real Estate
		Risk Management and Insurance
		Sales
		Sports and Leisure Studies

Investigative	Social	Conventional
Aeronautical/Astronautical Engineering	African-American Studies	Accounting
Allied Health (Medical Technologies)	Consumer Science and Education	Data Processing
Anthropology	Ethnic Studies	Library Science
Biology	Exercise and Sports Science	Secretarial Studies
Business Economics	Home Economics	
Chemistry	Human Development and Learning	
Civil Engineering	Integrative Studies (Elementary Education)	
Computer Engineering Technology	Literature	
Computer Science	Nursing	
Criminal Justice Studies	Physical Education	
Economics	Teacher Education	
Geography	Religion	
Geological Sciences	Social Work	
International Studies	Special Education	
Management	Theology	
Management Information Systems		
Mathematical Sciences		
Pharmacy		
Physical Sciences		
Physics		
Psychology		
Sociology		
Statistics		

References

Aiken, L. S., West, S. G., Sechrest, L., & Reno, R. R. (1990). Graduate training in statistics, methodology, and measurement in psychology. *American Psychologist, 45*(6), 721-734.

Alpert, D. (1985). Performance and paralysis: The organizational context American research university. *Journal of Higher Education, 56*, 241-81.

Andrich, D. (1978). A rating formulation for ordered response categories, *Psychometrica, 43*, 561-573.

Angelo, T. A., & Cross, K. P. (1993). Classroom assessment techniques: A handbook for college teachers (2nd ed.). San Francisco: Jossey-Bass.

Barnes, L. L. B., Bull, K. S., Campbell, N. J., & Perry, K. M. (1998). *Discipline-related differences in teaching and grading philosophies among undergraduate teaching faculty*. Paper presented at the Annual Meeting of the American Educational Research Association, April 13-17, San Diego, CA.

Barnes, L. L. B., Bull, K. S., Campbell, N. J., & Perry, K. M. (2001). Effects of academic discipline and teaching goals in predicting grading beliefs among undergraduate faculty. *Research in Higher Education, 42*(4), 455-467.

Becker, G. (2001). Controlling decremental and inflationary effects in reliability estimation resulting from violations of assumptions. *Psychological Reports, 89*, 403-424.

Becher, T. (1987). Disciplinary shaping of the profession. In B. R. Clark (ed.), *The Academic Profession: National, Disciplinary, and Institutional Settings*. Los Angeles: University of California Press.

Biglan, A. (1973a). The characteristics of subject matter in different academic areas. *Journal of Applied Psychology, 57*, 195-203.

Biglan, A. (1973b). Relationships between subject matter characteristics and the structure and output of university departments. *Journal of Applied Psychology, 57*, 3, 204-213.

Blank, R. (1978). An Organizational Model of Higher Education Institutions and Faculty Teaching Goals. *Sociological Inquiry, 48*, 1, 23-35.

Blank, R. (1976). An organizational model of academic institutions and the teaching goals of faculty members. (Doctoral dissertation, Florida State University, 1976). *Dissertation Abstracts International*, DAI-A 37/12, 7995.

Bond, T. & Fox, C. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ; Lawrence Erlbaum Associates, Inc.

Bradley, K. D., & Sampson, S. (Spring, 2005). A case for using Rasch Rating Scale analysis to assess the quality of measurement in survey research. *The Respondent*, 12-13.

Bradley, K. D., Cunningham, J. D., & Gilman, R. (2006). *Using the Rasch Rating Scale Model to Test the Fit and Function of MSLSS Scale across Groups*. Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 6-11, 2006).

Braxton, J. M., & Nordvall, R. C. (1985). Selective liberal arts colleges: Higher quality as well as higher prestige? *Journal of Higher Education, 5*, 538-554.

Brown, D. (2003). A comparison of teaching goals and objectives identified by faculty and peer learning facilitators within a transition seminar for new college students. (ERIC Document Reproduction Service No. ED 479 654).

