Research Articles

Student Perceptions of Teacher Evaluations: A Proposed Model

H. Joey Gray
Middle Tennessee State University
Department of Health & Human Performance
P.O. Box 96, Murfreesboro, TN 37132

Phone: (615) 904-8359 • E-mail: hjgray@mtsu.edu

Louis S. Nadelson
Boise State University

James A. Busser
University of Nevada, Las Vegas

Abstract

Students hold multifaceted perceptions of end-of-course teacher evaluations may guide their response behaviors when completing these evaluative measures. Research shows that student perceptions typically vary from those held by faculty and administration. Although research on student evaluations of teachers (SETs) is prevalent in the literature, research focusing on student perceptions is limited; SET research in the academic area of recreation is even more scant. Researchers have reported that students frequently do not understand the purpose of the evaluations and may not always provide the most accurate data. A model of the interaction of student variables that are reflective of the complexity of their interactions with evaluations can contribute to greater understanding of the process. Data was collected from 523 students majoring in recreation using a 52 item theoretically based survey. Path analysis showed a model of the interactions between five latent variables representative of student perceptions of teacher evaluations. Through the creation of a model that combines multiple variables, this investigation has enhanced understanding of student perceptions and behaviors associated with the teacher evaluation process. Results indicate that the measured variables can be attributed to five latent variables predictive of student perceptions of teacher evaluations. The hypothesized model predicts how these variables interact to contribute to student perceptions of teacher evaluations.

Background

Student Evaluations of Teachers (SETs) began in the 1920's (Doyle, 1983) and remain prevalent in most colleges and universities today. The use of SETs encompasses formative data for faculty such as course improvement (Hobson & Talbot, 2001), summative uses for administrative decision making such as tenure and promotion (Hobson & Talbot, 2001), and student selection of courses and faculty (Wilhelm & Comegys, 2004). Yet, how students, professors, and college administrators perceive evaluations may be very different (Marlin, 1987). Regardless of the use of SET data, all stakeholders are expecting accuracy in student responses. This has led researchers to attempt to isolate the factors that influence student responses.

Given the complexity of SETs, researchers have examined the impact of variables such as instructor personality (Cardy & Dobbins, 1986; Chonko, Tanner, & David, 2002; Clayson, 1999; Williams & Ceci, 1997), gender (Centra & Gaubatz, 2000; Dukes & Victoria, 1989), student expectations (Anderson & Miller, 1997; Crader & Butler, 1996; Dunegan & Hrivnak, 2003), and the teaching experience of instructors (Feldman, 1993). These provide justification to hypothesize that student perceptions of SETs influence how they complete their end of course evaluations. Yet, most studies have focused on instructor variables leaving a gap in the literature regarding students' perceptions of SETs.

Aside from the general absence of student perceptions of SETs, the majority of research related to this topic has been conducted in the academic areas of education, psychology, and business, leaving a gap in recreation research. Skepticism certainly exists regarding the accuracy of SET data and recreation curricula are not devoid of such scrutiny. However, there is a dearth of research investigating student perceptions of SETs in recreation curricula. Many colleagues in the field of recreation education have entertained comments from students, parents, and colleagues outside the field recounting blanket statements such as "recreation courses are easy" (Zuefle, 2004). If students and others outside the education community hold this view, it may impact the outcomes of SETs in recreation courses, especially when the course is more rigorous than the student expected. Dunegan and Hrivnak (2003) found that if a student holds a predetermined expectation and the instructor does not fit the norms of what is expected in the course, the student could perceive that a teacher is inadequate which in turn may impact the accuracy of SET data. Finally, Chonko, Tanner, and David (2002) noted contrary to common belief, teachers have less control over SETs because of students' varying expectations. Thus, a clear understanding of student perceptions of SETs in recreation is imperative to clarify how student perceptions may or may not impact SETs in recreation courses.

The published works on student perceptions all note the importance of research in this area and lack of comprehensive investigations studying these variables (Chen & Hoshower, 1998; Costin, Greenough, & Menges, 1971; Spencer & Schmelkin, 2002). The significance for understanding student's views of SETs is supported by Marlin (1987) who maintained that student perceptions of SETs

must be examined or the validity of the entire process is called into question. Researchers have investigated the background factors such as course characteristics (Koushki & Kuhn, 1982; Marsh & Dunkin, 1992; Wachtel, 1998), and instructor characteristics (Cardy & Dobbins, 1986; Chamberlin & Hickey, 2001; Chonko, Tanner, & David, 2002; Clayson, 1999; Williams & Ceci, 1997); however, limited research exists that is specifically aimed at determining the underlying factors that influence student perceptions.

Seriousness and Knowledge

Two broad themes found in the research on student perception of SETs are students' level of seriousness when completing the evaluations (Dunegan & Hrivnak, 2003, Jacobs, 1987; McKeachie, 1997; Marlin, 1987; Martinson & Ryan, 1981; Sheehan & DuPrey, 1999; Simpson & Siguaw, 2000; Smith & Carney, 1990; Spencer & Schmelkin, 2002) and student knowledge of the utility of evaluations (Smith & Carney, 1990; Marlin, 1987; Spencer & Schmelkin, 2002). The findings of this research revealed that students perceived that they completed SETs seriously, although, they were uncertain who viewed SET data and how the SET data was to be used. Research on student perceptions has made some progress; yet, understanding the influence student perceptions have on their SET responses may not be that clear cut. It appears that student perceptions are the result of a complex web of factors that interact and which are manifested in their responses on teacher evaluations.

The aim of our investigation was to propose a model of the underlying factors that influence student perceptions of SETs, specifically in the academic area of recreation. The completion of the end-of-course evaluations may be domain specific and recreation students may hold unique perceptions. The recreation discipline is a distinctive area of study given the uniqueness of the courses offered in the discipline and recreation students admit they possess lower expectations for recreation and leisure courses (Zuefle, 2004). Thus, when a teacher presents higher challenges than expected, a negative perception may be formed which may affect SET responses. This notion is not only restricted to students, the recreation discipline is often questioned by those in and outside of academia (Zuefle, 2004). At present no student perception of SET models are available for comparison. Thus, the findings could reveal that the model may be applied widely or the model necessitates modification for students in other academic domains. Regardless, the model will provide a specific picture of the factors that influence students enrolled in recreation courses. Possessing a deeper understanding of the factors that influence student's perceptions of teacher evaluations may better equip faculty and administrators to interpret SET data as well as structure SETs in a manner that allow for gathering data that is representative and useful.

