Undergraduate perception of introductory lecture and laboratory biology instructors at the University of Tennessee Knoxville

A Dissertation Presented for the
Doctor of Philosophy
The University of Tennessee, Knoxville

Katharina Denise Kendall
May 2013

Copyright © 2013 by Katharina Denise Kendall
All rights reserved.

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the encouragement of my family, Christina and Nathan Kendall, who offered considerable support throughout my graduate program. It also would not have been possible without the guidance of my advisor Dr. Elisabeth Schussler, and committee comprised of Drs. Randall Small, Kristin Rearden, and Aimee Classen. I am grateful to Raina Akin and Kate Weaver for their assistant throughout this project as well as the other faculty, staff, graduate students, and undergraduates in the Division of Biology and Department of Ecology and Evolutionary Biology who assisted and participated in my research at UTK. I also thank my fellow graduate students and faculty members in the Ecology and Evolutionary Biology Department for their feedback and collaboration as well as help administering surveys, particularly Matthew L. Niemiller, Dylan Dittrich-Reed, Jim Fordyce, Benjamin Fitzpatrick, Amanda Allison, Lisa Cantwell, Laura Rubio, and others who I have inadvertently forgotten. Funding for this dissertation was provided by the Division of Biology and Department of Ecology and Evolutionary Biology at the University of Tennessee, Knoxville.

ABSTRACT

Undergraduate students entering the higher education system are often unaware of the diverse teaching and learning community they will encounter, including the different instructor types who will teach their classes. In order to accommodate the growing numbers of enrolled students, the higher education system is increasingly reliant on contingent instructors such as non-tenure track faculty members and graduate teaching assistants (GTAs). This dissertation explores undergraduate student perspective of the different instructor types who teach introductory biology courses, with a focus on the University of Tennessee, Knoxville (UTK). The goal of this work is to provide insight regarding how perceived differences between GTAs and professors impact teaching effectiveness and student learning in order to make recommendations for professional development. Chapter one outlines a study that utilized qualitative and quantitative research methodologies to explore undergraduate perception of hypothetical GTAs versus professors to determine if students perceive differences that are independent of classroom setting. The second chapter describes a follow up study to further explore instructor type differences by having undergraduates rate their actual biology instructors (faculty members and GTAs). Concurrently, interviews were conducted to gain perspective about the instructional behaviors students perceived to be associated with words used to describe higher education instructors; results from this study can be found in chapter three. Chapter four reports on a quantitative study to identify the instructional behaviors which best predict teaching effectiveness of GTAs. Collectively, these studies provide critical insight into the impact of instructor type on instruction, and student learning, in

introductory biology courses. The results of this work are thereby used to make recommendations for GTA professional development, given that many GTAs are current instructors and could be future faculty, to promote and enhance student learning in the biological sciences.

TABLE OF CONTENTS

General Introduction 1
Chapter One 3
Chapter Two 4
Chapter Three 6
Chapter Four 7
Literature Cited 9
Chapter One 12
Abstract 13
Introduction 14
GTAs as Contingent Instructors 15
GTA Professional Development 16
Instructional Characteristics of GTAs 19
Project Rationale 21
Materials and Methods 22
Data Collection 22
Survey Validity and Reliability 25
Data Analysis 26
Results 29
Participants 29
Quantitative Analysis 30
Qualitative Analysis 32
Discussion 34
Proposed Explanations of Instructor Differences 38
Limitations 40
Recommendations for GTA Professional Development 41
Conclusions and Future Directions 43
Literature Cited 44
Appendix 49
Chapter Two 55
Abstract 56
Introduction 57
GTA Use: Benefits and Concerns 58
Students' Evaluations of Teaching 60
Project Rationale 61
Materials and Methods 63
Data Collection 63
Survey Design 64
Survey Validity and Reliability 66
Data Analysis 67
Results 69
Participants 69
Discussion 76

Conclusion 82
Literature Cited 83
Appendix 88
Chapter Three 98
Abstract 99
Introduction 100
Instructor Descriptions 101
Instructional Behaviors 102
Student Evaluations of Teaching 104
Project Rationale 105
Materials and Methods 106
Data Collection 106
Data Analysis 108
Results 110
Participants 110
Student Explication of Instructor Descriptions 111
Themes and Schema 118
Discussion 119
Changing Instructional Behaviors 120
Relationships Among Behaviors 121
Limitations 123
Conclusion 124
Literature Cited 124
Appendix 128
Chapter Four 132
Abstract 133
Introduction 133
Project Rationale 136
Methods 137
Data Analysis 139
Results 139
Conclusions 141
Limitations and Future Direction 144
Summary 145
Literature Cited 146
Appendix 151
General Conclusions 155
Literature Cited 161
Vita 163

LIST OF TABLES

Table 1. The CUCEI subscales and items used for the study 49
Table 2. QTI subscales and items used in the study 50
Table 3. Demographic percentages for overall study participants, as well as by scenario51
Table 4. Significant results for the Kruskal-Wallis ($\alpha < 0.05$) and Mann-Whitney U ($\alpha < 0.013$ with the Bonferroni correction) nonparametric tests 52
Table 5. Descriptive statistics for significant items for each scenario, including mean ± standard deviation, median, maximum, and minimum. 53
Table 6. Summary of themes (delivery technique, classroom atmosphere, and relationship) obtained from qualitative data in both the teaching and personal realms for GTAs and professors 54
Table 7. Survey items asking students to rate their instructor for thirteen descriptors as well as teaching effectiveness and student learning 88
Table 8. Descriptive statistics including mean, standard deviation (SD), median (med), minimum (min), and maximum (max) by instructor type for initial and final ratings of descriptors (in alphabetical order). 89
Table 9. Mean ± standard error for all initial ratings of descriptors (in alphabetical order) that had significant associations with (A) teaching experience, (B) instructor gender, or (C) instructor age (Spearman rank correlations, $\alpha = 0.05$). 95
Table 10. Mean ± standard error for all final ratings of descriptors (in alphabetical order) that had significant associations with (A) teaching experience, (B) instructor gender, or (C) instructor age in final ratings (Spearman rank correlations, $\alpha = 0.05$). 96
Table 11. Significant Spearman ($\alpha = 0.05$) associations of descriptors with instructor demographics for initial and final ratings. Non bold descriptors have weak associations, italic descriptors have moderate associations, and bold descriptors remain significant throughout the semester. 97
Table 12. Student instructor descriptions grouped into three overarching themes (delivery technique, classroom atmosphere, and relationship) within the teaching and personal realms of instructor type (modified from Kendall and Schussler, 2012)..... 128
Table 13. Undergraduate explanations of descriptive words, sorted into paired, positive and negative aspects. 129
Table 14. Undergraduate explanations of the descriptive words “organized” and “respect,” in which there were no negative terms or similar words associated with them. 130
Table 15. Twenty-one instructional behaviors included in the survey, divided into their respective sub-scales (teaching techniques, interpersonal rapport, passion for subject). 153
Table 16. Description of fixed effects and random intercepts for full and reduced models including DF, log (Likelihood), AIC, and BIC results. 154