Bruning, J. L., & Kintz, B. L. (1997). *Computational Handbook of Statistics* (4th ed.). New York: Addison-Wesley Educational Publishers Inc.

Bunderson, C. V. (2000). *Design experiments, design science, and the philosophy of measured realism: Philosophical foundations of design experiments*. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA. April 24-28.

Clark, B. R. (1980). Academic Culture. Higher Education Research Group, Yale University.

Colbeck, C. L., Cabrera, A. F., & Marine, R. J. (2002). *Faculty motivation to use alternative teaching methods*. Paper presented at the Annual Meeting of the American Educational Research Association, April 1-5, New Orleans, LA.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart and Winston, Inc.

Curtis, D. & Keeves, J. (1999). *The Course Experience Questionnaire as an institutional performance indicator*. Paper presented at the HERDSA Annual International Conference, July 12-15, Melbourne, Australia.

Donald, J. (1990). Knowledge and the university curriculum, in Conrad, C.F. and J.G. Haworth (eds.). *Curriculum in Transition: Perspectives on the Undergraduate Experience.* ASHE Reader series. Needham Heights, MA: Ginn Press.

Du, Y. (1995). When to adjust for Differential Item Functioning. *Rasch Measurement Transactions, 9*(1), 414.

Embretson, S. E. & Hershberger, S., (1999). Eds. *The new rules of measurement: What every psychologist and educator should know*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Fan, X. (1998). Item response theory and classical test theory: An empirical comparison of their item/person statistics. *Educational & Psychological Measurement, 58*, 357-380.

Fox, E. A. (1997). Examining factors which influence and significantly impact faculty teaching goals. (Doctoral dissertation, University of South Carolina, 1997) *Dissertation Abstracts International, 58*(03A), 0763.

Franklin, J., and Theall, M. (1992). Disciplinary differences: Instructional goals and activities, measures of student performance and student ratings of instructors. *Paper presented at the meeting of the American Educational Research Association,* San Francisco.

Gaff, J. G., & Wilson, R. C. (1971). Faculty cultures and interdisciplinary studies. *Journal of Higher Education, 42*(3), 186–201.

Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory: principles and applications.* Boston: Kluwer-Nijhoff Publishing.

Hambleton, R. K., Swaminathan, H., & Rogers, J. H. (1991). *Fundamentals of item response theory.* New York: Sage publications.

Hardy, K. A. (2002). Faculty teaching goals in the online environment. (Doctoral dissertation, Florida State University, 2002) *Dissertation Abstracts International, A 63/09*, 3094.

Hativa, N. (1997). *Teaching in a research university: Professors' conceptions, practices, and disciplinary differences*. Paper presented at the Annual Meeting of the American Educational Research Association, March 24-28, Chicago, IL.

Hativa, N., & Birenbaum, M. (2000). Who prefers what? Disciplinary differences in students' preferred approaches to teaching and learning styles. *Research in Higher Education, 41*(2), 209-236.

Hays, W. L. (1988). *Statistics* (4th ed.). Fort Worth: Holt, Rinehart and Winston.

Holland, J. L. (1966). *The psychology of vocational choice.* Waltham, Massachusetts: Blaisdell.

Hutchinson, S. R., & Lovell, C. D. (2004). A review of methodological characteristics of research published in key journals in higher education: Implications for Graduate research training. *Research in Higher Education, 45,* 4, 383-403.

Ironson, G. H. (1982). Use of chi-square and latent trait approaches for detecting item bias. In R. A. Berk (Ed.), *Handbook of methods for detecting test bias*. Baltimore, MD: Johns Hopkins University Press.

Jervis, F. M., & Cogdon, R. G. (1958). Student and faculty perceptions of educational values. *American Psychologist, 13*, 464-66.

Johnson, L. D. (1997). Faculty teaching goals at senior research universities. (Doctoral dissertation, Virginia Polytechnic Institute and State University, 1997). *Dissertation Abstracts International, 59*(12A), 4377.