Impact of Student Level of Seriousness and Knowledge

Common sense indicates that end-of-course evaluations are most useful and representative when students take the time to provide honest and constructive feedback. Research has revealed that faculty may be skeptical about the level of

seriousness (mindfulness) students hold when completing SETs, inferring that students view the process as a menial task or boring chore (McKeachie, 1997; Simpson & Siguaw, 2000). From an investigation focusing on student perceptions of the SET process, Jacobs (1987) reported that 40% of respondents indicated they had knowledge of students collaborating to give low evaluation scores in an attempt to "get back" at an instructor. Further, Dunegan and Hrivnak (2003) found that when student educational expectations were consistently met by the teacher, students completed the SETs mindlessly. However, when student expectations of the instructor were not met, students completed SETs in a more intentional manner.

There is also evidence to suggest that students generally attempt to complete SETs mindfully. Marlin (1987), for example, found that only 8 % of students reported that they were lackadaisical when completing SETs while over half of his study participants responded that they took their time when completing SETs and attempted to provide fair and accurate SET responses. Despite students' reporting intentions for their efforts to be fair, Marlin also found that students viewed SETs as a way to "let off steam" and indicated they felt that administrators and faculty held little value for student opinions.

Smith and Carney's (1990) multifaceted study of SETs involved the investigation of student perceptions, the seriousness with which students complete SETs, and the value placed on student opinions by faculty and administrators. The authors developed a 31-item measure, the Student Perceptions of Evaluations Questionnaire (SPEQ). They utilized a 5-point Likert scale and three open-ended questions to obtain information regarding student perceptions of SETs. Their findings revealed that students took the completion of SETs seriously; however, students did not have a thorough understanding of the use of SET data. In addition, students indicated limited awareness that instructors often used SET data to make improvements in teaching methods and course content. Almost certainly, the lack of student knowledge regarding the application of the data could have an impact on SET responses.

Perceived Value of Feedback.

According to Smith and Carney (1990), students also believed that they should have more influence regarding the evaluation of their instructors. Their investigation revealed that students had the misconception that SETs have little impact regarding an instructor's effectiveness. In actuality, students can have a great deal of influence when SET data is used in administrative decisions (Smith & Carney, 1990). Again, if students had a clear understanding of the impact their feedback has on administrative decisions, it could influence their perceptions and responses to SETs. Furthermore, research that revealed that students felt they should have more influence may be an indication that they believed that their feedback was not valued by administrators or instructors (Marlin, 1987; Spencer & Schmelkin, 2002). In fact, several investigations have revealed that students demonstrated uncertainty toward faculty and administrators' value of their feedback (Marlin, 1987; Spencer & Schmelkin, 2002; Wulff, Staton-Spicer, Hess,

& Nyquist, 1985). More specifically, Spencer and Schmelkin (2002) found that a combination of the lack of student awareness of the purposes/uses of SET data and student uncertainty of the value of their opinions most likely impacted their ability to mindfully complete end of course evaluations.

An understanding of student knowledge of the purposes and uses of SETs is an important factor in interpreting SET data. For example, if a student understands that SET data will influence instructor motivation to change teaching methods or course content, then the student may have a more serious attitude when completing SETs (Grimes, Millea, & Woodruff, 2004). This knowledge could also have an adverse effect if students perceive a sense of control in their responses. If students have knowledge that college administrators reviewed SET data for decisions about tenure and promotion, there is the potential for students to use SET responses in an attempt to influence an instructor's career.

On the other hand, a lack of knowledge of the purposes and uses of SETs could lead to a lack of perceived control and ultimately to mindless responses on SETs (Spencer & Schmelkin, 2002). Marlin (1987) found that students completing SETs held the misconception that administrators pay little attention to SET results. His study also revealed that between 65% and 70% of the students reported that administrators paid attention to SET results only a limited amount of time. Additionally, Marlin found that students lacked knowledge that SET data was used in tenure and promotion decisions. This is significant in light of the fact that 70% of the department chairs indicated they used SET data in tenure and promotion. This provides significant support when questioning the accuracy of student knowledge regarding who will review their SET responses. Knowledge of who will review the results of SETs could influence student honesty, candidness, and the accuracy of responses. Yet, there is insufficient research addressing the issue of student perceptions of the uses/purposes of SETs (Smith & Carney, 1990). Despite this notation by Smith and Carney over two decades ago, research exploring the link between student perceptions and SETs is still lacking. Furthermore, no research has been conducted to explore a comprehensive model of student perspectives of SETs.

Theory of Planned Behavior

The importance of student perceptions of SETs has become more evident over the past ten years as researchers have continued to make strides in this domain identifying common variables and factors that influence student opinion. In general, SET research and investigation of student perceptions lacks a theoretical foundation. A person's understandings or perceptions of situations ultimately influence their behavior; this is indeed the essence of the Theory of Planned Behavior (TpB). Ajzen and Fishbein's (2005) Theory of Planned Behavior demonstrates how one's attitude, perceptions of peer beliefs, and perceived level of control toward a behavior shape individuals' intentions and actions. Possessing a clear understanding of the perceptions students hold about SETs will yield insight as to what underlying factors influence their response behaviors when completing SETs,

which may ultimately provide a clearer interpretation of SET data. A theoretical model explaining student perceptions of SETs could greatly benefit investigations in the domain of recreation education.

Methods

Given the shift in student demographics such as age, sex, and race (Keller, 2001) and education expectations such as the role of the teacher and the delivery of content (James, 2002) along with the changing landscape of universities, it is essential to maintain engaged in research that explores student perceptions of SETs. Through the examination of student perceptions and behaviors one may take steps toward identifying the underlying factors that impact and influence their SET completion process.