LIST OF FIGURES

Figure 1. Mean ± standard error for descriptors (in alphabetical order) in which GTAs differed significantly from faculty members for initial ratings (Wilcoxon signed-rank tests [α = 0.005]). 90
Figure 2. Mean ± standard error for descriptors (in alphabetical order) in which GTAs differed significantly from faculty members for final ratings (Wilcoxon signed-rank tests [α = 0.005]). 91
Figure 3. Descriptors (in alphabetical order) in which significant changes in mean student ratings (from 1 – 6 +/- standard error) were found over the semester for GTAs only. GTAs are depicted with dashed lines and Faculty with solid lines. 92
Figure 4. Descriptors (in alphabetical order) in which significant changes in mean student ratings (from 1 – 6 +/- standard error) were found over the semester for Faculty only. GTAs are depicted with dashed lines and Faculty with solid lines. 93
Figure 5. Descriptors (in alphabetical order) in which significant changes in mean student ratings (from 1 – 6 +/- standard error) were found over the semester for GTAs and Faculty. GTAs are depicted with dashed lines and Faculty with solid lines. 94
Figure 6. Schema depicting positive instructional behaviors within the three themes (teaching techniques, interpersonal rapport, and passion for subject) identified from student perspectives, which lead to the overarching description of respect. 131

GENERAL INTRODUCTION

Higher education institutions utilize many different instructor types, including contingent instructors such as part-time (adjunct), non-tenure track faculty members, teaching postdoctoral fellows, and graduate teaching assistants (GTAs) to deliver courses to their undergraduate students. Recent surveys have found that contingent instructors represent nearly half the instructional staff at academic institutions (Jaeger, 2008; Baldwin and Wawrzynski, 2011). GTAs often represent a large part of these numbers, Sundberg et al. (2005) reported that 71% of undergraduate laboratory sections at comprehensive institutions and 91% at research institutions utilize GTAs for teaching (n = 65), and another survey of 153 graduate degree-granting institutions reported that 97% employ GTAs (Rushin et al., 1997).

GTAs have ambiguous roles within the higher education system. They are admitted to universities as students to pursue a graduate education, yet they are also often used as part-time instructors to provide instruction to undergraduate students. Departments may also have additional tasks for GTAs, including grading, handing out course material, and proctoring exams, making a generalization about typical GTAs duties difficult (Marincovich et al., 1998). Given the multiple roles GTAs play as students, researchers, and instructors in the academic learning community administrators are often unsure whether to classify GTAs as students or staff (Flora, 2007). Due to their ambiguous roles, however, GTAs play a key role in connecting the higher education community, which is comprised of teaching, research, and service (Sheridan, 1991).

Many administrators are aware that GTAs will be future members of academia; however, there is still a common view that no formal training is required to prepare GTAs for teaching (Rushin et al., 1997). When universities do offer training for GTAs it is typically a pre-academic year workshop (Rushin et al., 1997) or orientations focused on departmental and university policies with little focus on pedagogy. Frequently absent is background on how to deliver specific curricula, course planning, interdisciplinary connections, and assessment techniques (Marincovich et al., 1998; Shannon et al., 1998; Luft et al., 2004). Yet, there are also institutions which offer options such as formal courses, written training manuals, and pedagogical seminars for GTAs (Rushin et al., 1997; Golde and Dore, 2001). The offering of professional development at many institutions indicates that they are aware that GTAs should be provided with training; however, the diversity of options suggests that more research must be done in order to determine what aspects of training are most beneficial to GTAs.

This dissertation explores undergraduate student perception of instructors teaching introductory biology courses at the University of Tennessee, Knoxville (UTK). The goal of the dissertation is to indentify whether instructor type, title, and behaviors are perceived differently between GTAs and faculty members in order to make recommendations for professional development specific to GTAs. This in depth exploration will lead to the identification of specific training needs for GTAs at UTK, while providing insight for other institutions; however, other institutions must be aware that there is no single recipe for GTA fprofessional development and each institution must consider its own culture when establishing and modifying options available to their

GTAs (Hiiemae et al., 1991). Thus, a clarification of the culture present at UTK is essential in order for those from other institutions to make decisions regarding how this work may inform their own school.

UTK is a large research institution located in the southern United States. UTK enrolls about 21,000 undergraduate students, 9,000 graduate students, and has a student-instructor ratio of 15 to 1 (Undergraduate Admissions, 2012). The Division of Biology is based in the College of Arts and Sciences and encompasses three departments including Microbiology, Ecology and Evolutionary Biology, and Biochemistry, Cellular, and Molecular Biology. All students in these departments received a Biological Sciences degree upon graduation, with a concentration in one of the three departments. At any one time, there are about 800 declared majors in Biology. Students seeking a degree in Biology take a common set of core courses: Biodiversity, Cell Biology, Genetics, and Ecology. Each 4-credit course is delivered in lecture sections ranging in size from 120-225 students, with associated laboratory sections with 25 students each. Faculty members teach all of the lecture courses, and GTAs teach all of the laboratories. Instructors (faculty [tenure-track and lecturers] and GTAs) teaching these core courses at UTK are primarily from the three biology departments.

Chapter One

In previous work, Muzaka (2009) and Park (2002) explored the strength and weaknesses of GTAs as instructors, encompassing views from undergraduate students, graduate students, and faculty members. These studies found that GTAs are often thought of as beneficial as instructors because they are approachable, informal, understanding,

flexible, identifiable by undergraduates, less intimidating, enthusiastic, and good role models, yet problematic as instructors because they are nervous, limit access and contact time of students to academic experts, lack subject knowledge, teaching experience, confidence, and authority.

This study utilized two validated survey instruments to compare how students would rate hypothetical GTAs and professors teaching laboratory classes and discussion classes, in order to identify distinguishing characteristics of each instructor type. While professors and GTAs may teach similar content and students, they often teach different types of classes which may confound results comparing the two instructor types. Thus, data were collected on undergraduate student perception of each type of instructor in each of two classroom situations (laboratory and discussion). I hypothesized that there would be factors that differed between professors and GTAs. Undergraduate students enrolled in introductory biology courses at UTK in fall 2010 provided quantitative and qualitative data that were able to identify factors unique to each instructor type.