Kaya, N., Webb, J. D., & Weber, M. J. (2005). Faculty scholarly goals: Relationships between institutional, departmental, and individual's emphasis. *Education, 125*, 3, 446-453.

Kolb, D. A. (1980). Learning Styles and Disciplinary Differences. In Chickering, A. W., & Associates (Eds.). *The Modern American College*. San Francisco: Josey Bass.

Ladd, E. C., & Lipset, S. M. (1975). Technical Report: 1975 Survey of the American professorate. Storrs, CT: The University of Connecticut, Social Science Data Center.

Lattuca, L. & Stark, J. (1994) Will disciplinary perspectives impede curricular reform? *Journal of Higher Education, 65*, 401–426.

Lawrence, J. H., Hart, K. A., Mackie, C., Muniz, D., & Dickmann, E. (1990). A comparison of the teaching goals, assumptions, and practices of faculty in eight liberal arts disciplines. Paper presented at the Annual Meeting of the Association for the Study of Higher Education, Portland, OR. (ERIC Document Reproduction Service No. ED 326 115)

Lee, J. J. (2004). Comparing institutional relationships with academic departments: A study of five academic fields. *Research in Higher Education, 45*(6) 603-624.

Leverenz, T. R., & Lewis, B. R. (1981). An analysis of faculty consistency in the academic professions. (ERIC Document Reproduction Service No. ED 201 254).

Liebert, R. J., & Bayer, A. E. (1975). Goals in teaching undergraduates: Professional reproductions and client-centeredness. *The American Sociologist, 10*, 195-205.

Linacre, J. M. (1997). KR-20 / Cronbach Alpha or Rasch Reliability: Which Tells the "Truth"? *Rasch Measurement Transactions, 11*(3), 580-581.

Linacre, J. M. (2004a). *A user's guide to Winsteps Rasch-Model computer programs.* Chicago, IL: MESA Press.

Linacre, J. M. (2004b). WINSTEPS Rasch measurement computer program, Chicago: http://winsteps.com

Linacre, J. M. (2005). The partial credit model and the one-item rating scale model. *Rasch Measurement Transactions, 19*(1), 1000-1002.

Linacre, J. M. (2006). Facets Rasch measurement computer program. Chicago: Retreived from: http://www.winsteps.com/aftp/facets.pdf.

Lord, F. M. (1980). *Applications of item response theory to practical testing problems.* Hillsdale, NJ: Erlbaum.

Lovell, C. D., & Hutchinson, S. R. (2003). *Research training requirements in higher education doctoral programs*. Manuscript in preparation.

Lovell, C. D., Hutchinson, S. R., & Fairweather, J. R. (1999). *Graduate student research preparation in higher education: Too little, just right, or too much... Implications for our future*. Paper presented at the Annual Meeting of the Association for the Study of Higher Education. November, 1999, Miami, FL.

Marascuilo, L. A., & Slaughter, R. E. (1981). Statistical procedures for identifying possible sources of item bias based on chi-square statistics. *Journal of Educational Measurement, 18*, 229-248.

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*, 149-174.

Meaningful Measurement (2008). Requirements for Objective Measurement. Retrieved February 6, 2008. Available at: http://www.meaningfulmeasurement.com/Objective%20Measurement.pdf.

National Center for Postsecondary Improvement (1999). Grading skills by discipline: A comparison of faculty instructional goals. *Change*, May/June, 49-52.

Neumann, R., Parry, S., & Becher, T. (2002). Teaching and learning in their disciplinary contexts: A conceptual analysis. *Studies in Higher Education, 27*, 4, 405-417.

Peck, B. (2001). *Monocultural education for a multicultural population: The "ethnic disadvantage' issue revisited.* Unpublished doctoral dissertation, Murdoch University.

Platt, G. M., Parsons, T., & Kirshstein, R. (1976). Faculty teaching goals, 1968-1973. *Social Problems, 24*, 298-307.

Richter, P., Werne, J., Heerlein, A, & Sauer, H. (1998). On the validity of the Beck Depression Inventory. *Psychopathology, 31*, 160-168.

