

Fall 2011

Nonresponse bias in online course evaluations

Cassandra Jones
James Madison University

Follow this and additional works at: <https://commons.lib.jmu.edu/diss201019>



Part of the [Psychology Commons](#)

Recommended Citation

Jones, Cassandra, "Nonresponse bias in online course evaluations" (2011). *Dissertations*. 61.
<https://commons.lib.jmu.edu/diss201019/61>

This Dissertation is brought to you for free and open access by the The Graduate School at JMU Scholarly Commons. It has been accepted for inclusion in Dissertations by an authorized administrator of JMU Scholarly Commons. For more information, please contact dc_admin@jmu.edu.

Nonresponse Bias in Online Course Evaluations

Cassandra R. Jones

A dissertation submitted to the Graduate Faculty of

JAMES MADISON UNIVERSITY

In

Partial Fulfillment of the Requirements

for the degree of

Doctor of Philosophy

Assessment and Measurement

December 2009

Dedication

This manuscript is dedicated to my family and friends. To my mother, thank you for making it seem as though anything is possible. To my sisters Belinda and Leona, thank you for being there when I needed a little escape and support. To my friend Charlie, thank you for the reenergizing conversations.

Acknowledgements

I want to thank my committee for guiding me throughout this process. I must especially thank my chair and advisor, Donna Sundre, for the endless supply of encouragement.

This dissertation would not have been possible without the support of Marymount University, especially Michael Schuchert, the Executive Director of Institutional Effectiveness .

Table of Contents

Dedication.....	ii
Acknowledgements.....	iii
List of Tables.....	v
Abstract.....	vi
Chapter 1 – Introduction	1
Chapter 2 – Review of Literature.....	5
Definitions.....	7
Online surveys.....	7
Approaches to examining nonresponse.....	9
Theoretical perspectives on nonresponse.....	12
Online course evaluations.....	20
Summary of research.....	22
The current study.....	24
Chapter 3 – Methods.....	
Participants.....	26
Instruments.....	26
Courses.....	26
Procedure.....	27
Analysis.....	28
Chapter 4 – Results.....	34
Predicting course evaluation ratings.....	35
Predicting course evaluation participation.....	39
Determining nonresponse bias.....	41
Chapter 5 – Discussion.....	44
Research findings.....	46
Theoretical perspectives.....	47
Limitations.....	49
Future studies.....	50
Conclusions and implications.....	51
Appendix A.....	53
Appendix B.....	56
References.....	58

List of Tables

Table 1: Variables included in study	29
Table 2: Cross-classified model.....	30
Table 3: Response rate of study participants	34
Table 4: Cross-classified random effects model predicting online course evaluation ratings.....	37
Table 5: Hierarchical generalized linear model estimating predictors of online course evaluation participation.....	42

Abstract

Recently more universities have started administering course evaluations online. With the process no longer in the classroom, some students decide not to complete their course evaluations during their own time, resulting in concerns about online course evaluation results being biased because of lack of response. This study examined course evaluation results at a small diverse mid-Atlantic Catholic university. A cross-classified random effects model was used to capture student responses across all of their courses.

Nonresponse bias was examined by determining predictors of online course evaluation ratings and participation. Variables predicting both participation and ratings were considered to be a potential source of nonresponse bias. It was found that gender, ethnicity, and final course grade predicted online course evaluation participation. Only final course grade predicted online course evaluation ratings.

Chapter 1

Introduction

A typical end of the semester ritual at colleges includes course evaluations. A professor gives a nice talk about the semester and perhaps hands out candy, and then asks a volunteer to distribute forms to students and deliver the completed forms to the designated drop-off location. With advances in technology, sometimes this ritual includes students completing their course evaluations online during their own time instead of the in the classroom.

Online course evaluations have many advantages such as (a) taking up less class time, (b) being cheaper to administer, (c) not as influenced by professors, and (d) producing more written comments (Anderson, Cain, & Bird, 2005; Sorensen & Reiner, 2003). The disadvantages include students needing access to computers and producing low response rates (Anderson et al., 2005). Low response rates (nonresponse) and their potential to lead to nonresponse bias is the focus of this study.

Research problem

Course evaluation results are used by universities for course improvement and often in making tenure and promotion decisions. Because research has shown that online course evaluation response rates are low, a concern about nonresponse bias exists (Avery, Bryant, Mathios, Kang, & Bell, 2006; Heath, Lawyer, & Rasmussen, 2007; Leigle & McDonald, 2005; Thorpe, 2002). Bias in a general sense could be thought of as “systematic error” (Camilli, 2006, p.225). When thinking about nonresponse bias, this concern refers to potential respondents with certain characteristics not responding to the online course evaluation, and their ratings being different from the individuals who did

respond (Groves, Couper, Pressor, Tourangeau, Acosta & Nelson, 2006). If online course evaluation results are going to be used for decisions related to curriculum and personnel decisions, the results should represent the students in the class, and not a specific subset with potentially different opinions.

Understanding current response rates with online surveys and reasons for potential response behavior can be gleaned from studies from survey research.

Collectively, survey research in these areas provides three basic expectations for online course evaluation results.

First, meta-analyses examining online surveys found that when compared to other methods of survey administration, online survey rates were lower (Manfreda, Bosnjak, Berzelak, Haas & Vehovar, 2008; Shih & Fan, 2008). Research about online surveys in college settings also found low response rates, and statistically different rates of responses based on different student characteristics such as gender and ethnicity (Sax, Gilmartin, & Bryant, 2003; Sax, Gilmartin, Lee, & Hegedorn, 2008). These findings set the expectation for low response rates with online course evaluations and possible nonresponse bias.

The second expectation is that a student's willingness to complete an online course evaluation may be related to opinions about the course or general feelings about the university. Research examining organizational citizenship behavior suggested that individuals' opinions about an organization impacted their willingness to respond to organizational surveys (Spitzmuller, Glenn, Barr, Rogelberg, & Daniel, 2006; Spitzmuller, Glenn, Sutton, Barr, & Rogelberg 2007). With online course evaluations,

this suggests that completing an online course evaluation is influenced by student experiences at the university and within the classroom.

The third expectation is that response rates can be increased if one knows the important components of a survey to modify or incentives to offer. Online course evaluation studies suggest that unless an appropriate incentive can be found (Dommeyer, Baum, Hanna, & Chapman, 2004), students are not willing to complete an online course evaluation just because they like a course (Fidelman, 2007).

Online course evaluation studies have shown that certain groups of students complete online course evaluations at higher rates than others (Avery et al., 2006; Dommeyer et al., 2004, Fidelman, 2007; Layne, DeCristoforo, & McGinty, 1999; Liegle & McDonald, 2005; Thorpe, 2002). Identifying populations that may need additional encouragement to respond is a first step towards increasing response rates. The research on online course evaluation nonresponse has generally examined a few courses across a university or within an academic school. No study has considered that students take multiple courses at once, and generally complete their evaluations at once, not individually. Considering students in courses as a cross-classified structure will provide better estimates of predictors of student ratings and the probability of their response. Also, many of the studies have occurred at large universities. Examining nonresponse bias in online course evaluations at a small private university contributes to the research by investigating whether nonresponse and nonresponse bias issues at universities with specific characteristics or at institutions of higher education in general.

Purpose of study

The purpose of this study is to determine if nonresponse bias exists in online course evaluations. Previous studies examining nonresponse and nonresponse bias in online course evaluations have not considered that students take multiple courses throughout a semester, and thus are asked to complete multiple course evaluations. Previous studies have also occurred at larger universities than at universities such as the one included in this study. This study will use a conceptualization of nonresponse bias provided by Groves et al. (2006), which prescribes that in order for nonresponse bias to exist, characteristics of the nonrespondents must be related to the probability of response, and the variable of interest. In the current study, the variable of interest is course evaluation ratings. The major research question of this study is “Is there nonresponse bias in online course evaluations?” To examine nonresponse bias, this study will address the following questions:

1. Are student characteristics such as gender, ethnicity, cumulative GPA, course grade, institutional credits, and total credits related to online course evaluation ratings?
2. Is course size related to course evaluation ratings?
3. Are student characteristics such as gender, ethnicity, cumulative GPA, course grade, institutional credits, and total credits related to responding to the online course evaluation?
4. Is course size related to responding to an online course evaluation?

Chapter 2

Review of Literature

Completing a course evaluation during the last two weeks of the semester has been a ritual in the American college classroom since the 1920s (d'Apollonia & Abrami, 1997). Course evaluations have been used by students as an opportunity to express opinions about their experiences in the course, by faculty as a tool for course improvement, and by administrators as evidence when making decisions about tenure and promotions.

The administration of this process remained consistent until the 1990s when a few universities began experimenting with conducting course evaluations online. Although the exact numbers of universities conducting course evaluations online is unknown, efforts have been taken to provide an estimate. Hoffman (2002) conducted a study to discover how many universities were using online course evaluations. Seventeen percent (17%) of the sample of 500 reported using online course evaluations in some capacity for face-to-face courses.

Based on the usage of technology for many processes in everyday life, it is easy to anticipate the percentage of universities using online course evaluations increasing exponentially from 2002 to 2009, especially considering the advantages of an online process. Some of the advantages of online course evaluations over paper and pencil administration include: a) lower costs, b) no use of classroom time for the evaluation, c) students write more comments, and d) students are less influenced by faculty (Sorensen & Reiner, 2003). When surveyed, faculty acknowledge some of these advantages, but

also express a few concerns, specifically the lack of student participation (Anderson et al., 2005; Dommeyer, Baum, Chapman & Hanna, 2002).