Chapter Two

This chapter builds on the descriptors (instructor characteristics) identified from Chapter One, which were associated with different instructor types (GTAs and professors). Specifically, undergraduates viewed professors as being more structured, confident, in control, organized, experienced, knowledgeable, distant, formal, strict, serious, hard, boring, out of touch, and respected than GTAs. Conversely, GTAs were perceived as being more uncertain, hesitant, nervous, relaxed, laid-back, engaging, interactive, relatable, understanding, and able to personalize teaching than professors.

The goal of this chapter was to determine if these differences in descriptors held true if students were asked to rate their current professors and GTAs regarding each of these descriptors at the beginning and end of the semester.

Instructors of all types often receive evaluations of teaching from their students. While these evaluations are often used for purposes of retention or promotion, they also provide invaluable insight into instruction from the student perspective (Baird, 1987; Emery et al., 2003; Clayson et al., 2006; Wright et al., 2006; Zabaleta, 2007; Helterbran, 2008; Kogan et al., 2010). While these evaluations of teaching may be confounded by instructor demographics such as age, gender, and teaching experience, these variables can be accounted for in studies utilizing students' evaluations of teaching effectiveness.

This study utilized a survey using the instructor descriptors obtained in Chapter One. It asked students in eight lecture classes, and the associated lab sections, to rate their professor and GTA for each of the characteristics at the beginning and then at the end of the semester. Not only did this study seek to determine if instructor title impacts student perception in practice, but it sought to determine if these perceptions change over the course of a semester. Also, the study sought to determine if instructor descriptors are associated with instructor demographics (age, gender, and university teaching experience), or with student perception of teaching effectiveness or perception of their learning. I hypothesized that some instructional descriptors would hold true in practice, that student perception would change over the course of the semester, that student perception may be influenced by instructor demographics, and that these instructor descriptors would be correlated with perceived instructor effectiveness and student

learning. Undergraduate student ratings of their instructors revealed that student perception of instructors evolves over the course a semester.

Chapter Three

Student evaluation of their instructors is assumed to reflect student perception of teaching effectiveness and student learning; however, it is more likely that the results are a reflection of the complex interrelationship between students and instructors than strictly teaching factors such as knowledge, quality of instruction, and teaching expertise on the part of the instructor (Zabaleta, 2007). Thus, identification of instructor behaviors that contribute to excellent or poor student evaluations would provide invaluable insight into student perception of effective and ineffective teaching.

Pattison, Hale, and Gowens (2011) identified behaviors specific to effective teaching, including being knowledgeable, making material relevant, being prepared, having a plan, integrating material from various disciplines, being organized, and encouraging student questions. They also identified that some instructional behaviors, such as being tougher than necessary, not providing feedback, being unprepared or impatient, are perceived as ineffective in the educational environment (Pattison et al., 2011). Further clarification of these instructional behaviors would be beneficial to instructors wishing to interpret student comments of their teaching or improve student perception of their teaching (Pattison et al., 2011), and would also be useful for making targeted recommendations for GTA professional development.

This chapter asked students to provide qualitative descriptions of the instructor descriptors identified in Chapter One and utilized in Chapter Two to identify the

underlying instructional behaviors associated with each description. These results were then used to create a schema of relationships among the instructor characteristics. Through these undergraduate descriptions, this research identified specific behaviors that could be the focus of professional development to improve student learning in academic environments.

Chapter Four

The objective of this chapter was to further explore the themes (teaching techniques, interpersonal rapport, and passion for subject) identified in the schema of Chapter Three. Specifically, it researched whether these three themes are weighted equally for student perception of teaching effectiveness, or if certain themes are more important.

A student survey consisting of 21 instructional behaviors from the three identified themes was distributed to undergraduates enrolled in introductory biology laboratory courses (non-majors biology, biodiversity, and cellular biology). Students were also asked to rate their perception of their GTA's teaching effectiveness, their perceived grade in lab, and the amount they learned in the lab. Linear mixed models including student demographics (e.g., enrollment status, language, gender), perceived grade in lab, perceived student learning, and instructional behavior ratings as fixed effects, and teaching effectiveness as the response variable, were used to determine which theme(s) best predicted GTA teaching effectiveness. I hypothesized that the best model would not include any student demographics and would also exclude the passion for teaching

theme. Results from this research identified four topics that could be the focus of professional development sessions to support GTA teaching effectiveness.

Literature Cited

Baird JS (1987). Perceived learning in relation to student evaluation of university instruction. J Educ Psychol *79*, 90–91.

Baldwin RG, Wawrzynski MR (2011). Contingent faculty as teachers: what we know; what we need to know. Am Behav Sci *55*, 1485–1509.

Clayson DE, Frost TR, Sheffet MJ (2006). Grades and the student evaluation of instruction: a test of the reciprocity effect. Acad Mgmt Learn Educ *5*, 52–65.

Emery CR, Kramer TR, Tian RG (2003). Return to academic standards: a critique of students' evaluations of teaching effectiveness. Qual Assur Educ *11*, 37–46.

Flora BH (2007). Graduate assistants: students or staff, policy or practice? The current legal employment status of graduate assistants. J High Educ Pol Manag *29*, 315–322.

Golde CM, Dore TM (2001). At Cross Purposes: What the Experiences of Doctoral Students Reveal about Doctoral Education, Philadelphia, PA: Pew Charitable Trusts.

Helterbran VR (2008). The ideal professor: student perceptions of effective instructor practices, attitudes, and skills. Education *129*, 125–138.

Hiiemae K, Lambert L, Hayes D (1991). How to establish and run a comprehensive teaching assistant training program. In JD Nyquist, RD. Abbot, DH Wulff, & J Sprague (Eds.), Preparing the Professorate of Tomorrow to Teach (pp. 123-134). Dubuque, IA: Kendallmunt.

Jaeger AJ (2008). Contingent faculty and student outcomes. Academe *94*, 42–43.

Kogan LR, Schoenfeld-Tacher R, Hellyer PW (2010). Student evaluations of teaching: perceptions of faculty based on gender, position, and rank. Teach Higher Educ *15*, 623–636.

Luft JA, Kurdziel JP, Roehrig GH, Turner J (2004). Growing a garden without water: graduate teaching assistants in introductory science laboratories at a doctoral/research university. J Res Sci Teach *41*, 211–233.

Marincovich M, Prostko J, Stout F (eds.) (1998). The Professional Development of Graduate Teaching Assistants, Balton, MA: Anker Publishing.

Muzaka V (2009). The niche of graduate teaching assistant (GTAs): perceptions and reflections. Teach High Educ *14*, 1–12.