Rosser, V. J. (2005). Measuring the change in faculty perceptions over time: An examination of their worklife and satisfaction. *Research in Higher Education, 46*(1), 81-107.

Royal, K. D., Eli, J. A., & Bradley, K. D. (2005). *A pilot study exploring community college faculty perceptions of student outcomes.* Paper presented at the Mid-Western Education Research Association Annual Meeting, October 12-15, Columbus, OH.

Sampson, S. & Bradley, K. D. (2003). Sampson, S. and Bradley, K. D. (2003). Rasch analysis of educator supply and demand rating scale data: An alternative to the true score model. *Research Methods; The Forum.* Available at: http://aom.pace.edu/rmd/2003forum.html

Schwarze, C. C. (1996). A comparison of teaching goals of full-time and part-time college faculty. (Doctoral dissertation, Saint Louis University, 1996). *Dissertation Abstracts International, 57*(07A), 2904.

Shavelson, R. J., Ruiz-Primo, M. A., & Wiley, E. W. (2005). Windows into the mind. *Higher Education, 49*, 413-430.

Shepard, L. A., Camilli, G., & Williams, D. M. (1984). Accounting for statistical artifacts in item bias research. *Journal of Educational Statistics, 9*, 93-128.

Smart, J. C. (2005). Attributes of exemplary research manuscripts employing quantitative analyses. *Research in Higher Education, 46*, 4, 461-477.

Smart, J. C. (1982). Faculty teaching goals: A test of Holland's theory. *Journal of Educational Psychology, 74*(2), 180-188.

Smart, J. C., & Elton, C. F. (1975). Goal orientations of academic departments: A test of Biglan's theory. *Journal of Applied Psychology, 60*, 580-588.

Smart, J. C. & Ethington, C.A. (1995) Disciplinary and institutional differences in undergraduate education goals. In N. Hativa & M. Marincovich (Eds.), *Disciplinary differences in teaching and learning: Implications for practice*, pp. 49–57. San Francisco, CA: Jossey-Bass.

Smart, J. C., Feldman, K. A., & Ethington, C. A. (2000). *Academic disciplines: Holland's theory and the study of college students and faculty.* Nashville: Vanderbilt University Press.

Smeby, J. C. (1996). Disciplinary differences in university teaching. *Studies in Higher Education, 21*, 1, 69-80.

Smith, Jr., E. V., & Smith, R. (Eds.). (2004). *Introduction to Rasch measurement: Theory, models and applications*. Maple Grove, MN: JAM Press.

Stark, J. S., & Morstain, B. R. (1978). Educational orientations of faculty in liberal arts colleges: An analysis of disciplinary differences. *Journal of Higher Education, 49*, 420-437.

Swenson, C. D. (1997). Teaching goals of practitioner-educators: Improving teaching and learning in continuing higher education. *Continuing Higher Education Review, 61*, 88-103.

Trice, A. G., & Dey, E. L. (1997). Trends in faculty teaching goals: A longitudinal study of change. *Journal of College Student Development, 38*(5), 527-534.

Waugh, R. (1999). Approaches to studying for students in higher education: A Rasch measurement model analysis: *British Journal of Educational Psychology, 69*, 63-79.

Wilson, R. C., Gaff, J. G., Dienst, E. R., Wood, L., & Bavry, J. L. (1975). *College professors and their impact on students*. New York: Wiley.

Wright, B. D. (1999). Model selection: Rating scale or partial credit? *Rasch Measurement Transactions, 12*(3), 641-642.

Wright, B. D. (1997). Measurement for Social Science and Education: The history of social science measurement. Available at: http://www.rasch.org/memo62.htm. Retrieved on May 30, 2007.

Wright, B. D., & Masters, G. (1981). *The measurement of knowledge and attitude* (Research memorandum No. 30). Chicago: Statistical Laboratory, Department of Education, University of Chicago.

Wright, B. D., & Stone, M. (1979). *Best test design: Rasch measurement.* Chicago: MESA Press.