Once universities started experimenting with online course evaluations, they found that large percentages of students were not participating, which prompted a concern about the validity of the ratings (Hardy, 2003; Johnson, 2003; Layne et al., 1999). Research has shown that even with lower response rates, online course evaluation ratings are equivalent to paper and pencil ratings (Dommeyer et al., 2004; Leigle & McDonald, 2005; Thorpe, 2002). Although the ratings are not substantially different, there are still concerns about the lack of response from students, and implications for using the results from online course evaluations to make important changes in courses and decisions impacting the careers of instructors by administrators. Essentially, there is concern about potential nonresponse bias in online course evaluations.

The focus of this study is on the problem of nonresponse in online course evaluations and its potential to lead to biased results and potentially inappropriate inferences about courses and teaching. In addition, course evaluation results could have positive or negative impacts on an instructor's career as they are often used in personnel decisions (Ory, 2000). In essence, course evaluation results impact students and the instructors who receive them.

Because online course evaluations are surveys, this literature review will draw from research on nonresponse and nonresponse bias in survey research, theoretical perspectives of nonresponse, and, finally, to the still developing body of research pertaining to online course evaluations. This literature review will start with providing definitions of nonresponse and nonresponse bias.

Definitions

Nonresponse. There are two types of nonresponse: item and unit (Dillman, Eltinge, Groves, & Little 2002). Unit nonresponse occurs when a potential survey participant does not respond to an entire survey. Item nonresponse occurs when the survey recipient answers some of the items on a survey. When thinking about online course evaluations, the focus is on unit response, because the primary concern is about students not completing any item on the evaluation.

Nonresponse bias. This type of bias occurs when survey recipients with specific characteristics do not respond to the survey, and the reason for nonresponse is related to the topic being surveyed (Groves, 2006). With online course evaluations, the concern is that students who refrain from completing the survey are doing so for a reason that would impact course ratings. The reason for refraining would be the source of bias. When thinking about nonresponse and nonresponse bias in online course evaluations, online survey research gives a framework for what to expect with online course evaluations.

Online surveys

Online surveys are attractive because they are generally cheaper and less laborious to use than paper and pencil surveys. With online surveys, there is no concern about the cost of stamps and paper; however, other concerns emerge. The largest issue with online surveys is the lack of response. Because online course evaluations are online surveys, components of online research also relate to online course evaluations. Manfreda et al. (2008) and Shih and Fan (2008) conducted meta-analyses that compared response rates of surveys administered online or in paper. Both studies focused on the following

four features in exploring response rates in online surveys: (a) overall response rates by mode, (b) population type, (c) incentives, and (d) reminders.

First, both studies concluded that overall response rates were lower for online surveys than for surveys administered by mail, fax, face-to-face or telephone. Manfreda et al. (2008) found that across all 45 studies included in the meta-analysis, response rates for online surveys were 11% lower. Averaging across all 37 studies, Shih and Fan (2008) found response rates in online surveys 10% lower.

The two meta-analyses differed in their findings about the influence of population type on response rates. Manfreda et al. (2008) found that online survey response rates from the student population were 6% lower than surveys administered using other forms of administration. With Shih and Fan (2008), online response rates were 3% higher for college students than other populations. This finding was attributed to college students having more access to technology and being more comfortable with technology (Shih & Fan, 2008, p. 259). This finding provides a small amount of optimism for administering online surveys to college students.

The third feature is the influence of incentives on response rates. Manfreda et al. (2008) did not find significant differences based on use of incentives. One weakness in their meta-analysis was the inclusion of only three studies that used incentives and 42 studies that did not use incentives. They explained this as a problem by stating, “some of these non-significant results could have been caused by the low number of cases in certain categories, attenuating their statistical power” (Manfreda et al. 2008, p. 96). Shih and Fan (2008) reached a similar conclusion about the influence of incentives on response rates, and they had a sufficient number of studies included in the incentive and

no incentive conditions. This conclusion suggests that incentives may not be a viable tool in reducing nonresponse in online course evaluations.

Finally, both studies examined the influence of reminders on response rates. Manfreda et al. (2008) found significant differences in response rates by the number of contacts. Response rate differences increased between online surveys and other types of surveys as the number of contacts increased. Unfortunately, response rate differences were not in favor of online surveys. Shih and Fan (2008) found similar results with the difference in response rates for online and mail surveys increasing with the inclusion of one reminder. They concluded that reminders with an online administration could be equated to spam, which could influence the response rate. With online course evaluations, this finding creates the question of how many reminders are enough or too many.

In conclusion, both meta-analyses concluded that response rates for surveys administered online are lower than for surveys administered in other formats. Incentives and reminders were not found to be helpful tools in increasing online survey response rates. The only positive finding for online surveys came from Shih and Fan (2008), who concluded that college students responded at a higher rate than other populations to online surveys. The results from these meta-analyses emphasize the necessity of understanding how to examine nonresponse, because online survey response rates were consistently lower than other paper-based approaches.

Approaches to examining nonresponse

Many approaches exist for examining differences in nonresponse. These approaches describe how to gather data to compare respondents and nonrespondents, and

use the findings to see if these findings relate to the probability of the individuals responding to the survey, and the impact on final survey results. Sometimes the impact of survey results can be quantified and other times the impact must be examined and explained within the context of the study. Each approach provides a different perspective and comes with its own strengths and weaknesses. Four methods have been identified in the literature and each will be described briefly below.

One method is examining response rates across subgroups. Study participants are classified into groups. For example, Abraham, Maitland, and Bianchi (2006) used this approach to determine that “busy” and “not busy” people did not differ in their response rates. Thorpe (2002) reported differences in response rates by gender. The weakness of this approach is that it assumes that the subgroups are the only causes of nonresponse (Groves, 2006, p.654).

Another approach uses other information about the respondents and nonrespondents to examine differences in nonresponse patterns. Groves (2006) calls this “using rich sampling frame data.” (p. 654). This approach has also been referred to as record linkage (Porter & Whitcomb, 2005, p. 130). An advantage of this approach is that the same information is available for respondents and nonrespondents. In a higher education setting, one would use administrative data such as gender, grade point average (GPA), or number of credits completed. The disadvantage is that the information may not include all variables needed for detecting nonresponse bias. Also, as noted by Groves (2006), the matched data is subject to measurement error. For example, if one uses GPA as a variable, GPA is only as valid as the procedures used by the university to calculate and update the system to reflect an accurate and up to date student GPA. Within the

online course evaluation literature, Fidelman (2007) used this approach to examine nonresponse rate by using characteristics such as year in school and gender to predict participation in the online course evaluation.

A third approach to examining nonresponse bias is to analyze and identify differences between early and late responders (Porter & Whitcomb, 2005; Groves, 2006). Many online survey programs give a date of receipt for the survey, which makes this approach feasible. A strength of this approach is that there is an easy way to track respondents by time. The weakness of this approach is deciding the date by which to classify respondents as late (Porter & Whitcomb, 2005). The issue is making sure that the dates selected are relevant to the topic and population being surveyed. If one does not want to select absolute dates of early and late, there could be an analysis of the relationship between ratings and the number of days to completion. There has been little research to examine nonresponse and nonresponse bias in online course evaluations by early and late respondents.

The fourth and final approach to be described involves following up with nonresponders to discover why they did not respond. Kuh (2003) used this approach to contact students who did not complete the web or paper version of the National Survey of Student Engagement (NSSE). He found that students who did not complete NSSE were slightly more engaged than students who completed the web or paper version. With online course evaluations, one would contact students who did not respond after the end of the semester to ask why they did not respond. Before pursuing this approach with online course evaluations, one must consider how student perception and concern about anonymity and confidentiality of results would be impacted.

Out of all of the approaches described, some are more appropriate for online course evaluations than others. For example, asking students who did not respond to the online course evaluation their reasons for not responding might be difficult. Because this is a group that has already been uncooperative, there is a good chance that very few would actually respond to the follow-up request. Asking students why they did not respond may also trigger concerns about confidentiality. Obtaining additional information about survey students using the “rich sampling frame” approach is very straightforward in higher education. Universities collect administrative data (e.g. GPA, demographics, residence status) on students, making this strategy feasible. With this approach, the same information would be available for both respondents and nonrespondents, making it the best opportunity for comparing both groups. Although all of these approaches provide some information about respondents and nonrespondents, it is also important to consider what underlying processes might be occurring that make students complete an online course evaluation. Theory may provide guidance in understanding these underlying processes.

Theoretical Perspectives on Nonresponse

This section will discuss two theoretical perspectives explaining why individuals do or do not complete surveys. The first theory is organizational citizenship behavior. This focuses on an individual’s relationship with the organization as a reason for responding to the course evaluation. A contrasting theory, leverage-salience theory, uses a different approach and views the propensity to respond to a survey based on characteristics within the survey, and the circumstances surrounding the survey’s administration.

Responding to surveys as organizational citizenship behavior. In general, students complete online course evaluations during their own time, not during class time. Most professors will ask students to complete the evaluation, and explain how the results will be used so that students will understand the importance of their contribution. Because students are completing the evaluation at the end of the course, there will be no direct benefit for taking their own time to fill out this survey, unless an incentive is offered. Within the business community, using one's time without any personal gain for the betterment of the organization is called organizational citizenship behavior (OCB). As described by Organ (1988), "OCB represents individual behavior that is discretionary, not directly or explicitly recognized in the formal reward system, and that in the aggregate promotes the effective functioning of the organization" (p. 4). As long as there are no rewards for completing the online course evaluation, the behavior could be considered organizational citizenship behavior. The majority of the research in this area until recently has focused on behavior within the workplace.