Park C (2002). Neither fish nor fowl? The perceived benefits and problems of using graduate teaching assistants (GTAs) to teach undergraduate students. High Educ Rev *35*, 50–62.

Pattison P, Hale JR, Gowens P (2011).Mind and soul: connecting with students. J Legal Stud Educ *28*, 39–66.

Rushin JW, Saix JD, Lumsden A, Streubel DP, Summers G, Bernson C (1997). Graduate teaching assistant training: a basis for improvement of college biology teaching and faculty development? Am Biol Teach *59*, 86–90.

Shannon DM, Twale DJ, Moore MS (1998). TA teaching effectiveness: the impact of training and teaching experience. J High Educ *69*, 440– 466.

Sheridan JD (1991). A proactive approach to graduate teaching assistants in the research university: One graduate dean's perspective. In JD Nyquist, RD Abbot, DH

Wulff, & J Sprague (Eds.), Preparing the Professorate of Tomorrow to Teach (pp. 24-28). Dubuque, IA: Kendallmunt.

Sundberg MD, Armstrong JE, Wischusen EW (2005). A reappraisal of the status of introductory biology laboratory education in U.S. colleges and universities. Am Biol Teach *67*, 525–529.

Undergraduate Admissions (2012). Make Your Own Tradition. Retrieved from http://admissions.utk.edu/undergraduate/viewbook/Viewbook2012.pdf. Accessed 21. January. 2012.

Wright RE (2006). Student evaluations of faculty: Concerns raised in the literature, and possible solutions. Coll Student J *40*, 417-422.

Zabaleta F (2007). The use and misuse of student evaluations of teaching. Teach Higher Educ *12*, 55–76.

CHAPTER ONE

Does instructor type matter? Undergraduate student perception of graduate teaching assistants and professors

The following chapter has been published:

Kendall KD, Schussler EE (2012). Does instructor type matter? Undergraduate student perception of graduate teaching assistants and professors. CBE-Life Science Educ *11*, 187-199.

Abstract

Graduate teaching assistants (GTAs) are used extensively as instructors in higher education, yet their status and authority as teachers may be unclear to undergraduates, to administrators, and even to the GTAs themselves. This study explored undergraduate perception of classroom instruction by GTAs and professors to identify factors unique to each type of instructor, versus the type of classes they teach. Data collection was via an online survey composed of subscales from two validated instruments, as well as one open-ended question asking students to compare the same class taught by a professor versus a GTA. Quantitative and qualitative results indicated that some student instructional perceptions are specific to instructor type, and not class type. For example, regardless of type of class, professors are perceived as being confident, in control, organized, experienced, knowledgeable, distant, formal, strict, hard, boring, and respected. Conversely, GTAs are perceived as uncertain, hesitant, nervous, relaxed, laid-back, engaging, interactive, relatable, understanding, and able to personalize teaching. Overall, undergraduates seem to perceive professors as having more knowledge and authority over the curriculum, but enjoy the instructional style of GTAs. The results of this study will be used to make recommendations for GTA professional development programs.

Introduction

There is a growing dependence on contingent instructors (part time, non-tenure track faculty, and graduate teaching assistants [GTAs]) at research universities (Johnson, 2011); specifically, Baldwin and Wawrzynski (2011) and Jaeger (2008) indicate that these contingent instructors may contribute to roughly half of the instructional staff. According to one survey in the biological sciences discipline in the United States (n = 65), GTAs are responsible for teaching 71% of undergraduate laboratory sections at their comprehensive institutions and 91% at their research institutions (Sundberg et al., 2005). Similarly, Rushin et al. (1997) found that 97% of 153 graduate schools surveyed in the United States used GTAs to teach laboratories and/or lectures in biology courses. While there may be variability in these numbers from institution to institution, they highlight the reliance of higher education on GTA employment.

The majority of contingent instructors provide instruction to lower-division courses, making the likelihood of contact with them greatest for first year students. Given the importance of first year coursework to student retention, understanding the impact of contingent instructors on student learning is critical (Benjamin, 2002; Jaeger, 2008). However, few studies have been conducted on the quality of teaching students receive from contingent instructors, and those that have been conducted report conflicting results or are limited by sample size (Umbach, 2007).

For instance, Bolge (1995) reported that contingent instructors do not differentially affect student outcomes, because there were no differences in final grades.

Yet Johnson (2011) determined that contingent instructors typically give higher grades. It has also been reported that full-time tenure-track faculty devote proportionately more time to students than do contingent instructors (Benjamin, 2002). Similarly, Jaeger (2008) noted that contingent instructors are generally less accessible and less available even though students note that out-of-class interactions are most important for their education. Umbach (2007) suggested that contingent instructors are typically less effective in how they work with undergraduate students than tenured/tenure track faculty. Yet, Johnson (2011) reported that instructor type does not impact student retention, while O'Neal et al. (2007) found that GTA enthusiasm increased the likelihood of student retention. Baldwin and Wawrzynski (2011) posited that there is sufficient evidence to merit concern about the teaching quality of contingent instructors, and they are proponents of targeted professional development strategies for different types of contingent instructors. For this reason, this study focuses on identifying factors that undergraduates perceive as different between GTAs and professors in order to make suggestions specifically for GTA professional development.

GTAs as Contingent Instructors

While GTAs are used extensively as undergraduate instructors, they are admitted to universities to pursue graduate education at their institutions. Hence, some define a GTA as a graduate student pursuing a master's or doctoral degree who is used part-time to provide instruction to undergraduates, while also being involved in research and departmental degree requirements (Bos et al., 1980; Park and Ramos, 2002). However,

departments may have jobs for GTAs other than instruction, which makes a generalizable definition difficult (Marincovich et al., 1998).

The perceived role of a GTA at an institution varies based on who is asked. For instance, undergraduates perceive GTAs as holding a status between students and academics, while GTAs see themselves as students with teaching responsibilities (Park, 2002; Muzaka, 2009). Faculty members consider GTAs to be research students who are also academic apprentices (Park, 2002; Muzaka, 2009). Their role is sometimes puzzling to administrators and policy makers who are often unsure whether to classify GTAs as students or staff (Flora, 2007). Some studies attempting to clarify GTA responsibilities have gone as far as calling GTAs "donkeys of the department" due to their immense workload, level of responsibility, and restricted autonomy (Park and Ramos, 2002). Thus, GTAs occupy a unique position in the academic system as both researchers and faculty/teachers in training, whether by choice or not (Golde, 1998; Marincovich et al., 1998; Park, 2002; Park and Ramos, 2002; Muzaka, 2009).