In an effort to determine if the process that students utilize in completing end-of-course evaluations could be modeled, the Student Perceptions of Student Evaluations of Teachers questionnaire (SPSET) was developed. The SPSET was created to measure students' perceptions of the teacher evaluation process. The SPSET was created through the application of the theoretical framework of the Theory of Planned Behavior (TpB). The TpB was selected for the framework of SPSET because it requires that one must understand perceptions to understand behavior. Specifically, the TpB explains there are existing factors that influence an individual's intention and as such one's intentions directly influences their behavior. An individual's intention toward a behavior is, in turn, a function of their attitude perception of peer beliefs and sense of control in the situation (Ajzen, 2002).

These underlying factors of attitude, peer beliefs, and control are hypothesized to encompass overall perceptions. As related to SETs, it would then be presumed that students' attitudes or perceptions of the evaluation process, what they believe their peers' beliefs about the evaluation process may be, and how much control they feel that they have may all directly affect the behavior students engage in when completing SETs. Thus, understanding these perceptions undergraduate college students hold toward SETs may provide a clearer understanding, and interpretation of SET data. The purpose of the present investigation was to propose a model of the underlying student factors that influence the completion of SETs. Our proposed model is intended to be a predecessor of a TpB based model focusing on the student variables associated with the SET process and should be considered as a step toward a complete theoretical model of the phenomenon.

SPSET Instrument Development

The development of the Student Perceptions of Student Evaluations of Teachers questionnaire was based on the work of Smith and Carney (1990). Smith and Carney designed a 31-item questionnaire to examine student perceptions of teacher evaluations in introductory psychology and education courses. Specifically, Smith and Carney were interested in (1) student perceptions of the uses of SETs, (2) student misconceptions of SETs, (3) the level of seriousness (attitude) with which students take the opportunity to evaluate their teachers, and (4) student perceptions of the value professors assign to student feedback and comments.

Using the Smith and Carney questionnaire as a model, 52 seven-point Likert scale items were developed for the SPSET to obtain data in the following five categories: (a) demographics, (b) knowledge of purposes (uses) of SETs, (c) seriousness with which students respond to SETs, (d) perceived value of SET feedback, and (e) accuracy of SET responses. In an effort to avoid systematic responses by subjects, the items were presented in a randomized order as suggested by Ajzen (2002). Additionally, several questions targeting the same information used a semantic differential response scale to test for accuracy of responses (Ajzen, 2002). Specifically, for questions that were repeated using positive and negative anchors, one question would utilize the positive anchor and the repeated question would utilize the negative anchor (e.g., definitely true – definitely false; extremely unlikely – extremely likely). Active voice verbiage was incorporated with the anchors along with a degree of novelty to reduce mindless and repetitive responses. Additionally, a unipolar measure seven-point optimal Likert scale was incorporated to measure belief strength and outcome evaluation. The seven-point scale was selected to provide increased answer choices and afford better discrimination, resulting in the higher likelihood of detecting true differences (Ajzen, 2002) and assisted with identifying variability and differences in responses. Thus, in an effort to increase the SPSET reliability both verbiage and visual differentiation were incorporated.

The SPSET was examined for both validity and reliability via a panel of experts and pilot testing. The panel of experts was comprised of a dean who oversees the pedagogical aspects of the college, the University Director for the Center of Evaluation, and the undergraduate coordinator for the recreation curriculum. The panel reviewed the questionnaire for content, reliability, and clarity of word choice. Upon review, panel comments were integrated and modifications were made which increased the content validity of the instrument. Major efforts were made to keep the questionnaire short in order to increase the likelihood of cooperation from participants, and the questions clear and simple. Each of the questions in the major categories was asked in several different ways throughout the pilot study to obtain reliability.

The SPSET was then pilot tested in a health class consisting of 58 undergraduate students at a Carnegie Extensive Research institution. The health class was selected because of the potential for the inclusion of recreation majors. Yet, the health course was distal to the sequence of the coursework of the recruited participants, which allowed for the determination of instrument validity while controlling for the potential contamination of the study sample. Subjects were asked to identify confusing or inappropriate items, ask questions for clarification, and provide feedback regarding clarity of the questions. It was determined that approximately eight minutes was needed to complete the SPSET questionnaire. Reliability analysis for internal consistency resulted in a Cronbach alpha coefficient of .810 for the overall instrument indicating the questionnaire had an acceptable level of reliability. Cronbach's alpha coefficient results for questions in each category were as follows: knowledge α=.673, seriousness α=.866, and accuracy α=.686. The questions in the value of feedback category resulted in a Cronbach's alpha coefficient

of .476. Upon removal of a question regarding the seriousness level with which instructors review SET data from this category, the Cronbach's alpha coefficient increased to .663. This increase indicated subjects may have misinterpreted the question or the question was placed in the incorrect category. The pilot study sample size and the acceptable levels of the Cronbach's alpha coefficients achieved from the initial analysis indicated that some adjustments to the instrument were necessary to assure that the validity and reliability were suitable for assessing student perceptions of SETs. Upon review of the pilot data and additional feedback from the panel of experts, four unnecessary questions were deleted, four new questions were added, and minor adjustments were made for clarity of word choice. Given the minor adjustments to the instrument, it was decided that no additional pilot data was necessary and that the survey was ready for data collection. The SPSET item total remained at 52 questions.

Procedures for Data Collection

A letter was sent a letter via campus mail and departmental email to all spring 2006 instructors in the recreation department requesting approval to utilize class time to solicit student participation in this study prior to the collection of data. As a result of this solicitation, instructors of 37 courses agreed to participate in the study. Data was not collected from those students who did not wish to participate or had previously participated. Surveys were distributed and collected by the same research team member and returned to the primary researcher for analysis.

Response Rate

Thirty-seven recreation course instructors with an approximate total enrollment of approximately 1,749 students permitted members of the research team to collect data in their courses. Ultimately, the number of subjects was determined by (1) the number of instructors who agreed to allow the researcher to utilize class time to present the survey to the students enrolled in their recreation course, and (2) students who chose to self-select to participate in the survey. This census approach to data collection allowed us to gather data from 523 undergraduate subjects enrolled in at least one course in the recreation department during spring 2006 semester. Since students were highly likely to be enrolled in multiple recreation courses in the semester and based on the total enrollments the researchers were confident that 523 undergraduate was an adequate number of subjects to represent the student opinion and perceptions of SETs.