VITA

Date of birth: December 23, 1979

Place of birth: Grundy, Virginia

Educational institutions attended and degrees already awarded:

University of Kentucky, Masters Degree, 2003

Alice Lloyd College, Bachelors Degree, 2002

Southwest Virginia Community College, Associate of Arts & Science Degree, 2000

Scholastic and professional honors:

Best Paper Award, 2008 College Teaching & Learning (TLC) Conference, "An investigation of 'honesty check' items in higher education course evaluations"

Association Council Nominee, Midwestern Educational Research Association (MWERA), 2008

Alumni Graduate Fellowship, University of Kentucky, 2006-2007; 2007-2008

Teacher Who Made a Difference Award, University of Kentucky, 2006-2007

Chancellor's List, University of Kentucky, 2005-2006; 2006-2007

Who's Who Among Students at American Universities and Colleges, University of Kentucky, 2006-2007

Ambassador, College of Education, University of Kentucky, 2004-2006

Barbara Ayres Threadgill Scholarship for Development, University of Kentucky, 2005-2006

John Edwin and Gwendolyn Gray Partington Scholarship, University of Kentucky, 2003-2004; 2004-2005; 2005-2006

Kentucky Graduate Scholarship, University of Kentucky, 2003-2008

Salutatorian, Alice Lloyd College, 2002-2003

Caney Scholar, Alice Lloyd College, 2003-2008

Semester in London Scholarship, Alice Lloyd College, 2002

Who's Who Among Students in American Colleges and Universities, Alice Lloyd College, 2002

Who's Who Among Students in American Junior Colleges, Southwest Virginia Community College, 1999; 2000

Academic All-American Nominee, Southwest Virginia Community College, 1999

Harlan and Vera Walls Scholarship, Southwest Virginia Community College, 1999

First Union Bank Scholarship, Southwest Virginia Community College, 1998

Mark VanMeter Memorial Scholarship, Southwest Virginia Community College, 1998

Grundy Lions Club Scholarship, Southwest Virginia Community College, 1998

Professional publications:

Bradley, K. D., Royal, K. D., Cunningham, J. C., Eli, J. A., Weber, J., (2008). What is 'Good' Research? A consideration of Ethics, Methods, and Theory. *Mid-Western Educational Research Journal, 21*, 1, 26-35.

Royal, K. D., Eli, J. A., & Bradley, K. D. (Accepted). A Pilot Study Exploring Community College Faculty Perceptions Regarding Student Outcomes. *Community College Journal of Research and Practice.*

Royal, K. D., Drake, G., Lineberry, G.T., & Murphy, W. (2007). Investigating Diversity: A Study of One Institution's Main and Extended Campus Undergraduate Engineering Programs. *Proceedings of the* Spring 2007 *American Society for Engineering Education Illinois-Indiana Section Conference.* Available at: http://www.asee4ilin.org/Conference2007program/Papers/Conference%20Papers/Session%205B/Royal.pdf

Bradley, K. D., Sampson, S. O., & Royal, K. D. (2006). Applying the Rasch Rating Scale Model to Gain Insights into Students' Conceptualization of Quality Mathematics Instruction. *Mathematics Education Research Journal, Special Issue: Research in Mathematics Education: Insights from Rasch Measurement.*

Royal, K. D. (2006). [Review of book *500 Tips for Open and Online Learning (2nd Edition)]*. *Education Review*, April. Available at: http://edrev.asu.edu/brief/apr06.html.

Royal, K. D., & Bradley, K. D. (2005). Interactive Television Courses and Students' Satisfaction. *Journal of Distance Learning Administration, 8*(4).

Royal, K. D., Bradley, K. D., and Lineberry, G. T. (2005). Evaluating Interactive Television (ITV) Courses: An Identification of Factors Associated with Student Satisfactions. *Journal of Distance Learning Administration, 8*(2).

Signature of student:

Kenneth Darrell Royal

Kenneth Darrell Royal