More recently, researchers have examined the applicability of OCB theory to survey completion in other settings. Spitzmuller et al. (2006) studied the relationship between predictors of organizational citizenship behavior (OCB) and survey response within the collegiate setting. Specifically the researchers examined college students' opinions of procedural justice, perceived organizational support, social exchange, and reciprocation wariness as they relate to their organization (the university) and intentions to participate in future university surveys. Procedural justice was described as behavior used by a member to "get back" at the organization for previous unfair or unjust decisions. Perceived organizational support relates to how the individual rates the

organization as being caring and supportive (Rhoades & Eisenberger, 2002). Social exchange relates to how well the individual believes that there is a mutual exchange of caring between the individual and organization (Spitzmueller et al. 2006). Finally, reciprocation wariness was described as more of a personality characteristic in which individuals are afraid of being exploited if they give more than expected in an organization.

The sample consisted of undergraduates in upper level business courses. Graduate assistants administered a survey that contained questions about satisfaction with university services, predictors of organizational citizenship behavior, and intention to complete future surveys. Several weeks after the in-class survey, a web survey about university shopping and dining options was sent to students who participated in the in-class survey.

After receiving the second set of survey responses, the researchers divided students into three categories based on their intention to respond to future surveys: active nonrespondents, passive nonrespondents, and respondents. Active nonrespondents were students who indicated they had no intention of responding to future surveys and did not respond. Passive nonrespondents were students who indicated that they planned to respond to future surveys, but did not respond to the follow-up survey. Respondents were students who indicated they would, and did respond to the follow-up survey.

A one-way MANOVA was conducted to examine differences in respondents and nonrespondents. The independent variable had three levels (active nonrespondents, passive nonrespondents, and respondents). The dependent variables were the predictors of OCB behavior (procedural justice, perceived organizational support, social exchange,

and reciprocation wariness). Spitzmuller et al. (2006) found that nonrespondents and respondents differed from passive nonrespondents and respondents in all four areas. Active nonrespondents were more likely to rate the university as lower in procedural justice, perceived organizational support, and social exchange. In addition, active nonrespondents were also more likely to be higher on reciprocation wariness than responders. They also examined demographic characteristics of students and discovered that there were significant differences in response by gender, with more women than men electing to respond in all three groups.

To further make the connection between organizational citizenship behavior and survey response, Spitzmuller et al. (2007) conducted follow up research that examined OCB qualities, rather than predictors of OCB. This study asked subjects if they participated in OCB behaviors. They examined survey response using Organ's (1988) dimensions of OCB: a) altruism, b) conscientiousness, c) civic virtue, and d) courtesy. The definitions of each dimension as described by Organ (1988) are straightforward. Altruism refers to behavior that is performed for the good of someone else, without expecting anything in return. Conscientiousness is behavior that involves "doing a little more" such as staying late at work double-checking a work document. Unlike altruism, conscientiousness is not necessarily directed towards helping another person. Civic virtue relates to being involved in activities that generally help the organization such as a governance group. Finally, courtesy refers to being polite.

This research examined the relationship between OCB and survey response with two studies. In the first study, firefighters were asked to complete a survey which contained items addressing the OCB dimensions using a scale developed by Podsakoff,

MacKenzie, Moorman, and Fetter (1990). Researchers asked firefighters their intention to complete future surveys. Researchers found courtesy to be the only OCB dimension related to survey response behavior. Unfortunately one weakness of this study is that the firefighters were only asked their intention to respond, not to actually respond to another survey.

In the second study, the sample was comprised of college students. Representing the university's office of Institutional Research as the official sponsor of this research, graduate assistants administered a survey packet in undergraduate upper level business courses. This packet contained survey items that addressed the following: OCB behavior, university policies, and intention to complete future surveys. A follow-up web survey was sent by the office of Institutional Research that asked satisfaction questions about university services.

Data analysis first consisted of dividing students into the same three response categories from their 2006 study. A one-way MANOVA was conducted to examine differences in respondents and nonrespondents. The independent variable had three levels (active nonrespondents, passive nonrespondents, and respondents). The dependent variables were the OCB dimensions (altruism, conscientious, civic virtue, and courtesy). They found that active nonrespondents were significantly less likely than respondents and passive nonrespondents to participate in helping behaviors (altruism), optional activities at the university (civic virtue), and "courteous behavior," (p. 465). There were no significant differences among the three groups on the conscientiousness variable. More females than males were classified as respondents. There were no differences based on ethnicity.

Both Spitzmueller et al. (2006) and Spritzmueller et al. (2007) found significant differences in survey response rates for individuals based on their stated opinions about an organization, and their OCB behaviors. Although usually applied in the business sector, the studies conducted by Spitzmueller and associates (2006, 2007) provide support for the generalizability of this theory to other settings. This theory has many dimensions that may prove useful to the study of students completing course evaluations. For one, like employees and employers, there is a power structure setup with students and instructors, and administrators (Spitzmueller, 2006). When thinking about the predictors of OCB, students have relationships with their university in which they may experience a need for procedural justice or organizational support. Finally, when thinking about students displaying OCB, the university experience provides many opportunities for students to be conscientious, and display civic virtue, courteousness, and conscientiousness inside and outside the classroom. These findings present several interesting thoughts about the connection between OCB behaviors and responding to an online course evaluation. With the Spitzmueller et al. (2007) study, there was no significant difference in response rates between conscientiousness and the response groups. When thinking about conscientiousness in a classroom setting, one often thinks about students who perform well academically. Studies have found grade point average to be a significant predictor of participating in college surveys (Porter & Whitcomb, 2005; Porter & Umbach, 2006). If academic performance is considered a form of conscientiousness, then one may expect student online course evaluation nonresponse to vary by student grade point average. Higher education research has also captured students' feelings about their institutions through measures of sense of belonging

(Hausmann, Schofield & Woods, 2007) and interactions with faculty (Kim & Sax, 2009) with findings varying by gender and ethnicity. Using the OCB studies as a foundation, the higher education studies suggest that students completing online course evaluations may also vary by demographic characteristics such as ethnicity and gender. Although OCB studies focus on the individual's relationship with the organization as a factor in completing a survey, other perspectives such as leverage salience theory, suggest that changing the characteristics of a survey may influence the decision to complete a survey.

Leverage salience theory. In order to minimize survey nonresponse, it is important to know what leads the transition from survey recipient to survey respondent. Leverage salience theory attempts to define and measure positive and negative attributes of survey designs and processes to enhance participation. With leverage salience theory (Groves, 2000), there are a variety of positive and negative survey characteristics guiding the decision to respond. When there are more positive influences, there is more leverage towards responding. Usually the potential survey respondent's interest in the topic is given as a starting point, and various survey characteristics are considered potential factors in leveraging the decision to participate in the survey. Studies testing the viability of leverage salience have manipulated characteristics of the surveys (i.e., interest, length) or the procedures (i.e., incentives, follow ups) to determine impact on response rates. A critical component of this theory is that the leverages are interacting with individual factors or considerations (e.g. interest in survey topic, socioeconomic status) that may or may not be known.

Groves et al. (2006) provides an example of testing this theory. They tested this theory by manipulating incentives (no incentive or incentive given), survey interest (low

and high) and community involvement (low or high). It was predicted that the cash incentive would provide the extra leverage needed for individuals who were not high in community involvement. For individuals with high community involvement, it was predicted that the incentive would not significantly influence the decision to respond to the survey. The researchers predicted correctly with low community involvement survey recipients with incentives responding at much higher rates than low community involvement recipients without incentives. For high community involvement recipients, the incentives were not as influential.

Using this theory with online course evaluations, one must try to identify the attributes that might “tip the scales” and provide “leverage” for deciding to respond. Dommeyer et al (2004) offered a miniscule increase in the final grade, an in-class presentation about how to log-in to the online evaluation, and early receipt of the final grade. The intent was to see which if any of these conditions would influence course evaluation response. The courses which offered the grade increase received the highest response rates. In this example, the incentive of a grade provided the “leverage” needed to persuade more students to complete the online evaluation. Although this is a great example of leverage-salience theory, some instructors may not want to give grade incentives to increase their response rates. An important component in using leverage-salience theory is finding the most effective tool to use as leverage. As indicated in the meta-analysis by Shih and Fan (2008), incentives did not increase online response rates to be equal to or higher than surveys administered by mail. With the Dommeyer et al. (2004) study, a small grade increase was an effective leverage.

Fidelman (2007) explored if liking a course was a predictor of responding to the online course evaluation. She found that liking a course was not related to the decision to respond to course evaluation but was related to a student's evaluation of that course. This means that students who liked the course gave higher evaluations than students who did not like the course. Although one might consider this result an indication of nonresponse bias, it was not. If liking the course was a predictor of course evaluations and response, nonresponse bias would have been a concern (Groves et al., 2006).

Knowing that liking a course is not a predictor of completing an online course evaluation is important because it provides information about where and how leverage could be used. In Groves et al. (2006), there was the expectation that individuals not interested in the topic would respond to an incentive. With online course evaluations, finding the appropriate leverage is important. Before using leverage, it is important to know where and with whom the leverage is needed. The current study will capture that information by exploring student characteristics as predictors of online course evaluation nonresponse.

Online Course Evaluations

The research on online course evaluations is heavily focused on the issue of response rates, or more specifically, the lack of response. This section will discuss results from studies about nonresponse in online course evaluations, and considerations for future research.