Analysis

The *Student Perceptions of Student Evaluations of Teachers* survey was returned completed by the 523 students was first analyzed for demographic data. Table 1 below provides the demographic information of the participants involved in this study.

TABLE 1
Demographic Distribution of Participants in the Study

Year in College		**Gender**		**Age**	
Freshmen	88	Male	220	18 & 19	127
Sophomore	134	Female	303	20	147
Junior	156			21	127
Senior	145			22 - 42	122

The 52 item *Student Perceptions of Student Evaluations of Teachers* survey instrument was then analyzed for reliability using Cronbach's alpha as an indicator. The reliability of the instrument was determined to have a Cronbach's alpha of .824, indicating a moderate to high level of instrument stability.

As discussed previously the *SPSET* survey contained items used to measure the latent variables associated with the various aspects of student perceptions of the evaluation process. Although the instrument items were developed to specifically measure different aspects of student perceptions of the evaluation process, further analysis indicated that there was not always adequate item discrimination. This resulted in situations in which some items were measuring more than one latent variable, possibly leading to confounding measurement situations. Given the large number of items of the *Student Perceptions of Student Evaluations of Teachers* instrument, actions were taken to minimize this situation. Our first objective in the analysis process was to determine which items were adequately measuring the latent variables of interest. Factor analysis of the original *SPSET* survey revealed the following latent variables: knowledge of purposes (uses) of SETs, seriousness with which students respond to SETs, perceived value of SET feedback, and accuracy of SET responses. Similarities were noted between the measured variables and the related latent variables in the original survey and the latent variables that were exposed in the process of developing our model. This suggests that our model development is not purely exploratory, proposing a model, but also somewhat confirmatory, confirming a model, (Thompson, 2004), yet different from Ajzen & Fishbein's (2005) TpB model that was used as a guide for the development of the original SPSET survey.

Given the potential confounding variables in the data due to discrepancies in latent variable representation and possible inadequate item discrimination, our first objective was to narrow the items used in analysis to those that were most representative of the constructs being measured. To meet this object a bivariate correlational analysis was conducted using all 52 items of the SPSET instrument. The correlational analysis provided evidence for the relative strength of item relationships, allowing for the preliminary categorization of items. Both the level of significance and correlation values were used as quantitative criteria for categorizing items. Items were considered for grouping if they were significant at the .01 level. If an item's correlation was significant with more than one group, it

was either categorized based on a conspicuously larger correlation value, or was eliminated from analysis.

The categorizations of the 52 items led to the creation of seven distinct groups each representing a latent variable corresponding to the different aspects of student perceptions of teacher evaluations. To assure consistency and content validity a content analysis of the items within each of the groups was conducted. Unexpectedly, the content analysis of the categorized items revealed that two pairs of item groups were essentially measuring the same construct. Given the two overlapping pairs of variables, a decision was made to eliminate one of the groups from each pair leaving a single group of items to represent each of these latent variables. The final result of this analysis was the generation of five groups of items with each group representing one of the five latent variables being considered for further analysis. This allowed the researchers to meet the objective to reduce the potential confounds in the data due to inadequate item discrimination. In addition the selection of five distinct latent variables for further consideration provided a parsimonious effect on the development of a hypothesized model for student perceptions of teacher evaluations.

To simplify further analysis the researchers decided to reduce the number of items representing each latent variable to three, for a total of 15 items. The retention of items was based on content analysis (applying logic and conjecture for justification), and the strength of the correlational values (which provided quantitative justification). The Cronbach's Alpha reliability was again tested using the remaining 15 items producing a value of .658, indicating a drop in stability yet still within the moderate and acceptable range.

Following the initial categorization of items and subsequent conditioning of the data (i.e. reverse coding the negative phrased items), the researchers conducted an exploratory factor analysis (EFA) using SPSS. The EFA was used to examine the retained 15 items to determine if the categorization was truly consistent with the expected latent variable representation. The EFA parameters were principal component extraction, extraction requiring Eigenvalues over 1.0, varimax rotation, and the suppression of loading values less than .50. The EFA successfully extracted five components, each represented by three items, with the loadings accounting for 61.67% of the variance.

The EFA provided further empirical justification for the selection of the items representing the five latent variables. The groups of items and the corresponding latent variables are presented in Table 2.

TABLE 2

Item Placement of the Student Perceptions of Student Evaluations of Teachers Questionnaire and Corresponding Latent Variable

Latent Variables (Factors)	**88 Items from Student Perceptions of Student Evaluations of Teachers (SPSET)**
IMPACT – of student opinion	1. Most administrators (Deans, Chairs, University Officials) do not care about student evaluations of instructors. 2. How seriously do you think the majority of administrators take teacher evaluations? 3. Most instructors do not care about student evaluations of their teaching ability.
PROCESS – in which students complete evaluations	1. How often do you read each question on teacher evaluations before you respond to the question? 2. How often do you take your time when completing teacher evaluations? 3. On teacher evaluations, I often rate all the answers the same rather than read all of the questions (for example, you put "A" for every answer).
ACCURACY – of student responses	1. How often do you feel your teacher evaluation responses are higher than they actually should be? 2. How often would you say other students in your classes (whose opinion you value) give a teacher higher marks than the student thought the teacher deserved? 3. How often do you give a teacher higher marks than you think he/she deserved?
AUDIENCE – who uses/sees the evaluations	1. Who do you think reads/reviews teacher evaluations? a) Dean of the School (the boss of the chair & instructor) b) Other University Officials/Administrators c) Department Clerical Staff
PURPOSE – of evaluations	1. Teacher evaluations are used to: a) Give instructors pay raises based on teaching ability (merit). b) Make decisions about the promotion of an instructor. c) Make decisions about tenure of an instructor (job security).