GTA Professional Development

Given the reliance on GTAs for a majority of science laboratory teaching at U.S. universities, the need for effective professional development is a necessity. GTAs want preparation and guidance throughout their teaching experiences to improve not only their instructional ability but also their overall teaching experience (Bond-Robinson and Rodriques, 2006). Given that GTAs may be future faculty members at academic institutions, there is also a need to carefully mentor GTAs and provide them with advanced instructional assignments as they progress in their teaching experience (Braxton

et al., 1995); however, GTA assignments are typically made to cover departmental needs and not the professional development needs of future faculty members (Austin, 2002). The reality is that GTAs often feel unprepared for their teaching assignments (Dudley, 2009), and most universities offer no formal professional training at all (Rushin et al., 1997).

A study by Rushin et al. (1997) surveyed graduate schools (n = 153) in the United States to determine what training opportunities they offer to GTAs. Although the most common response was that no formal training was required, the second most common approach was a pre-academic year workshop (Rushin et al., 1997). The next most common GTA development opportunities, in decreasing prevalence, were: a semester long college teaching seminar, a formal college teaching course, and training by a professor (Rushin et al., 1997). Typical professional development activities included teaching multiple courses during the graduate program, videotaping for self evaluation, written training manuals, weekly meetings prior to teaching, and semester long courses (Rushin et al., 1997), activities similar to those suggested by graduate students to improve their professional development as teachers (Nyquist et al., 1999). These findings are also similar to Marincovich et al. (1998), who found that the topics most commonly addressed in GTA professional development included teaching in a nontraditional setting, ethical issues, communication skills, developing reflection habits (including evaluating teaching skills), and obtaining frequent feedback (including close mentorship).

The Golde and Dore (2001) Pew Charitable Trusts study also asked doctoral students (n = 4114) from 11 Arts and Sciences departments about preparation for

teaching. More than half of these doctoral students indicated they were required to teach during their degree program, and these students also reported that they were interested in, prepared for, and confident in teaching laboratory sections, and lecture courses and in leading discussions (Golde and Dore, 2001). Yet, Golde and Dore (2001) were unable to determine whether this confidence was merited since graduate students clearly noted that their programs did not prepare them for these instructional roles.

Several researchers have indicated that even if departments provide instructional training, GTAs often receive minimal amounts of the pedagogical and content instruction information that they need (Marincovich et al., 1988; Shannon et al., 1998; Luft et al., 2004). Frequently absent from their training is the necessary background on delivering specific curriculums, course planning, and assessment, or skills such as interdisciplinary connections, interactive pedagogy, instructional design, and teamwork (Marincovich et al., 1998; Shannon et al., 1998; Luft et al., 2004).

The offering of GTA professional development indicates that universities understand the importance of training; however, the diversity of training available, and the reported lack of consistency in GTA training manuals (Lowman and Mathie, 1993) also indicates that more research needs to be done on what aspects of training are most beneficial to GTAs. Shannon et al. (1998), for instance found that techniques such as microteaching, coteaching, and practice simulations were the most important aspects of developing teaching effectiveness in GTAs. They argue, however, that more progress will be made in determining the aspects essential to training when agreement is reached regarding the purpose and definition of GTA professional development (Shannon et al.,

1998). This study responds to this call by identifying factors that undergraduates perceive to be different between GTAs and professors as one way to better understand what aspects need to be addressed in GTA professional development.

Instructional Characteristics of GTAs

Given the multiple roles of GTAs at universities, investigation of how the instructional role of a GTA might differ from that of a professor at a university seems warranted. Studies on this topic have been limited by several factors, however, including the overarching fact that many undergraduates are often unaware that those teaching them are GTAs (Park, 2002). There are also cultural differences, for instance, quality assurance, employment practices, and graduate recruitment procedures, between the higher education systems in which these studies have been performed (United States vs. United Kingdom) that hinder comparability of studies in different countries (Park, 2002). GTA studies are also difficult to compare because most literature has focused on instructors who are tenure line or ranked faculty members and few compare them directly with GTAs (Bos et al., 1980). GTAs also often teach different types of classes (e.g., laboratories or small discussions) than do faculty members.

However, several studies have looked specifically at GTA instructional characteristics and can be used to inform studies on this topic. One such study conducted by Dudley (2009) explored the role of a GTA, from the perspective of GTAs, to determine how they are able to balance their many responsibilities and still remain effective in the classroom. GTAs reported experiencing nervousness and intimidation as well as concern that their teaching contradicted that of the professor. Dudley (2009)

acknowledged that teaching requires training, practice, and revision and that many GTAs draw from their own experiences as students to determine the best teaching style for their own classes. An important concluding factor is that GTAs must be able to overcome jitteriness while establishing authority in the classroom in order to be successful in the teaching environment (Dudley, 2009).

In another study, Park (2002) asked staff and GTAs to identify one positive and one negative aspect of GTA teaching. The positive aspects determined by this study included GTA characteristics of being approachable, informal, identifiable by undergraduates, enthusiastic, and good role models. Negative aspects included that the use of GTAs limits the access and contact time of students to academic experts (in terms of subject knowledge and teaching experience). Muzaka (2009) used methods similar to Park (2002), yet this study also took into account the perspective of undergraduates. Undergraduates, GTAs, and staff responded to open-ended questions asking about the most beneficial and most problematic aspects of GTA-led group seminars. Muzaka (2009) found that, from the undergraduate perspective, GTAs lack overall subject knowledge, teaching experience, confidence, control and authority, and are nervous. Yet undergraduates also indicated that GTAs identify easily with undergraduates and are understanding, flexible, informal, enthusiastic, approachable, and less intimidating. Characteristics identified from GTA and staff perspectives included GTA lack of ability to guarantee consistency, lack of overall subject knowledge, and limited teaching skills (Muzaka, 2009).

While the previous studies were qualitative, Bos et al. (1980) quantitatively analyzed student ratings of their GTAs to identify characteristics that may impact the effectiveness of instruction. This study resulted in the generalizations that GTA ratings are not impacted by undergraduate student major and gender of the GTA, yet they are influenced by degree held (higher degrees result in higher ratings), teaching experience, and GTA age (those in their late twenties are rated higher than those in their early twenties; Bos et al., 1980).