Examination of nonresponse bias. Determining nonresponse bias in online course evaluations requires a relationship between predictors of response and predictors of ratings. Research has found that response rates for online course evaluations are lower

than paper evaluations (Avery et al., 2006; Dommeyer et al., 2004, Layne et al., 1999). Many studies have examined gender, expected course grade or course grade, and ethnicity as variables to predict who will respond to the online course evaluation. For gender, research has found that females responded at higher rates than males (Avery et al., 2006; Fidelman, 2007). In some cases, students with higher course grades respond at higher rates than students with lower course grades (Layne et al., 1999; Avery et al., 2006; Fidelman, 2007), but not always (Liegle & McDonald, 2005; Thorpe, 2002). Finally, Caucasian students generally respond at higher rates than students of color (Avery et al., 2006; Fidelman, 2007).

To determine nonresponse bias, they calculated differences in mean scores for paper and pencil and online administrations for different demographic groups (Thorpe, 2002; Liegle & MacDonald, 2005) or created models to predict course evaluation response and ratings using the same variables (Fidelman, 2007; Avery, 2006).

The same studies also examined differences in evaluation ratings by student characteristics. Because the variables do not differ significantly by student demographic characteristics and evaluation scores, it was determined that nonresponse bias did not exist (Avery et al., 2006; Dommeyer et al., 2004, Fidelman, 2007; Layne et al., 1999; Liegle & McDonald, 2005; Thorpe, 2002).

Considerations for future research. Based on the review of previous online course evaluation studies, there are opportunities to add to the research about nonresponse and nonresponse bias. There are two areas where additional studies could further research pertaining to online course evaluations.

First, current studies about online course evaluations have not considered that students take multiple courses per semester, and may complete all of their evaluations at the same time (Avery et al., 2006; Dommeyer et al., 2004, Fidelman, 2007; Layne et al., 1999; Liegle & McDonald, 2005; Thorpe, 2002). Conducting an analysis that considers the cross-classified structure of students with courses will provide better estimates for predicting course evaluation ratings, probability of response, and ultimately identifying nonresponse bias.

Second, many of the previous studies occurred at large universities (Avery et al.; Dommeyer et al., 2004; Fidelman, 2007, Thorpe, 2002) with homogeneous populations (Avery et al., 2006; Fidelman, 2007). The current study will occur at a small, private Catholic university with a diverse population. This change in population will provide the opportunity to explore whether nonresponse and nonresponse bias patterns are similar across all college student populations and across the student populations at this university.

Summary of Research

Online course evaluations come with many advantages, but also with concerns about student nonresponse resulting in nonresponse bias. Explanations for online course evaluation nonresponse could be found in survey research and studies about online course evaluations. A summary of the literature review is as follows:

1. The majority of online surveys obtain smaller response rates when compared to other modes of administration (Manfreda et al., 2008; Shih & Fan, 2008). Offering incentives did not close the response rate gap between online surveys and paper based surveys (Shih & Fan, 2008). A closer view of nonresponse

studies conducted with college populations reveals statistically significant differences in participation by student gender, GPA, and ethnicity (Porter & Whitcomb, 2005; Sax et al. 2003; Sax et al. 2008).

2. The decision to complete an online survey could be influenced by a variety of factors. In some cases, the relationship between the individual and the organization makes a difference (Spitzmueller et al., 2006; Spritzmueller et al., 2007). Individuals' interactions with survey characteristics may also trigger nonresponse or response behavior (Groves et al. 2006). Modifying survey characteristics or incentives requires understanding the population being surveyed to identify and use the most effective incentives and survey techniques. This point was emphasized by Fidelman (2007) discovering that liking a course was unrelated to responding to the online course evaluation.
3. Research about online course evaluations has found that faculty and students acknowledge the pros and cons of online course evaluations, with the largest con being low response rates (Donovan, Mader & Shinsky, 2007; Dommeyer, Baum & Hanna, 2002; Dommeyer, Baum, Chapman, & Hanna, 2002). Investigations of nonresponse bias revealed differences in response rates across groupings of students (e.g. GPA, residency, gender, ethnicity) (Avery et al, 2006; Fidelman, 2007; Thorpe), but no statistically significant differences in mean evaluation ratings across the groups (Avery et al., 2006; Heath et al., 2007; Leigle & McDonald, 2005; Thorpe, 2002). The majority of these studies occurred at large universities, using a small subset of courses. None of these studies considered that

students take multiple courses, and if their response patterns were consistent across all courses.

The Current Study

The focus of this study is to examine nonresponse bias in online course evaluations. Previous studies have only included specific courses at a university or courses within a specific department. This study will include all undergraduate courses that received an evaluation resulting in an analysis with student responses across multiple courses. This study will further the research of online course evaluations by including the following:

1. A cross-classified model will be analyzed to account for students taking multiple courses. Only the Fidelman (2007) study considered the structure of students within courses. This study analyzed the data as students nested within courses. The current study will account for students taking multiple courses at the same time, and provide a view of students' behavior across all of their courses.
2. The population in the study will be from a small diverse campus. The populations from the majority of studies have occurred at large universities with homogeneous populations. This will contribute to the literature by providing future researchers the opportunity to see if online course evaluation nonresponse and nonresponse bias is pervasive across all institutions, or only certain types of institutions.

The major research question in this study is: "Is there nonresponse bias in online course evaluations?" As defined by Groves (2006), "nonresponse in a survey estimate arises when the set of mechanisms that influence the participation decision is related to the variables involved in the estimate" (p. 722). For nonresponse bias to exist in online

course evaluations, the variables included in the study must significantly predict course evaluations ratings and participation in the online course evaluations. The following questions will be addressed to answer the major research questions:

1. Are student characteristics such as gender, ethnicity, cumulative GPA, course grade, institutional credits, and total credits related to online course evaluation ratings?
2. Is course size related to course evaluation ratings?
3. Are student characteristics such as gender, ethnicity, cumulative GPA, course grade, institutional credits, and total credits related to responding to the online course evaluation?
4. Is course size related to responding to an online course evaluation?

Nonresponse bias will be indicated by variables that predict both online course evaluation ratings and response.

Chapter 3

Methods

Participants

This study was conducted at a small Catholic University located on the east coast. The university's population was ethnically diverse with the student body being 51% Caucasian, 7% Asian/Pacific Islander, 15% African American, 9% Hispanic. The majority of students are commuters (70%) and 30% are transfer students. Study participants were undergraduate students who enrolled in a face-to-face course during spring 2009. This research study focused on the question, "Is there nonresponse bias in online course evaluations?"

Instruments

This study used the university's course evaluation form for face-to-face courses (Appendix A). This course evaluation form was created by a university committee consisting of faculty and the Coordinator of University Assessment in 2002. The instrument consisted of eleven Likert type items, four demographic questions, and four open-ended questions.

For the purpose of this study, a total score was calculated as the sum of items 4 – 9. These items were consistent in their format and focus. Previous examinations of course evaluation data with items 4-9 yielded internal consistencies of .90, and only one factor when conducting an exploratory factor analysis.

Courses

This study included results from online course evaluations collected during the spring 2009 semester. Only undergraduate courses that met face-to-face were included in

this study. Courses that were independent research, student teaching, one credit, and had enrollments less than 10 were excluded from this study.

Procedure

Distribution of online course evaluations. Data were collected as part of the university's regular course evaluation process. The process for distribution of course evaluations during spring 2009 was as follows:

1. Instructors received notification the week before the online course evaluation started from the Vice President for Academic Affairs (VPAA) requesting that they remind their students to complete the online evaluation.
2. Students received notification the week before the online course evaluation started from the Vice President for Academic Affairs (VPAA) requesting them to complete the online evaluation.
3. Students received an email from the office of Institutional Effectiveness for each course in which they were enrolled. All emails were sent to students' university email addresses on the first day of the course evaluation period. For Spring 2009, course evaluations were available from April 14 to April 28.
4. To access the online evaluation, students entered their course section and student ID number. Both pieces of information were included in the email message.
5. Students received three reminder notices. The reminders were generated from the survey program, which kept track of who responded. The planned reminder notices for spring 2009 were sent on April 20, April 23, and April 27 (see Appendix B).

6. Students were also reminded about course evaluations through flyers, table tents in the cafeteria, and reminder cards distributed on campus (Appendix C).

Reminder cards were distributed at the main cafeteria and at the library checkout desk.

Retrieval of course evaluation data. Course evaluation data were collected online via survey software. Once the course evaluation period ended, an excel spreadsheet was downloaded by the office of Institutional Effectiveness for analysis.

Analysis

Descriptives. Descriptives of respondents' and nonrespondents' demographics and course size were examined, focusing on the number and percentage of students responding in each category. Coefficient alpha was used to review internal consistency of survey scores. A total score was calculated using items 4-9. Items were on a Likert scale, on which they were scored with the most positive response was scored as a 5 and the least positive scored a 1.

Exploring Nonresponse Bias. Determining nonresponse bias requires two components: respondent and course characteristics as predictors of actual survey response; and a relationship between the item of interest and characteristics of respondents. The object of interest in this study was course evaluation ratings obtained from adding items 4-9 on the course evaluation. Gender, ethnicity, cumulative GPA, course grade, institutional credits, and total credits were included as predictors (see Table 1 for details).