The Measurement Model

Following the conditioning and reorganization of the data, a measurement model confirmatory factor analysis was completed to determine the acceptability of the factor structure. This measurement model analysis provided the support necessary to develop a hypothesized model representing student perceptions of Student Evaluations of Teachers (SETs). The researchers conducted the confirmatory factor analysis of the measurement model using EQS 6.1 and consulted Byrne (2006) for interpreting and presenting the results. The measurement model was constructed by establishing correlations between the five latent variables, with each of these factors represented by the three previously established measured variables (see Figure 1). Both the fit indices and the factor correlations were used as empirical indicators for justification in the formation of a hypothesized model.

Multiple fit indices measures were assessed to determine the data including the normed chi-square (chi-square/df), comparative fit index (CFI), goodness of fit index (GFI), the adjusted goodness-of-fit index (AGFI), root mean square (RMS) and the root mean-square error of approximation (RMSEA) as suggested by Byrne (2006) and Kline (2005).

The initial measurement model including examining the correlations between the latent variables indicated a moderate to good fit. The results revealed that most fit statistics were above or nearing the suggested threshold values for an acceptable model (Normed chi-square=1.71, CFI=.958, GFI=.966, AGFI=.949, Standardized RMS=.042, RMSEA=.037, and the 90% Confidence Interval of the RMSEA between .026 and .048).

An assessment of the fit indices indicated that the inclusion of all five factors in a model would be justified. Further, the resulting correlations, presented in Table 3, revealed a reasonable level of correlation between the factors, indicating that there was an acceptable level of consistency in the measures, and no obvious occurrences of collinearity of measurement.

TABLE 3

Correlations Among The Five Study Latent Variables

	IMPACT	PROCESS	ACCURACY	AUDIENCE	PURPOSE
IMPACT	1.00				
PROCESS	.187	1.00			
ACCURACY	.156	.242	1.00		
AUDIENCE	.125	-.036	.030	1.00	
PURPOSE	.346	.162	.061	.041	1.00

A Proposed Model

Based on the outcome of the factor analysis, the content analysis of the factors, and the literature, it is apparent that student perceptions of teacher evaluations are a complex interaction of variables. Our analysis indicates that perceptions

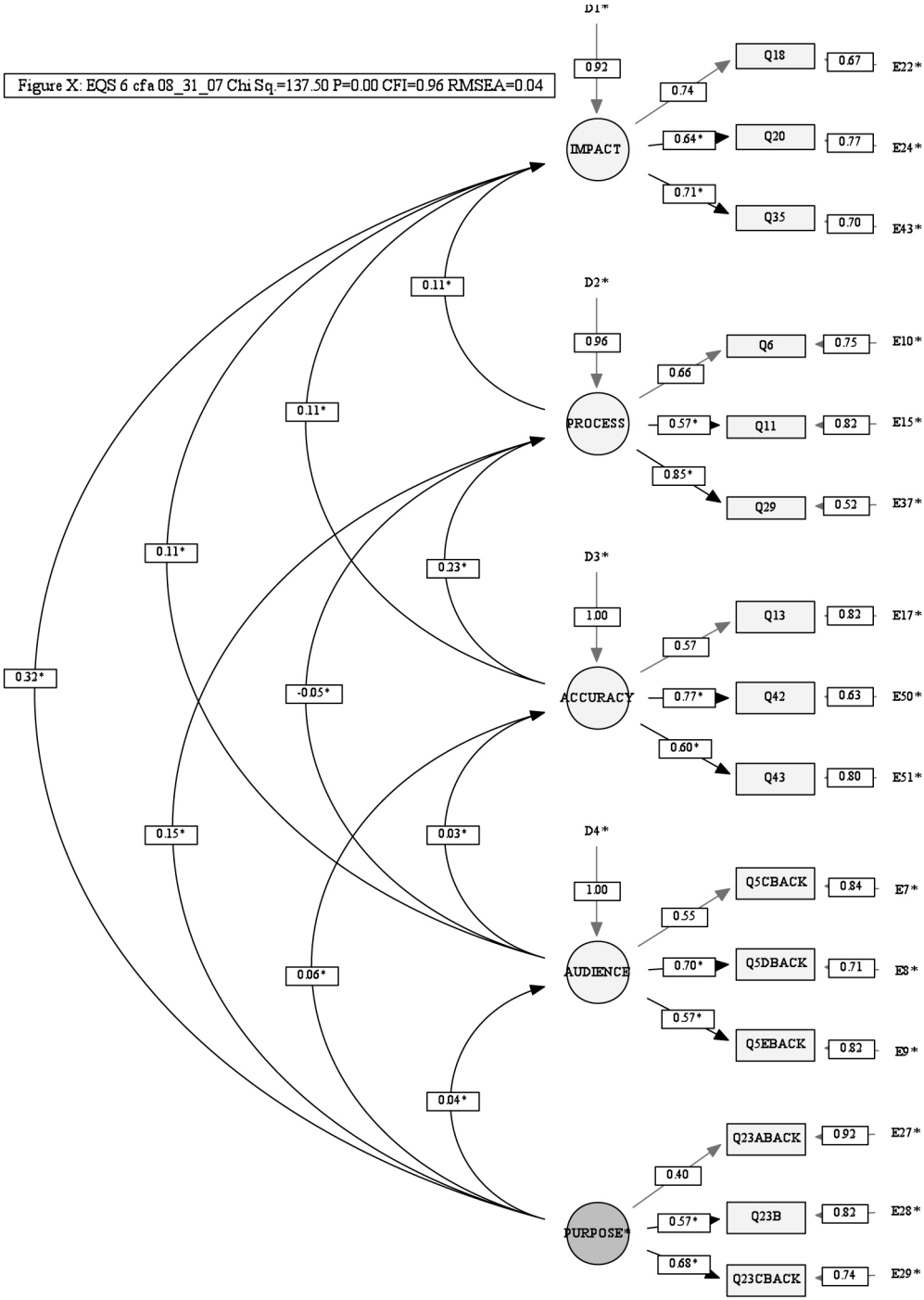

Figure 1. Factor loadings and standard errors are shown for each item and correlations are shown between factors.

of evaluation includes factors related to student opinion of the impact of the evaluations, the process in which they complete the evaluations, the accuracy of which their opinion is interpreted, their perception of the intended audience, and their perception of the purpose of the evaluations. This provides further support for the position that student behaviors are influenced by perceptions of the emphasis and manner in which student opinion is interpreted by the administration and faculty, intended audience, the impact of the evaluations, and the purpose of the evaluations. Taking these factors into consideration, the researchers are proposing the following model as representative of student perceptions of student evaluations of teachers (see Figure 2).