Project Rationale

Previous studies have found that there are instructional characteristics of GTAs that impact the perception of teaching and learning at universities; however, there have been few direct comparisons of student perceptions of professors and GTAs, particularly those employing the same methodologies on the same undergraduate student population. The goal of this study is to provide a quantitative and qualitative comparison of professors and GTAs from the perspective of undergraduate students at one institution in order to identify factors useful for GTA professional development. Although professors and GTAs may teach similar content and students, they often teach different types of classes, which could confound the identification of variables specific to each type of instructor. Thus, this study is designed to identify factors unique to each type of instructor, versus the type of classes they teach, by collecting data on undergraduate student perception of each type of instructor in each of two classroom situations (laboratory and discussion). The hypothesis for this study is that undergraduates will

perceive a difference between professors and GTAs related to the instructor and not just the classroom environment.

Materials and Methods

Data Collection

Data were collected through an online survey (hosted by surveymonkey.com) composed of subscales from two published and validated instruments, the College and University Classroom Environment Inventory (CUCEI) and the Questionnaire of Teacher Interaction (QTI).

These instruments were chosen for this study as they best captured the previously identified instructional characteristics of GTAs (e.g., uncertainty, approachable, etc.), and both had been used individually and together in previous studies in the college environment (Treagust and Fraser, 1986; Coll et al., 2002). The CUCEI was designed to incorporate the dimensions identified by Moos (1979) in his work on classroom environment, and includes the subscales of: Personalization, Innovation, Student Cohesiveness, Satisfaction, Task Orientation, Innovation, and Individualization (Treagust and Fraser, 1986; Coll et al., 2002). The QTI was originally designed in the Netherlands to explore how individuals mutually influence each other and consists of 77 items; however, it was shortened to 48 items for use in the Australian science education environment (Coll et al., 2002). The shortened QTI covers the subscales of: Leadership, Understanding, Uncertain, Admonishing, Helpful/Friendly, Student Responsibility and Freedom, Dissatisfied, and Strict (Coll et al., 2002).

Using both the CUCEI and QTI instruments in their entirety would result in a survey requiring responses to 97-items, not all of which were necessarily relevant to the study. Given the voluntary nature of the survey, and since not all of the subscales were directly related to potential differences between GTAs and professors (based on previous research), it was determined that only certain subscales of each instrument would be used in the study. Therefore, this project utilized only the CUCEI subscales of Personalization, Involvement, Task Orientation, and Individualization. For the QTI, only the sub-scales of Leadership, Uncertain, Helpful/Friendly, and Strict were used. The subscales were chosen to best capture aspects that were similar to the results of previous research, such as positive aspects of GTA instructors (e.g., approachability, informality, enthusiasm, and less intimidating, which related to the CUCEI and QTI subscales of Helpful/Friendly, Personalization, Task Orientation, Leadership, and Strict), and negative aspects (e.g., lack of experience, control, and content knowledge, and limited contact time, which related to the CUCEI and QTI subscales of Leadership, Task Orientation, Uncertain, Individualization, and Involvement) (Bos et al., 1980; Park, 2002; Dudley, 2009; Muzaka, 2009). The subscale questions, as well as the Likert choices undergraduates were given for each, are shown in Tables 1 and 2.

The survey was sent via email in September 2010 (4 weeks into the Fall semester) to undergraduate students enrolled in majors and nonmajors general biology courses (introductory biology, plant biology, biodiversity, cell biology, genetics and ecology) at a large research institution in the southern United States. Each of the selected courses has a lecture and laboratory component; the laboratories are all taught by GTAs, who teach two

or three laboratory sections per semester, and the lecture classes are all taught by PhD-level faculty (tenureline or lecturers). Undergraduate students received the email, which contained a description of the project and a link to the survey, from their lecture instructor on behalf of the researcher. The communication was done via the lecture and not the laboratory because more undergraduates would receive the information at one time, and much of the course communication is routinely done through the lecture instructor. The project description indicated that responding to the survey was voluntary and was not related to the course grade; additionally, undergraduates were assured that their instructor would not see their responses to the survey. Throughout the survey, undergraduates were asked to focus on generic and not specific instructors.

Prior to beginning the survey, undergraduate students responded to demographic questions about their gender, enrollment status, major, native language, and other biology courses they had completed. Student participants then answered the same set of survey questions but were assigned (by their last name for each lecture class) to only one of four scenarios: 1) imagine a small discussion class (20-25 students) taught by a professor, 2) imagine a small discussion class (20-25 students) taught by a graduate teaching assistant, 3) imagine a lab class (20-25 students) taught by a professor, or 4) imagine a lab class (20-25 students) taught by a graduate teaching assistant.

After respondents were informed of their scenario, they responded to items from the CUCEI. The original directions for the CUCEI were included, and students were told to respond to each item using a four point scale of “strongly disagree,” “disagree,”

“agree,” and “strong disagree”. The order the items were presented in was similar to the methodology in Coll et al. (2002) and Treagust and Fraser (1986) to be consistent with previous studies. After students completed the CUCEI items, they were presented with directions for the QTI. The response options were on a five point scale ranging from “0” (the instructor never displays this behavior) to “4” (the instructor always displays this behavior). The QTI items were presented in an order similar to that used by Coll et al. (2002). Students were unable to change their responses once they had completed a page. Undergraduates were also asked to confirm their scenario at the end of the survey to allow the researcher to remove responses from undergraduates who might have forgotten their scenario midsurvey. After responding to the CUCEI and QTI items, all undergraduate students then responded to the same open-ended question: If the same class was taught by a professor versus a graduate teaching assistant, how do you think the classes would be different?

No incentives were offered for participation, and all procedures were reviewed and approved by the Institutional Review Board for Human Subjects.

Survey Validity and Reliability

Because this project used two pre-existing instruments (both previously validated) in the same survey, with four different scenarios, internal consistency/reliability estimates were calculated for each subscale, as well as for each subscale for each scenario, using Cronbach’s alpha. For the overall data, these values ranged from 0.655 to 0.876, depending on the subscale, and were higher than those reported for previous uses of the instruments in college environments, except for the subscale of Individualization

(Treagust and Fraser, 1986;Coll et al., 2002). For the subscales by scenarios, the reliability measures varied from 0.580 to 0.911 for each subscale. From these results, it was judged that the instruments were reliable to use for the project.

Validity of the open-response question was judged via face validity, in that most students responded to the question and gave responses that were consistent with the intent of the question. No students wrote that they did not understand what the question was asking.

Data Analysis

Quantitative Data. Data from the CUCEI and QTI sub-scales were coded independently due to differences in the Likert scaling. The CUCEI responses were coded in the following manner: strongly agree = 5, agree = 4, disagree = 2, and strongly disagree = 1. Nonpositive items (see Table 1) were "flipped" and then coded: strongly agree = 1, agree = 2, disagree = 4, or strongly disagree = 5. The QTI was coded in the following manner: responses of "always" were scored as "4," and then coding progressed down to "never" which was scored as "0." No questions from the QTI were nonpositive Student responses for each scenario were then compiled for comparison.