Table 1

Variables included in study

Variable	Description
Course evaluation rating	Consisted of items 4 – 9 on the course evaluation Possible range from 6 to 30
Course evaluation participation	0 = did not complete online course evaluation 1 = completed online course evaluation
Course size	Actual course size. Course enrollments ranged from 10 to 48 students.
Ethnicity	The variables were dummy coded (0 or 1) and entered as separate groups with Caucasian being the comparison group. The following groups were included: Asian, Black, Hispanic, Native American, 2 or more races, non-resident alien
Final course grade	The numerical value of the course grade was used and not the letter grade. Points for course grade ranged from 0 to 4 There were linear and quadratic terms for this variable.
Gender	A dummy variable where 0 = males 1 = female
Grade point average (GPA)	The numerical value of the grade point average will be used and not a letter grade. Points for grade point average range from 0 to 4. There were linear and quadratic terms for this variable.
Institutional credits	This variable included the number of credits taken only at the university where the study was conducted
Total credits	The variable included the total number of credits taken at the university where the study was conducted and other institutions. There were linear and quadratic terms for this variable.

Predicting course evaluation ratings. Because most students take more than one course during a semester, an analytic technique is needed to model possible dependencies due to students taking multiple courses, and multiple students taking the same course. As seen in Table 2, students are represented in the columns and courses in the rows. The number of students enrolled in each course, and the number of courses taken will vary with the student. The responses of students within the same course may be more similar than the responses of student within different courses. Correspondingly, the responses of the same student regarding different courses may be related. Because of the potential lack of independence in students rating multiple courses and responses from multiple students in the same course, an analytic technique needed to be selected that would account for this dual dependency. A cross-classified analysis model allows for analyzing the variables that are within each cell (characteristics unique to each student in each class), variables that are the characteristics of the columns (courses) and variables that are the characteristics of the rows (students).

Table 2

Cross-classified model

	Course 1	Course 2	Course 3	Course 4
Student 1	X		X	X
Student 2		X	X	
Student 3				
Student 4	X			X
Student 5	X	X	X	X

The cross-classified model can be more formally described as follows:

1. Level 1 – within the cells formed by students crossed with courses

With this study, the course grade is unique to each student and each course. The model is as follows: Course evaluation rating = $\pi_{0jk} + \pi_{1jk}(\text{Course Grade})_{jk} + \pi_{2jk}(\text{Course Grade})^2_{jk} + e_j$

2. Level 2 – between students and between courses

Student and course characteristics were included at this level. The model can be represented as follows:

$$\pi_{0jk} = \theta_0 + \gamma_{01}(\text{Coursesize})_j + \beta_{01}(\text{Gender})_k + \beta_{02}(\text{Ethnicity})_k + \beta_{03}(\text{Cumulative GPA})_k + \beta_{04}(\text{Cumulative GPA})^2_k + \beta_{05}(\text{Institutional Credits})_k + \beta_{06}(\text{Total Credits})_k + \beta_{07}(\text{Total Credits})^2_k + b_{00j} + c_{00k}$$

$$\pi_{1jk} = \theta_1$$

$$\pi_{2jk} = \theta_2$$

θ_0 is the expected rating when all predictors are set to zero, γ_{01} is the linear effect of the size of course j , β_{01} is the effect of student k 's gender, $\beta_{02} - \beta_{07}$ are the slopes for the effects of other student characteristics as indicated, b_{00j} is the random effect of course j , and c_{00k} is the random effect of student k . θ_1 and θ_2 are the mean linear and quadratic effects, respectively, of course grade. There are no random terms predicting π_{1jk} and π_{2jk} , indicating they are held constant across combinations of students and courses (Raudenbush & Bryk, 2002).

To allow for easier interpretation of the parameters, the continuous variables were grand-mean centered, allowing the intercept to be interpreted as the mean of the dependent variable. To conduct the analysis, the following steps were taken:

1. Ran the unconditional model – This was the model without any predictors.
2. Reviewed the variance components to determine the following: (a) percentage of variance between students, (b) percentage of variance between classes (c) percentage of residual variation not explained by either the main effect of student or the main effect of class.
3. Entered the predictors - All predictors were reviewed for statistical significance. Non-significant predictors were removed and the model was re-run using only statistically significant predictors. This continued until the model only contained statistically significant predictors.

Predicting participation in online course evaluations. The second component of determining nonresponse bias involved determining if a relationship existed between relevant predictors of ratings and participation in the online course evaluation. All of the predictor variables were the same as in the previous model; only the dependent variable changed. With this analysis, the dependent variable was of the form 0 (did not participate) or 1 (participated). As described by Raudenbush and Bryk (2002), a non linear model with a logit link and a Bernoulli distribution for the residuals was appropriate for this form of response (0 or 1). This model could more simply be described as a logistic regression model extended to account for the dependencies due to student and course. Because this was still a cross-classified model, the model retained the within and between levels as explained earlier; however, the dependent variable was reviewed differently. The analysis produced log-odds which can be converted to odds and ultimately into the probability of responding.

Determining nonresponse bias. Nonresponse bias was determined by using the “rich data frame” approach (Grooves, 2006). This approach used variables that were available from both nonrespondents and respondents to determine if there is nonresponse bias. Variables that were both significant predictors of course evaluation ratings and predictors of completing the course evaluation were potential sources of nonresponse bias. Statistically significant variables from the analysis predicting course evaluation scores, and probability of responding to the course evaluation were compared for common predictors.

Chapter 4

Results

This study included online course evaluation records from spring 2009. After removing 62 courses with enrollment sizes less than 10, a total of 348 course sections and 2011 students were included in the analysis. The average course enrollment was 22. Six (6) course evaluation questions were added to form a total score (see Appendix A). The average course evaluation rating was 25.90, with ratings ranging from 6 to 30. Internal consistency reliability (coefficient alpha) of the items was calculated and found to be .92.

Table 3

Response rate of study participants

Characteristics	N (Across all course sections)	Response Rate
Gender		
Female	1463	0.51
Male	538	0.36
Ethnicity		
Asian	171	0.46
Black	293	0.35
Hispanic	236	0.43
White	897	0.51
Native American	8	0.48
Non-resident alien	110	0.41
Two or more races	131	0.44
Other	165	0.49

The major research question of this study was “Is there nonresponse bias in online course evaluations?” The research questions address predictors of course evaluation

ratings and predictors of course evaluation participation. The specific research questions were

1. Are student characteristics such as gender, ethnicity, cumulative GPA, course grade, institutional credits, and total credits related to online course evaluation ratings?
2. Is course size related to course evaluation ratings?
3. Are student characteristics such as gender, ethnicity, cumulative GPA, course grade, institutional credits, and total credits related to responding to an online course evaluation?
4. Is course size related to responding to an online course evaluation?

Results from predicting online course evaluation ratings will be explored first, followed by a review of results from predicting course evaluation participation.

Predicting online course evaluation ratings

The research questions for predicting online course evaluation ratings were

1. Are student characteristics such as gender, ethnicity, cumulative GPA, course grade, institutional credits, and total credits related to online course evaluation ratings?
2. Is course size related to course evaluation ratings?

A cross-classified random effects model was used to analyze predictors of online course evaluation ratings (see Table 2). First the variance components were examined to determine if sufficient variance existed to justify this model. This was accomplished by running the unconditional model – a model with no predictors. The analysis produced a t-ratio, which was used to determine statistical significance. Standard errors were also

calculated to estimate the accuracy of the parameters. See Table 4 for all estimates. As suggested in Raudenbush and Bryk, (2002), intraunit correlations were calculated to examine the percentages of variance between students and between courses. Nineteen percent (19%) of the variance was between students, and 25% between courses.

Determining the percentage of variance in course ratings related to both students and courses is a unique feature of the cross-classified model. Next, all of the variables were analyzed in the model (Model 1). The only significant predictor was course grade. The final model only included the significant predictor – course grade. The results of this model are interpreted as the average course evaluation rating increasing by 1.00 on the 30-point rating scale for every one unit change in course grade. This result can be summarized as course evaluation ratings increasing with course grades.

The answers to the research questions are as follows:

1. Are student characteristics such as gender, ethnicity, cumulative GPA, course grade, institutional credits, and total credits related to online course evaluation ratings?

Only course grade was related to course evaluation ratings in this study.

2. Is course size related to course evaluation ratings?

Course size was not a significant predictor of course evaluation ratings in this study.

Table 4

Cross-Classified Random Effects Model Predicting Online Course Evaluation Ratings

Fixed Effects	Unconditional Model			Model 1			Final Model		
	Coeff.	se	t-ratio	Coeff.	se	t-ratio	Coeff.	se	t-ratio
Intercept	25.853	0.156	165.675	25.982	0.484	53.589	25.853	0.155	167.208
Student Variables									
Gender (Female)				-0.299	0.219	-1.359			
Asian				0.399	0.335	1.192			
Black				-0.012	0.301	-0.041			
Hispanic				0.047	0.300	-0.156			
Native American				0.399	1.533	0.260			
2 or more races				0.160	0.393	-0.408			
Non-Resident Alien				0.616	0.421	1.463			
Other				-0.445	0.328	-1.358			
Cumulative GPA				-0.042	0.865	-0.049			
Cumulative GPA ²				-0.020	0.153	-0.095			
Institutional credits				-0.001	0.003	-0.304			
Total Credits				-0.014	0.009	-1.640			
Total Credits ²				0.000	-	2.087			

Table 4

Cross-Classified Random Effects Model Predicting Online Course Evaluation Ratings (continued)

Fixed Effects	Unconditional Model	Model 1			Final Model		
Course Variable							
Course Size		0.004	0.018	0.236			
Student x Course Variables							
Course Grade		1.007*	0.350	2.877	1.001*	0.093	10.791
Course Grade ²		-0.004	0.067	-0.058			

*p < .01

Predicting online course evaluation participation

The second set of research questions involving prediction of participation in online course evaluations were

3. Are student characteristics such as gender, ethnicity, cumulative GPA, course GPA, institutional credits, and total credits related to responding to an online course evaluation?
4. Is course size related to responding to an online course evaluation?