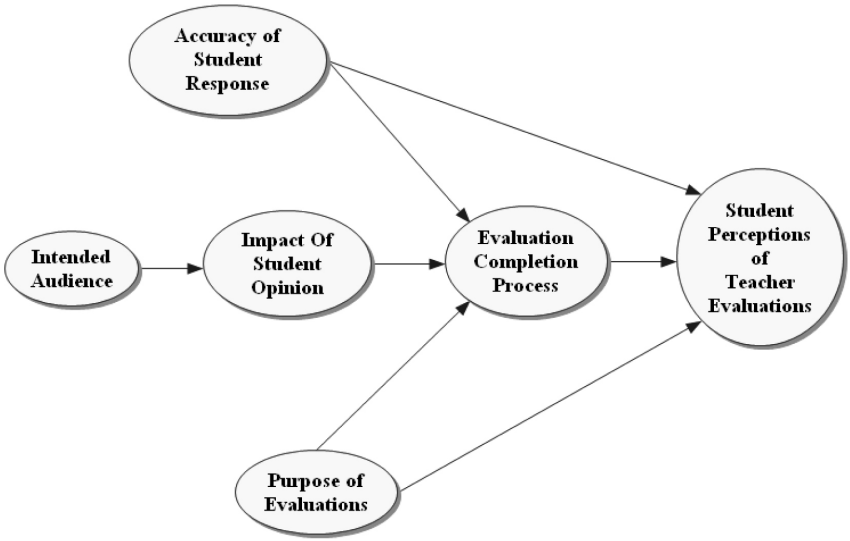

Figure 2. Our hypothesized model for student perceptions of student evaluations of teachers.

In this hypothesized model the process by which students complete teacher evaluations is directly influenced by their perceptions. Student perceptions of the purpose of the evaluations and the accuracy of evaluation interpretation are both direct and indirect measures of their perceptions. Further, the process is directly influenced by the student perceptions of impact which has an antecedent of student perceptions of the intended audience, which are indirect measures of their perceptions. The hypothesized model reflects the notion that the factors measured can be used directly and indirectly to predict student perceptions of teacher evaluations.

The confirmation of the proposed model will require the collection of additional related data that corresponds to the latent variables posited to represent student perceptions of SETs. This data could be collected using the modified

SPSET questionnaire or a different questionnaire developed to measure the same variables. The accuracy and application of this model depends on the reproducibility of acceptable levels of additional confirmatory factor analysis.

Discussion

This work contributes to SET research by providing an empirical model for the student perspective and by investigating an under-explored academic area of research within SET literature. Notably, this is the first such study investigating student perceptions who are recreation majors. The analysis of the 52-item *SPSET* indicated that fewer items can be used to capture the range of variables influencing student completion of SETs. Of the 52 items of the *SPSET* the final analysis used only 15 items, in contrast to existing research that suggests there are many factors influencing how students complete SETs (Anderson & Miller, 1997; Cardy & Dobbins, 1986; Centra & Gaubatz, 2000; Chonko, Tanner, & David, 2002; Clayson, 1999; Crader & Butler,1996; Dukes & Victoria, 1989; Dunegan & Hrivnak, 2003; Feldman, 1993; Grimes, Millea, & Woodruff, 2004; Marlin, 1997; Smith & Carney, 1990; Williams & Ceci, 1997). This study determined that the collection of five variables that are representative of student perceptions of SETs may be attainable with a relatively brief instrument. The three measurement variable items selected to represent each of five latent variables revealed sufficient discrimination of measurement while maintaining an acceptable level of reliability. Further, content analysis supports the construct validity of the items as important considerations of student perceptions of SETs.

Existing literature supports this model. Specifically, both Smith and Carney (1990) and Marlin (1987) found perceived value of feedback (as labeled in the *SPSET* model as Impact of Student Opinion) to be a contributing factor of student responses on SETs noting that students did not perceive that their SETs responses were valued by instructors or administrators. An effort to relay specific uses and/or examples of how students' SET data has been or will be used may increase student's perceptions of how their SET data is valued by both instructors and the administrators and therefore should be included in any model representing the SET process.

Level of Seriousness

The level of seriousness students hold when completing end-of-course evaluations (Process with which Students Complete SETs in model) is a re-occurring theme in SET literature (Dunegan & Hrivnak, 2003, Jacobs 1987; McKeachie, 1997; Marlin, 1987; Martinson & Ryan, 1981; Sheehan & DuPrey, 1999; Simpson & Siguaw, 2000; Smith & Carney, 1990; Spencer & Schmelkin, 2002). This model reveals continued support for the potential impact students' attitude toward SETs completion may have on SET data. In short, while the majority of research found that students perceive they complete SETs seriously (Smith & Carney, 1990; Marlin, 1987; Martinson, & Ryan, 1981), evidence exists that at times students hold a careless approach in completing evaluations (Dunegan & Hrivnak, 2003). However, further exploration into the nature of the motivating factors exposes

multifaceted influences on the SET process. This model begins to depict how various factors may interact to influence student SET responses.

Accuracy of Student Responses

The accuracy of student responses (labeled under the same name in the *SPSET* model) is central to SET research. Marsh (1984) asserts that the validity of SETs has been well established; however, research continues to explore this area. This analysis revealed that student perceptions of the use and importance of SETs is linked to the accuracy of their responses. The interaction of perceptions and accuracy suggests a need to educate students on the importance of taking the time to provide honest feedback and be thoughtful in their responses to teacher evaluations.