The quantitative data were analyzed using nonparametric methodology because the Likert-type choices are ordinal. In ordinal data, there is no guarantee that students perceive the difference between intervals on the point scale as equal distances (e.g., "agree" and "strongly agree" are at the same distance as "agree" and "disagree"). This theoretical lack of equal distances violates assumptions for parametric methodology, requiring the data to be analyzed via nonparametric tests (Huck, 2008).

For each of the CUCEI and QTI items, student responses among each of the four scenarios were compared using Kruskal-Wallis one-way analysis of variance (ANOVA) tests (SPSS Statistics 19.0). Items in which statistically significant differences were found ($\alpha < 0.05$) were then compared using pairwise comparisons (Mann-Whitney U-tests, with Bonferroni approach to control for type I error across tests [$\alpha < 0.013$]). Pairwise comparisons were performed to determine whether the differences were due to the classroom setting or instructor type; the former were considered to be "classroom variables" and the latter were considered to be "instructor variables." Significant classroom variables were identified by comparing professors in a discussion class with professors in a lab class and GTAs in a discussion class to GTAs in a lab class, while instructor variables were identified by comparing professors in a discussion class with GTAs in a discussion class and professors in a lab class to GTAs in a lab class.

The majority of undergraduate participants were first-year students (53%); consequently, Kruskal-Wallis one way ANOVA tests were performed to determine whether student enrollment status affected the results. It was found that responses from first-year students did not differ from second-, third-, fourth-year and beyond students; therefore, undergraduate students were grouped together for all analyses.

***Qualitative Data*.** Open-ended responses (n = 127) to the question about differences between a class taught by a GTA versus a professor were first sorted into responses that indicated there would be no difference and those that stated there would be a difference. Undergraduate student responses in which differences were perceived (n = 110) then underwent thematic analysis using a "grounded theory" approach (Corbin and

Strauss, 1990; LeCompte, 2000) in which the researchers let the results emerge from the data without preconceived ideas about what students might articulate. Responses were analyzed independently by two researchers (each author of this paper) who read and reread the responses and took notes on the differences undergraduates articulated about the two different types of instructors. Factors that arose consistently were grouped and given a name (key word) and description, and then each researcher tallied the number of times those keywords or descriptions appeared in student responses. Each researcher then compiled her findings and only then did the researchers compare the results. The identified key words were then grouped into themes via discussion between the two researchers until both were in agreement.

As an example of this process, student responses yielded the key word "relatable" as one possible emergent characteristic of GTAs. Student responses that indicated that GTAs identified with them, related to them, knew what it was like to be in a student's shoes, were classified into this category. This key word of "relatable" was then merged with other keywords (such as "respect," "boring," and "approachable") into a theme that was entitled "relationship," because they all were thought to be articulating how students and instructors interact on a personal level (one-on-one, not just in a classroom delivery setting).

Reliability of these results was first obtained by the concordance of key words and descriptions between the two authors. Even if the key words were different, the ideas captured from the student responses were the same, and the researchers discussed the final title of the key word to reach consensus. The results also aligned with several of the

quantitative results of this study (e.g., uncertain and nervous), as well as the results of previous studies (e.g., approachable, relatable, uncertain, nervous, and limited control [Park, 2002; Dudley, 2009; Muzaka, 2009]). These multiple sources of verification of the results were considered evidence of the reliability of the findings.

Results

Participants

The survey was sent to a potential undergraduate pool of 2586 undergraduate students. From this pool, 387 began the survey (15.0%), while 225 undergraduates completed the survey (8.7%). Undergraduates who were minors or who could not remember their scenario at the end of the survey were removed, leaving 184 total respondents (7.1%). This response rate was likely a result of the survey being voluntary with no incentives for participation. Of the 184 respondents, survey completion for each scenario was: 59 for a discussion class taught by a professor; 43 for a discussion class taught by a GTA; 38 for a lab class taught by a professor; and 44 for a lab class taught by a GTA. There were 167 undergraduates who responded to the open-ended question. After removing responses in which the instructor they were referring to was indeterminable (e.g., “They are more understanding”), 127 respondents remained for analysis.

Individuals who completed the survey were mostly freshman (first year; 53%), non-biology majors (74%), female (72%), native English speakers (95%). Second- and third-year students comprised 21 and 18%, respectively, of the respondents, with 8% more being fourth year or beyond. Twenty-six percent of the students were biology majors, with 4% concentrating in Ecology and Evolutionary Biology, 10% in

Biochemistry & Cellular and Molecular Biology and 3% in Microbiology. Most of the respondents were currently enrolled in a majors' biology course (62%), while the rest were currently enrolled in non-majors courses. The majority of respondents had not completed another biology course (63%); however, 37% had completed at least one other semester-long lecture/lab biology course, such as first semester non-majors biology (11%), biodiversity (17%), and cell biology (13%). A complete summary of the demographics for overall survey respondents as well as demographics for each scenario is shown in Table 3.

Quantitative Analysis

***Item Analysis*.** Items from each survey in which significant differences in the medians among the scenarios were found (Kruskal-Wallis; $\alpha < 0.05$) are shown in Table 4. Descriptive statistics for each of these items for each scenario are shown in Table 5. These items included: students knowing what to do in class ($\chi^2 = 7.95$, df = 3, $n = 183$, p = 0.047), having a say in how class time is spent ($\chi^2 = 7.821$, df = 3, $n = 184$, p = 0.050), class being disorganized ($\chi^2 = 8.264$, df = 3, $n = 184$, p = 0.041), students being allowed to choose activities and how they will work ($\chi^2 = 9.004$, df = 3, $n = 181$, p = 0.029), students having opportunities to express their opinions in class ($\chi^2 = 10.976$, df = 3, $n = 184$, p = 0.012), activities being clearly and carefully planned ($\chi^2 = 13.09$, df = 3, $n = 184$, p = 0.004), the teacher talking enthusiastically about the subject ($\chi^2 = 24.098$, df = 3, $n = 184$, p = 0.000), the teacher being uncertain ($\chi^2 = 16.20$, df = 3, $n = 184$, p = 0.001), the teacher being hesitant ($\chi^2 = 10.628$, df = 3, $n = 183$, p = 0.014), and the teacher knowing what to do ($\chi^2 = 17.319$, df = 3, $n = 184$, p = 0.001). These significant

differences spanned items from five of the eight subscales of the two instruments, including Task Orientation, Individualization, Involvement, Leadership, and Uncertain.