With this analysis, the dependent variable was of the form 0 (did not participate) or 1 (participated). As described by Raudenbush and Bryk (2002), a nonlinear model with a logit link and a Bernoulli distribution for the residuals is appropriate for this form of response (0 or 1). This model could more simply be described as a logistic regression model extended to account for the dependencies due to student and course. The coefficients from the logistic regression model were transformed to odds ratios, which above 1 indicate an increase in the odd of the event occurring. Because this is still a cross-classified model, the model retained the within and between levels as explained earlier; however, the dependent variable was modeled differently because it was dichotomous. The steps to running this analysis were the same as the steps for predicting course evaluation results. An unconditional model was attempted, but did not converge after 100 iterations. Variance due to course was zero, which indicated lack of variance across courses. Due to lack of variance, a 2-level model with responses nested within students was used instead. The model was as follows:

$$\text{Level 1: } \eta = \pi_{0j} + \pi_{1j} (\text{Final Grade})_{jk} + \pi_{2j} (\text{Final Grade})_{jk}^2$$

$$\text{Level 2: } \pi_{0j} = \beta_{00} + \beta_{01}(\text{Gender})_k + \beta_{02}(\text{Ethnicity})_k + \beta_{03}(\text{Cumulative GPA})_k + \\ \beta_{04}(\text{Cumulative GPA})_k^2 + \beta_{05}(\text{Institutional credits})_k + \beta_{06}(\text{Total Credits})_k + \\ \beta_{07}(\text{Total Credits})_k^2 + r_{0j}$$

$$\pi_{1j} = \beta_{10}$$

$$\pi_{2j} = \beta_{20}$$

where $\eta = \ln \frac{P(y_{jk} = 1)}{1 - P(y_{jk} = 1)}$, r_{0j} is a random term for the intercept for student j ,

and other terms are as defined before.

With the 2 level-model, course-specific responses within students were at level 1 and students were at level 2. Table 5 displays the results with all predictors (Model 1) and only significant predictors (Model 2). The results indicate that the female students' odds of participating were 2.3 times the odds for the male students. For each unit increase in course grade, the odds of participating were multiplied by a factor of 1.2. The odds of participation for Black students were approximately 0.5 times the odds for White students.

3. Are student characteristics such as gender, ethnicity, cumulative GPA, course grade, institutional credits, and total credits related to responding to an online course evaluation?

Yes. In this study, gender, ethnicity, and course grade were related to responding to an online course evaluation.

4. Is course size related to responding to an online course evaluation?

No, course size was not related to responding to an online course evaluation in this study.

Determining nonresponse bias

According to Grooves (2006), a variable becomes a potential source of nonresponse bias when it is related to both who participates and the actual variable of interest. Using this definition, course grades was the only potential source of nonresponse bias in this study. Although many variables predicted who responded, only course grades predicted course evaluation ratings and participation.

Table 5

Hierarchical Generalized Linear Model Estimating Predictors of Online Course Evaluation Participation

	Model 1				Model 2			
	Coeff.	se	t-ratio	odds ratio	Coeff.	se	t-ratio	Odds ratio
Intercept	-0.671	0.148	-4.549	0.511	-0.762	0.128	-5.944	0.467
Within-Student Predictors								
Course Grade	0.322**	0.107	3.012	1.451	0.216**	.046	4.736	1.241
Course Grade ²	-0.036	0.022	-1.675	0.941				
Between-Student Predictors								
Gender (Female)	0.800**	0.152	5.265	2.223	0.842**	0.147	5.726	2.322
Asian	-0.221	0.243	-0.911	0.801				
Black	-0.697**	0.203	-3.433	0.499	-0.734*	0.073	-8.347	0.479
Hispanic	-0.217	0.214	-1.018	0.804				
Native American	0.497	1.101	0.451	1.643				
2 or more races	-0.338	0.273	-1.241	0.713				
Non-Resident Alien	-0.421	0.299	-1.409	0.656				
Other	0.186	0.248	0.750	1.204				
Cumulative GPA	0.351	0.599	0.585	1.420				
Cumulative GPA ²	0.079	0.110	0.723	1.082				

*p < .05, **p < .01

Table 5

Hierarchical Generalized Linear Model Estimating Predictors of online course evaluation participation (continued)

	Model 1				Model 2			
	Coeff.	se	t-ratio	odds ratio	Coeff.	se	t-ratio	odds ratio
Institutional credits	-0.000	0.002	-0.372	0.999				
Total Credits	-0.012	0.006	-1.823	0.988				
Total Credits ²	-0.000	0.000	0.869	1.000				

*p < .05, **p < .01

Chapter 5

Discussion

The purpose of this study was to examine nonresponse bias in online course evaluations. A two-tiered approach was used to explore this issue by examining predictors of online course evaluation ratings and predictors of online course evaluation participation. Online course evaluations have traditionally produced lower than desired response rates (Avery et al., 2006; Fidelman, 2007; Layne et al. 1999). Increases in online course evaluation response rates have been observed when some form of intervention was employed (Dommeyer et al., 2004, Johnson, 2003). As a result of lower response rates and the potential impact of inferences made on the basis of online course evaluations, concerns about nonresponse bias developed. This study was designed to address the issue of nonresponse bias at one institution of higher education.

Although the purpose of this study was to address a concern at one institution, the contribution of this study to online course evaluation research is multi-faceted. First, many of the previous studies about online course evaluations compared response rates and student characteristics of paper and pencil course evaluations to online course evaluations (Avery et al., 2006; Fidelman, 2007; Layne et al. 1999; Leigle & McDonald, 2005; Thorpe, 2002). This study was not a comparison study, and occurred at a university that is only in its second year of online course evaluations. This study also differed from previous studies because it occurred at a small, religiously affiliated university with a total undergraduate enrollment of less than 3,000 students. The majority of previous studies occurred at larger universities and usually incorporated only a few departments and/or classes. The student population at this institution was also diverse with the

following ethnic composition of students: 51% White, 7% Asian, 15% Black, 9% Hispanic. Another difference is that this study used a multi-level approach to analyze potential sources of nonresponse bias. Although previous studies such as Fidelman (2007), used hierarchical linear regression to examine course evaluation ratings, this study used a cross-classified random approach to examine nonresponse and course evaluation ratings across a student's entire course schedule.

The major research focus of this study was, "Is there nonresponse bias in online course evaluations?" This important issue was addressed via the following research questions:

1. Are student characteristics such as gender, ethnicity, cumulative GPA, course grade, institutional credits, and total credits related to online course evaluation ratings?
2. Is course size related to course evaluation ratings?
3. Are student characteristics such as gender, ethnicity, cumulative GPA, course grade, institutional credits, and total credits related to responding to an online course evaluation?
4. Is course size related to responding to an online course evaluation?

The following sections discuss the results of the study by reviewing the research questions within the context of previous research about online course evaluations and theoretical considerations for responding to online course evaluations.

Research Findings

Are student characteristics such as gender, ethnicity, cumulative GPA, course grade, institutional credits, and total credits related to online course evaluation ratings? This study found online course evaluation participation to increase with higher course grade and for female students. The results of this study corroborate those of Liegle and McDonald (2005) who found that weaker students were less likely to complete an online course evaluation than a paper evaluation and generally gave faculty lower course evaluation ratings than students who were academically stronger (p. 8).

Is course size related to course evaluation ratings? This study found no relationship between course evaluation ratings and course size. This suggests that students do not vary their ratings based on course size.

Are student characteristics such as gender, ethnicity, cumulative GPA, course grades, institutional credits, and total credits related to responding to an online course evaluation? This study found gender, ethnicity, and course grades to be predictors of participation, with the odds of females participating being higher than males. Avery et al. (2006) and Thorpe (2002) also had similar findings. Gender was not a predictor of participation for Layne et al. (1999). This finding is also similar to other studies examining nonresponse bias in student surveys in higher education (Porter & Whitcomb, 2005; Sax et al. 2003). Patterns of participation varying by ethnicity have also been observed in other studies with white students participating at higher rates. In this study Black students were not as likely to participate as White students. In Avery et al. (2006) Asian students were less likely to participate than other ethnic groups. Thorpe (2002) developed minority and non-minority classifications and found that non-minority

students responded at higher rates. Also similar to other studies, this study found that online course evaluation participation increased with course grade (Liegle & McDonald, 2005; Thorpe, 2002). Unlike previous studies, a relationship between online course evaluation participation and overall GPA (Thorpe, 2002; Layne et al, 1999) and year in school (Layne et al., 1999) was not observed.

Is course size related to responding to an online course evaluation? This study did not find a relationship between course size and participation in online course evaluations. This finding differed from Avery et al (2006) in which participation in online course evaluations decreased with class size. The courses included in their study ranged from 22 to 325 students. The range of course sizes in this study was not large with courses varying from 10 to 44 students. This lack of variability in course size could have contributed to not finding a difference in participation.

Is there nonresponse bias in online course evaluations? For nonresponse bias to exist, variables must be related to online course evaluation participation and ratings (Groves, 2006). Only course grade predicted both, suggesting it may be source of potential nonresponse bias. The findings from this study and others provide insight into student behavior when presented with an opportunity to complete an online course evaluation. Reviewing these results within the context of theoretical perspectives provides a different explanatory viewpoint on the results.