Knowledge of Uses of SETs

Further, students were uncertain as to the audience (who uses/sees evaluations) of SETs. This model revealed most were aware that instructors and chairs reviewed SET data, but were unclear if the Dean or University officials reviewed the data. Additionally, student perceptions of the purpose of evaluations (as labeled in the *SPSET* model as Purposes/Uses of SETs) were confusing for students. Marlin (1987), reports that students' perceive end-of course evaluations as having little impact on faculty careers. Marlin asserts that students perceive that nobody pays attention to the outcome evaluations with little consideration being given to student opinion by faculty and administrators. Our findings are consistent with the assertions of Marlin, indicating that over time student perceptions that SETs are of low consideration and interest by faculty and administrators. Our results indicate student perceptions of faculty and administrator consideration of SETs has remained an influential issue. An increased awareness in the purposes, uses, and the value placed on SET responses would provide impetus for educating students about the seriousness and impact of SETs. Our proposed model provides a framework increasing student awareness of SETs. Such an effort may ultimately strengthen the accuracy of SET data in general. Understanding how the factors influencing SET interaction is the essence of our proposed *SPSET* model.

This study has lent further evidence to the fact that student attitudes and perceptions of faculty and administrators' interest in their opinions influences engagement in the evaluation process. Sheehan and DuPrey (1999) report that students exhibit different behaviors and provide different answers if they perceive end-of-course evaluations as inconsequential rather than an opportunity to provide honest and accurate feedback of faculty performance and course organization. This position is further supported by our finding that indicates the level of seriousness students hold when completing end-of-course evaluations (Process with which Students Complete SETs in the model) was an important factor and needs to be included into a model of student perceptions of SETs.

Conclusion

The proposed model developed from a desire to have a theoretical foundation from which we could base analysis of student perceptions of and completion behaviors of SETs. The SPSET model includes the variables identified as indicators of the student aspects of the SET process. Although the model was intended to apply the TpB (Ajzen & Fishbein, 2005), the outcome indicates that recreation student related SET variables may be unique and do not conform as predicted to the TpB and, therefore, are in need of additional investigation. Further, if the SPSET model is representative and supported by confirmatory factor analysis, then examination of additional individual student characteristics such as gender, age, levels of education, may need to be explored as variables of the student SET perceptions model.

Implications

Faculty and administrators can better interpret and assess student behaviors and outcome results on SETs by measuring attending to the factors that influence recreation students' overall perceptions or levels of understanding of this process. Latent variables, such as perceptions of the use of SET, cannot be measured directly and therefore are assessed using items representative of these variables. The proposed model identifies the latent variables and corresponding measurement items that interact and influence student perceptions of end-of-course teacher evaluations.

Zuefle (2004) maintained that some students hold lower course expectations along with the common perception that recreation classes lack rigor. These perceptions (arguably misperceptions) may influence student SETs responses in a manner that results in outcome that differ from SETs from other academic courses. The proposed model depicts the major variables that interact in end-of-course teacher evaluations, specifically in the academic discipline of recreation. Now that these latent variables have identified and their interaction determined, the investigation of recreation students' perceptions of SETs needs to continue. The SPSET model will assist recreation faculty and administrators in developing additional items specifically designed to measure these latent variables and assess overall student perceptions. Through an enhanced understanding of the major latent variables that influence student perceptions, recreation faculty and administrators can better interpret and access student behaviors and performance on SETs, elucidating the outcomes in perspective of the process.

While the SPSET model depicts the major latent variables that pertain to students enrolled in recreation courses it has potential application to other academic disciplines. Most notably, the SPSET model enables the ability to perform a comparative study may reveal student perceptions indeed differ in recreation courses as opposed to other academic courses. Accuracy of SET responses is dependant upon a better understanding of student perceptions of SETs. The proposed model depicts it is important that faculty ensure students enrolled in recreation courses are well informed regarding the uses and purposes of SETs and the need for stu-

dent to take completing SET seriously as their opinions are valued by the faculty and administration. Recreation student possess a range of perceptions influencing their completion of SETs suggesting there is justification for maintaining perspective and remaining cautious in the interpretation of SETs among recreation faculty and administrators.

Future Research

Additional areas that warrant further investigation include the development of a parsimonious and conciliatory instrument that could be widely applied to assess student perceptions of SETs. Other variables may need to be included such as region in which an institution is located, course of study of the students, tier of the institution, and student demographics, which all may prove to be important considerations in the study of the student perceptions of SETs. Additionally, other forms of data such as interviews, focus groups, or open ended questions may need to be considered to provide explanations of the reasoning, interpretation, and justification that student use when answering forced response and Likert scale SET questionnaires. Finally, given the dynamic nature of student perceptions their perceptions of SETs should be continually monitored for consistency and for change. For example, behavior and belief systems of students vary widely among the 'baby boomers', 'generation X', and 'generation Y' (Smola & Sutton, 2002).

Future studies might include interviews of students which may provide additional insight into their thoughts and interpretations of the specific aspects of teacher evaluations. Data collection from a wider range of students including non-recreation majors and students from other institutions would provide additional confirmatory data for the proposed model. If setting, region and other university specific variables are correlated with student perceptions of SETs the restriction to a single institution may not capture the complexity and variability of the process. Finally, the analysis process was used in the formation of the latent variables post-hoc, and even though planned a-priori, the data may not be as predicted which constrained the outcome to a predicted model and limited the opportunity to confirming the model. However, now that the model has been developed there is opportunity for additional confirmatory analysis. The confirmation of the model will have to come from the collection of additional data that is representative of the variables the researchers have determined to be interacting as indicators of students' perceptions of SETs. The confirmation of the application of the model may also be expanded to different population of students such as graduate students, other majors, and on-line courses.

References

Ajzen, I. (2002). University of Massachusetts, Ajzen Website. *Constructing a TpB questionnaire: Conceptual and methodological considerations*. Retrieved September 22 2005, from http://www.people.umass.edu/aizen/pdf/tpb.measurement.pdf.

Ajzen, I., & Fishbein, M. (2005). The influence of attitudes on behavior. In D. Albarracín, B. T. Johnson, & M. P. Zanna (Eds.), *The handbook of attitudes* (pp. 173-221). Mahwah, NJ: Erlbaum.

Anderson, K., & Miller, E. D. (1997). Gender and student evaluations of teaching. *PS, 30*(2), 216-219.

Byrne, B. (2006). *Structural equation modeling with eqs: Basic concepts, applications, and programming.* Thousand Oaks, CA: Sage.