Theoretical perspectives

Organizational citizenship behavior. Connections can conceptually be made to the findings of this study and research examining the relationship between organizational citizenship behavior and survey participation. In the Spitzmeuller et al. (2006) study,

students who indicated that they would not respond to future survey requests rated their university lower on characteristics such as procedural justice, perceived organizational support, and social exchange. This study found that online course evaluation ratings and participation were both significantly related to course grade, with participation and ratings both increasing with course grade. When thinking about the dynamics in a classroom, one can make connections between students believing that they are treated fairly (procedural justice), and given necessary support to master material (perceived organizational support) with course grades. Given the dynamics in a classroom, there is the possibility that students who do not perform as well in courses may potentially rate procedural justice and perceived organizational support similar to the nonrespondents in Spitzmueller et al. (2006). Given the results of this study, testing the theory more directly may provide additional insight into the reasons for online course evaluation nonresponse.

Leverage salience theory. The connection between the results of this study and leverage salience theory is more direct. Leverage salience theory essentially states that the decision to complete a survey is based on a combination of factors, and additional factors such as incentives can influence the decision to participate in a survey (Groves et al., 2000). Previous research using leverage salience theory has demonstrated that individuals more interested in a topic were more likely to complete a survey about that topic without incentives (Groves et al., 2004). The results from this study showed that without any incentives, students who receive higher course grades were more likely to complete the online course evaluation. Although there were other characteristics that predicted participation and non-participation, course grade was the only variable that predicted course evaluation ratings, making it the only potential source of nonresponse

bias. As stated as an explanation for their findings, Groves et al. (2006) wrote, “when the topic of the survey is relevant to the sample person but generates negative thoughts, unpleasant memories, or reminders of embarrassing personal failings, then the topic may suppress participation despite its personal relevance” (p. 734). Although written to summarize the findings of their study, those words could apply to college student willingness to participate in course evaluation procedures, particularly when participation requires outside of class behaviors and initiative. Perhaps students who receive lower course grades do not want to be reminded about their courses, and simply want to complete their course work and not be bothered with completing an online course evaluation. Using leverage salience theory as a guide, this finding suggests that students who do not perform as well in a course may need additional incentives or “leverage” to encourage them to complete the online evaluation. Previous studies such as Dommeyer et al (2004) used incentives and information as leverage, but no studies have explored the impact of incentives and if they alleviate or contribute to nonresponse bias.

Limitations

Although this study found one source of nonresponse bias, there were certain characteristics unique to this study that may have influenced those results. For one, the study occurred at a small, private university with total enrollment of less than 4,000 students. The majority of studies reviewed in this research project occurred at larger universities (Avery et al., Dommeyer et al., 2004; Fidelman, 2007; Johnson, 2003; Layne et al., 1999; Leigle & McDonald, 2005; Thorpe, 2002). Another limitation of this study is that it did not examine differences by academic schools in addition to viewing results across the entire university. Although taking the approach to limit analyses to one

academic discipline would present a narrow picture of nonresponse in online course evaluations, conducting analyses for the entire population and by academic disciplines or schools would provide a complete picture for response patterns. This study also did not examine response rates across multiple semesters with the same professors (Avery et al., 2006), faculty characteristics such as length of time teaching or tenure, and course characteristics such as being an elective or requirement.

Future Studies

Given that online course evaluation results are often used in important decisions such as faculty tenure and promotion, working towards eliminating or at least minimizing nonresponse and potential nonresponse bias is important. Ideally all students would comply and complete all online evaluations. Because all students do not complete the evaluations and concerns about nonresponse bias exist, future research should address practical concerns about online course evaluations and underlying reasons for nonresponse.

The major practical concern is increasing response rates and making sure that all students participate. Conducting research to improve this issue will be valuable as more and more universities use online course evaluations. Dommeyer et al. (2004) is an example of such a study because the researchers offered a minute course grade incentive, which resulted in higher online course evaluation than paper evaluation response rates. Johnson (2003) also suggested holding reporting of early grades for students who did not complete the online course evaluation. Trying different approaches to increase participation of students and publishing those findings would help other universities that are struggling with the concern of nonresponse bias. For example, although constraints

may exist with extending the time into exams, perhaps online course evaluations could be made available earlier than the last few weeks of the semester. As described in Wachtel (1998), studies have shown that course evaluation ratings change little from the middle to the end of the course. Another practical suggestion is testing the impact of additional email reminders.

Understanding why certain characteristics become potential sources of nonresponse bias is just as important as understanding appropriate techniques to use for reducing nonresponse. Future studies could also use organizational research as a guide for understanding why students do not participate in online course evaluations. In the Spitzmueller et al. (2006; 2007) studies, students were asked their intentions to respond within the context of the organization. Future studies could also examine nonresponse within the context of student ratings of belongingness or use other frameworks that examine students' sense of connectedness to their educational experience such as student engagement.

There is also room for future studies to include dimensions that include course and faculty characteristics. Issues surrounding nonresponse could be related to electives versus non-electives. Faculty characteristics such as number of years teaching, full-time versus adjunct, or tenured versus untenured may impact online course evaluation ratings or participation.

Conclusions and Implications

The purpose of this study was to examine potential sources of nonresponse bias in online course evaluations. In this study, analyses found that course evaluation ratings and the odds of participating in online course evaluations increased with course grade.

These results suggest that students who are not performing as well in classes are less inclined to participate in the online course evaluation, resulting in course evaluation results that do not include the perspectives of those students. Based on these results, future studies should be conducted at this university to find strategies to increase the response rates of students, particularly those not inclined to participate, and to understand why students refrain from the online course evaluation process. Because this is the only study that has found a potential source of nonresponse bias in online course evaluations, it is suggested that future studies replicate the process to determine if these results are specific to this population or if they exist in other populations. This study should be repeated at this university to determine if course grade is a consistent source of nonresponse bias, or just an anomaly this one semester. In addition to course grade, gender and ethnicity were predictors of not participating in the online course evaluation with men and black students less likely to respond. Even though these characteristics were not considered sources of non-response bias because they did not predict participation and ratings, the findings once again suggest that only select portions of the population are participating in the process. Ultimately, with any university-wide process for improvement and feedback, participation is desired from all groups in the population.

Appendix A

Spring 2009 Student Evaluation of Course Effectiveness

This evaluation gives you the opportunity to comment on your experience in this course. We ask that you think about the impact this course has had on your learning. Please answer each item below honestly and thoughtfully, as your feedback is very important to us and will benefit students who take this course in the future. Summary results will be released to your instructor only after grades have been filed and will not be linked to individual student responses.

1) How would you rate (insert course)as a learning experience?

- Excellent
- Good
- Fair
- Poor
- Very Poor

2) What is your rating of (insert instructor's name)'s teaching?

- Excellent
- Good
- Fair
- Poor
- Very Poor

3) How much do you feel you have learned in (insert instructor's name)compared to other courses you have taken at Marymount?

- Much more than most courses
- More than most courses
- About the same as other courses
- Less than most courses
- Much less than most courses
- This is my first course at Marymount

4) The instructor was well prepared for each class with lectures, exercises, assignments, handouts, or other materials to support the day's topic.

- Strongly Disagree
- Disagree
- Neither Disagree Nor Agree
- Agree
- Strongly Agree

5) Instructor (insert instructor's name) explained course content in a way that I could understand.

- Strongly Disagree
- Disagree
- Neither Disagree Nor Agree
- Agree
- Strongly Agree

6) The instructor was responsive to student questions.

- Strongly Disagree
- Disagree
- Neither Disagree Nor Agree
- Agree
- Strongly Agree

7) Instructor (insert instructor's name) used the class time well (employed a balance of teaching methods, handled student questions efficiently, stayed on topic, etc.).

- Strongly Disagree
- Disagree
- Neither Disagree Nor Agree
- Agree
- Strongly Agree

8) The instructor encouraged me to take responsibility for my own learning.

- Strongly Disagree
- Disagree
- Neither Disagree Nor Agree
- Agree
- Strongly Agree

9) Instructor (insert instructor's name) encouraged students to participate actively in class.

- Strongly Disagree
- Disagree
- Neither Disagree Nor Agree
- Agree
- Strongly Agree

10) The feedback I received on my work in this course helped me learn.

- Strongly Disagree
- Disagree
- Neither Disagree Nor Agree
- Agree
- Strongly Agree

11) The standards for evaluating my work were consistently applied.

- Strongly Disagree
- Disagree
- Neither Disagree Nor Agree
- Agree
- Strongly Agree

12) What did you like about (insert instructor's name)and/or the instructor's teaching of it?

13) Which teaching methods or strategies contributed most to your learning (lectures, technology, group activities, projects, readings, handouts, etc.)?

14) What suggestions can you offer that would make this course a better experience?

15) Is this course:

- A required course for your major/program
- An elective
- A gen ed core requirement
- Other

16) On average, how many hours per week have you spent on this course outside of the class?

- Fewer than 2
- 2 to 4
- 5 to 9
- 10 to 12
- More than 12

17) How many sessions of this class have you missed this semester?

- 0 to 2
- 3 to 5
- 6 or more

18) What grade do you expect to receive in (insert course name) ?

- A
- A-/B+
- B
- B-/C+
- C/C-
- D or below
- Not Applicable

19) What is your current class level?

- Freshman
- Sophomore
- Junior
- Senior
- Graduate
- Other

20) If you have other comments about this course, please feel free to share them below.

Thank you for completing this evaluation! If you encounter any problems while completing this on-line evaluation, please contact the -----

Appendix B

FIRST EMAIL MESSAGE

Dear Marymount University Student:

Marymount University conducts a survey to evaluate the effectiveness of instruction at the end of every semester and summer session. Information from course evaluations is used for course improvement and faculty promotion decisions. Your feedback is very important to the academic process.

To complete the evaluation please enter the following information:

1. Course section: MU 101 A
2. Student ID number: |UNIQUEID|

Please go to [LINK](#) to complete the course evaluation.

If you have any questions or difficulty with this process, please contact the Office of Institutional Effectiveness at ira@marymount.edu or call 703-284-3809 for assistance.

REMINDER 1

Dear Marymount University Student,

This is just a reminder to complete your course evaluation for |UDF1|. Your feedback is very important.

To complete your course evaluation, you will need the following information:

1. Course section: |UDF1|
2. Student ID: |UNIQUEID|
3. Course evaluation link: |LINK1|

(Student ID number is used for administrative purposes only. Faculty only receive reports with aggregated results.)

If you have problems accessing the evaluation, please contact ira@marymount.edu or 703-284-3809.

FINAL REMINDER

Dear Marymount University Student,

The last day to complete your course evaluation is Tuesday, December 9. Do not miss this opportunity to share your feedback about the course. Course evaluation results are used by faculty for course improvement and as a component in tenure and promotion decisions.

Completing the course evaluation is a simple process. You just need three pieces of information:

1. Course section: |UDF1|
2. Student ID: |UNIQUEID|
3. Course evaluation link: |LINK1|

Once the course evaluation process ends, the Office of Institutional Effectiveness will analyze the results and give each faculty member a report with summary results. These reports do not include any individual identifying information.

If you have questions about the course evaluation process, please contact ira@marymount.edu or xxx-xxx-xxxx.

References

- Abraham, K., Maitland, A. & Bianchi, S. (2006). Nonresponse in the American time use survey: Who is missing from the data and how much does it matter. *Public Opinion Quarterly*, 70(5), 676-703.
- Anderson, H., Cain, J., & Bird, E. (2005). Online student course evaluations: Review of literature and a pilot study. *American Journal of Pharmaceutical Education*, 69(1), 34-43.
- Avery, R., Bryant, W., Mathios, A., Kang, H., & Bell, D. (2006). Electronic course evaluations: Does an online delivery system influence student evaluations. *Journal of Economic Education*, 37(1), 21-37.
- Camilli, G. (2006). Test fairness. In R. Brennan (Ed.), *Educational Measurement* (pp. 221-256). Westport, CT: American Council on Education and Praeger Publishers.
- d'Apollonia, S. & Abrami, P. (1997). Navigating student ratings of instruction. *American Psychologist*, 52(11), 1198-1208.
- Dillman, D., Eltinge, J., Groves, R. & Little, R. (2002). Survey nonresponse in design, data collection, and analysis. In D. Dillman, J. Eltinge, R. Groves & R. Little (Eds.). *Survey Response*, (pp. 3-26). New York: Wiley
- Dommeier, C., Baum, P., Chapman, K., & Hanna, R. (2002). Attitudes of business faculty towards two methods of collecting teaching evaluations: Paper vs. online. *Assessment & Evaluation in Higher Education*, 27(5), 455-462.

- Dommeier, J., Baum, P., Hanna, R., & Chapman, K. (2004). Gathering faculty teaching evaluations by in-class and online surveys: Their effects on response rates and evaluations. *Assessment & Evaluation in Higher Education, 29(5)*, 611-623.
- Donovan, J, Mader, C., & Shinsky, J. (2007). Online vs. traditional course evaluation formats: Student perceptions. *Journal of Interactive Online Learning, 6(3)*, 158-180.
- Fidelman, C. Course evaluation surveys: In-class paper surveys versus voluntary online surveys. Ph.D. dissertation, Boston College, United States -- Massachusetts. Retrieved July 27, 2009, from Dissertations & Theses: Full Text. (Publication No. AAT 3301790).
- Groves, R. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly, 70(5)*, 646-675.
- Groves, R., Couper, M., Pressor, S., Singer, E., Tourangeau, R. Acosta, G., & Nelson, L. (2006). Experiments in producing nonresponse bias. *Public Opinion Quarterly, 70(5)*, 720-736.
- Groves, R., Pressor, S., & Dipko, S. (2004). The role of topic interest in survey participation decisions. *Public Opinion Quarterly, 68(1)*, 2-31.
- Groves, R., Singer, E., & Corning, A. (2000). Leverage-saliency theory of survey participation description and an illustration. *Public Opinion Quarterly, 64*, 299-308.

- Hardy, N. (2003). Online ratings: Fact and fiction. In D. Sorenson & T. Johnson (Eds.), *New Directions for Teaching and Learning (96)*, (pp. 31-38). San Francisco: Jossey Bass.
- Hausmann, L., Schofield, J., & Woods, R. (2007). Sense of belonging as a predictor of intentions to persist among African American and White first-year college students. *Research in Higher Education, 48(7)*, 803-839.
- Heath, N., Lawyer, S., & Rasmussen, E. (2007). Web-based versus paper and pencil course evaluations. *Teaching of Psychology, 34(4)*, 259-261.
- Hoffman, K. (2003). Online course evaluation and reporting in higher education. In D. Sorenson & T. Johnson (Eds.), *New Directions for Teaching and Learning (96)*, (pp. 25-29). San Francisco: Jossey Bass.
- Johnson, Trav. (2003). Online student ratings: Will students respond. In D. Sorenson & T. Johnson (Eds.), *New Directions for Teaching and Learning (96)*, (pp. 49-59). San Francisco: Jossey Bass.
- Kim, Y., & Sax, L. (2009). Student-faculty interaction in research universities: Differences by student gender, race, social class, and first generation status. *Research in Higher Education, 50*, 437-459.
- Kuh, G. (2003). The national survey of student engagement: Conceptual framework and overview of psychometric properties. Retrieved from National Survey of Student Engagement website: [fhttp://nsse.iub.edu/pdf/conceptual_framework_2003.pdf](http://nsse.iub.edu/pdf/conceptual_framework_2003.pdf)
- Layne, B., DeCristoforo, J., & McGinty, D. (1999). Electronic versus traditional student ratings of instruction. *Research in Higher Education, 40(2)*, 221-231.

- Leigle, J. & McDonald, D. (2005). Lessons learned from online vs. paper-based computer information students' evaluation system. *Information Systems Education Journal*, 3(37), 1-14.
- Manfreda, L., Bosnjak, M., Berzelak, J., Haas, I., & Vehvovar, V. (2008). Web surveys versus other survey modes: A meta-analysis comparing response rates. *International Journal of Market Research*, 50(1), 79-104.
- Organ, D. (1988). *Organizational citizenship behavior: The good soldier syndrome*. Lexington, MA: Lexington Books.
- Ory, J. (1990) Student ratings of instruction: ethics and practice. In Theall, M. & Franklin, J. (Eds.) *Student ratings of instruction: issues for improving practice* (pp. 63–74). San Francisco, CA: Jossey-Bass.
- Podsakoff, P., MacKenzie, S., Moorman, R. & Fetter, R. (1990). Transformational leader behaviors and their effects on followers' trust in leader, satisfaction, and organizational citizenship behaviors. *Leadership Quarterly*, 1(2), 107-142.
- Porter, S., & Umbach, P. (2006). Student survey response rates across institutions: Why do they vary? *Research in Higher Education*, 47(2), 229-247.
- Porter, S., & Whitcomb, M. (2005). Nonresponse in student surveys: The role of demographics, engagement and personality. *Research in Higher Education*, 46(2), 127-151.
- Raudenbush, S. & Bryk, A. (2002). *Hierarchical linear models: Applications and data analysis methods*. (2nd ed.). Thousand Oaks, CA: Sage.
- Rhoades, L. & Eisenberger, R. (2002). Perceived organizational support: A review of the literature. *Journal of Applied Psychology*, 87(4), 698-714.

- Sax, L., Gilmartin, S., & Bryant, A. (2003). Assessing response rates and nonresponse bias in web and paper surveys. *Research in Higher Education, 44*(4), 409-432.
- Sax, L., Gilmartin, S., Lee, J., & Hagedorn, L. (2008). Using web surveys to reach community college students: An analysis of response rates and response bias. *Community College Journal of Research and Practice, 32*, 712-729.
- Shi, T., & Fan, X. (2008). Comparing response rates from web and mail surveys: A meta-analysis. *Field Methods, 20*(3), 249-271.
- Sorenson, D., & Reiner, C. (2003). Charting the uncharted seas of online student ratings of instruction. In D. Sorenson & T. Johnson (Eds.), *New Directions for Teaching and Learning (96)*, (pp. 1-24). San Francisco: Jossey Bass.
- Spitzmuller, C., Glenn, D., Barr, C., Rogelberg, S. & Daniel, P. (2006). If you treat me right I might reciprocate: Examining the role of exchange in organizational survey response. *Journal of Organizational Behavior, 27*, 19-35.
- Spitzmuller, C., Glenn, D., Sutton, M., Barr, C., & Rogelberg, S. (2007). Survey nonrespondents as bad soldiers: Examining the relationship between organizational citizenship and survey response behavior. *International Journal of Selection and Assessment, 15*(4), 449-459.
- Thorpe, S. (2002). *Online student evaluation of instruction: An investigation of non-response bias*. Paper presented at 42nd Annual Forum for the Association for Institutional Research, Toronto, Ontario, Canada.
- Wachtel, H. (1998). Student evaluation of college teaching effectiveness: A brief review. *Assessment & Evaluation in Higher Education, 23*(2), 191-21