Cardy, R. L., & Dobbins, G. H. (1986). Affected appraisal accuracy: Liking as an integral dimension of evaluating performance. *Journal of Applied Psychology, 71*(4), 672-678.

Centra J. A., & Gaubatz, N. B. (2000). Is there gender bias in student evaluations of teaching? *Journal of Higher Education, 71*(1), 17-33.

Chamberlin, M. S., & Hickey, J. S. (2001). Student evaluations of faculty performance: The role of gender expectations in differential evaluations. *Educational Research Quarterly, 25*(2), 3-14.

Chen, Y. & Hoshower, L. B. (1998). Assessing student motivations to participate in teaching evaluations: An application of expectancy theory. *Issues in Accounting Education*, *13*(3), 531–549.

Chonko, L. B., Tanner, J. F., & David, R. (2002). What are they thinking? Students' expectations and self-assessments. *Journal of Education for Business, 77*(5), 271-279.

Clayson, D. E. (1999). Students' evaluation of teaching effectiveness: Some implications of stability. *Journal of Marketing Education, 21*(1), 68-75.

Costin, F., Greenough, W. T., & Menges, R. J. (1971). Student ratings of college teaching. *Review of Educational Research, 41*(5), 511-535.

Crader, K. W., & Butler, J. K. (1996).Validity of students' teaching evaluation scores: The Wimberly-Faulkner-Moxley questionnaire. *Educational and Psychological Measurement, 56*(2), 304–314.

Doyle, K. O. (1983). *Evaluating teaching.* Lexington, MA: Lexington Books.

Dukes, R. L., & Victoria, G. (1989). The effects of gender status, and effective teaching on the evaluation of college instruction. *Teaching Sociology, 17*(4), 447-457.

Dunegan, K. J., & Hrivnak, M. W. (2003). Characteristics of mindless teaching evaluations and the moderating effects of image compatibility. *Journal of Management Education, 27*(3), 280-303.

Feldman, K. A. (1993). College students' views of male and female college teachers: Part I – evidence from students' evaluations of their classroom teachers. *Research in Higher Education, 34*(2), 151-211.

Grimes, P. W., Millea, M. J., & Woodruff, T. W. (2004). Grades-who's to blame? Student evaluation of teaching and locus of control. *Journal of Economic Education, 35*(2), 129-147.

Hobson, S. M., & Talbot, D. M. (2001). Understanding student evaluations. *College Teaching, 49*(1), 26-31.

Jacobs, L. C. (1987). University faculty and students' opinions of student ratings. *Indiana Studies in Higher Education* (No. 55). Bloomington, Indiana University: Bureau of Evaluation and Testing.

James, R. (2002). Students' changing expectations of higher education and the consequences of mismatches with reality. In P. Coaldrake (Ed.), *Responding to Student Expectations*. Paris: OECD.

Keller, G. (2001). The new demographics of higher education. *The Review of Higher Education, 24*(3), 219-235.

Kline, R. (2005). *Principles and practice of structural equation modeling*. New York: Guilford Press.

Koushki, P. A., & Kuhn, H. A. J. (1982). How reliable are student evaluations of teachers? *Engineering Education, 72*(3), 362-367.

Marlin, J. W. (1987). Student perceptions of end of course evaluations. *Journal of Higher Education, 58*(6), 704-716.

Marsh, H. W. (1984). Students' evaluation of university teaching: Dimensionality, reliability, validity, potential biases, and utility. *Journal of Educational Psychology, 76*(5), 707-754.

Marsh, H. W., & Dunkin, M. J. (1992). Students' evaluations of university teaching: A multidimensional perspective. In J.C. Smart (Ed.), *Higher education: Handbook of theory and research* (pp.143-223). New York: Agathon Press.

Martinson, D. L., & Ryan, M. (1981). What do J-students think about teacher evaluations? *Journalism Educator, 36*(2), 53-55.

McKeachie, W. J. (1997). Student Ratings: The validity of use. *American Psychologist, 52*(11), 1218-1225.

Schmelkin, L., Spencer, K., & Gellman, E. (1997). Faculty perspectives on course and teacher evaluations. *Research in Higher Education, 38*(5), 575–592.

Sheehan, E., & DuPrey, T. (1999). Student evaluations of university teaching. *Journal of Instructional Psychology*, *26*(3), 188–193.

Simpson, P. M., & Siguaw, J. A. (2000). Students' evaluation of teaching: An exploratory study of faculty response. *Journal of Marketing Education, 22*(3), 199-213.

Smith, M. C., & Carney, R. N. (1990, April). *Students' perceptions of the teaching evaluation process*. Paper presented at the annual meeting of the American Educational Research Association, Boston, MA. Retrieved September 22, 2005, from http://www.cedu.niu.edu/~smith/Unpubs/AERA90.pdf.

Smola, K. D., & Sutton, C. D. (2002). Generational differences: Revisiting generational work values for the new millennium. *Journal of Organizational Behavior, 23*(4), 363–382.

Spencer, K. J., & Schmelkin, L. P. (2002). Student perspectives on teaching and its evaluation. *Assessment & Evaluation in Higher Education, 27*(5), 397-409.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis*. Washington, DC: American Psychological Association.

Wachtel, H. K. (1998). Student evaluation of college teaching effectiveness: A brief review. *Assessment & Evaluation in Higher Education, 23*(2), 191-211.

Wilhelm, W. B., & Comegys, C. (2004). Course selection decisions by students on campuses with and without published teaching evaluations. *Practical Assessment Research & Evaluation, 9*(16). Retrieved September 29, 2005, from http://www.pareonline.net/getvn.asp?v=9&n=16.

Williams, W. M., & Ceci, S. J. (1997). How'm I doing? Problems with students' ratings of instructors and courses. *Change, 29*(5), 12-23.

Wulff, D. H., Staton-Spicer, A. Q., Hess, C. W., & Nyquist, J. D. (1985, August). The student perspective on evaluating teaching effectiveness. *ACA Bulletin, 53*, 39-47.

Zuefle, D. M. (2004). Do folks in the recreation field get the last laugh? *Parks & Recreation, 39*(4), 14-16.