***Pairwise comparison*s.** To determine whether the significant differences among scenarios were a result of classroom or instructor variables, pairwise comparisons were completed between classroom scenarios and then between instructor scenarios for each significant item. The comparisons among classroom scenarios (a GTA teaching a lab vs. teaching a discussion, or a professor teaching a lab vs. teaching a discussion) found no differences between any of the scenarios (Table 4). These results indicate that undergraduates in this study perceive instructor qualities to be the same in discussion classrooms as in lab classrooms.

For instructor variables (a GTA vs. a professor teaching the same type of class), there were no significant differences in responses between a professor teaching a discussion class and a GTA teaching a discussion class. However, undergraduates perceived differences in a professor teaching a lab class and a GTA teaching a lab class. In the situation of the professor and GTA both teaching a lab class, significant differences in undergraduate responses occurred in the factors of planning (U = 570.500, n = 81, p = 0.004), enthusiasm (U = 373.500, n = 81, p = 0.000), uncertainty (U = 553.500, n = 81, p = 0.005), hesitation (U = 518.000, n = 81, p = 0.002), and knowing what to do (U = 464.500, n = 81, p = 0.000) spanning items from three (Task Orientation, Leadership, and Uncertain) of the eight subscales from the two instruments used in this study(Table 4). Thus, when professors and GTAs are in lab classrooms, students in our study indicated that professors are more enthusiastic and have activities more clearly and

carefully planned, while GTAs are more uncertain, hesitant, and act as if they do not know what to do. These results are similar to those obtained from the qualitative data (see below), except the factor of enthusiasm, which was not mentioned by undergraduates in the open ended responses.

Qualitative Analysis

While ~13% (n = 17) of the undergraduate student respondents indicated that they would not perceive a difference in a class taught by a professor versus a GTA, analysis of the responses from undergraduates who did perceive differences (n = 110) generated two overall themes: factors relating to the teaching realm (further subdivided into themes of delivery technique and classroom atmosphere) and factors related to the personal realm (particularly in regard to relationship). The delivery technique, classroom atmosphere, and relationship themes are explained in Table 6 by using the key words and main descriptions researchers used to characterize undergraduate perceptions of the differences between GTAs and professors in the open-ended question. The numbers in parentheses after each bold descriptive key word indicate the number of times responses were grouped into those key words. A single student response could contain anywhere from zero to three key words (mean = 1.4 key words). These perceived differences between GTAs and professors were identified as instructor variables and not classroom variables since participants were asked about how GTAs and professors would differ in teaching the same type of class.

The theme of delivery technique includes student responses that seemed to reflect the characteristics of the instructors as teachers, specifically how they deliver the course

material to students and how confident the instructor appears in the classroom. This theme incorporated responses that referenced classroom control, organization and preparedness, knowledge level, and teaching experience. Undergraduates described professors as being more experienced, structured, confident, knowledgeable, organized, and in control in the classroom as compared with GTAs who undergraduates described as more hesitant, nervous, and uncertain. For instance, student 58 stated "TAs tend to be unorganized, without a strong curriculum to back them up, nor do they have the teaching experience that gives them the courage to stand before students and confidently teach the material." Meanwhile, student 153 elaborated on this by saying, "Professors are usually more confident and know the subject or lab better. GTAs are always a little uncertain and don't always answer questions well." The key words that comprised this theme were the most common responses by undergraduates to the open-ended question, especially in terms of the knowledge (n = 21), structure/organization (n= 17), and experience (n = 15) of professors being greater than that of GTAs.

The classroom atmosphere theme grouped student comments that seemed to be about student-instructor relationships within the context of classroom instruction. Undergraduates in the study articulated that the classroom atmosphere with a professor was more distant, strict, and formal as compared with that of a GTA whose classroom was seen as more relaxed, laid-back, and personalized. Furthermore, some undergraduates expressed that professor-led classrooms had higher expectations and were more serious than GTA-led classrooms, which some undergraduates indicated were more interactive and engaging. For instance, student 60 said "The professor might be more

strict rather than a graduate teaching assistant who might be a little more laid back because they can relate to the stress of college and how they might have had a bad teacher or hard professor." Student 156 explained, "Professors are typically old and out of touch with the students. They are set in their ways about their classes, and tend to lecture blandly. Graduate TAs seem to be more engaging." In this theme the most prevalent distinction made by undergraduates was the strictness (n = 13) and formality (n = 9) of professors in comparison with GTAs, who were perceived as more relaxed and laid-back (n = 9).

The relationship theme featured comments that seemed to be about person-to-person interactions, such as the comfort level and identifiability between students and instructor. For example, professors were perceived as being boring and out of touch, yet respected, while GTAs were seen as approachable, understanding, and relatable. Student 132 explained, "People may feel that they could approach a grad assistant, whereas they might be threatened by a professor." Student 163 commented, "I believe we would respect the Professor more than the Graduate Teaching Assistant. Just because the TA's are younger and look pretty much like our friends therefore making them look younger and less serious." In terms of the relationship aspect, undergraduates noted their respect for professors (n = 14) and their ability to relate to GTAs (n = 19) as the primary descriptors.

Discussion

Bos et al. (1980) suggested caution when comparing different studies regarding GTA instructional abilities with ones for ranked faculty members; our study has

addressed this concern through comparison of GTAs and professors in the same study using the same sample population and instruments. This study has further addressed the call for developing targeted professional development (Shannon et al., 1998; Baldwin and Wawrzynski, 2011) by identifying aspects that distinguish GTAs from faculty and that could be the focus of professional development.

We hypothesized that undergraduates would perceive differences between GTAs and professors that were independent of classroom variables, and found evidence that supported this hypothesis. Undergraduates in this study perceived professors as being more structured, confident, in control, organized, experienced, knowledgeable, distant, formal, strict, serious, hard, boring, out of touch, and respected than GTAs. Conversely, GTAs were perceived as more uncertain, hesitant, nervous, relaxed, laid-back, engaging, interactive, relatable, understanding, and able to personalize teaching than professors.

Significant differences between instructor type were found for five of the 52 items surveyed in the quantitative portion of this study, while no significant classroom variables were identified. Thus, although students do see many similarities between professors and GTAs, there are also core factors that undergraduates at one institution say are different between GTAs and professors independent of what classes they teach. These results were supported by the qualitative data analysis, which independently confirmed four of the five instructor differences from the quantitative analysis. In addition, the open-response format allowed students to add additional variables that were not identified from the survey.

Although it remains to be seen whether the results of this study are broadly applicable (see *Limitations*), they can be used as the starting point for investigating classroom practice and pondering the stereotypes that students may carry with them into classrooms. For instance, undergraduates appear to have positive feelings about how professors organize and understand the content they teach, and respect them overall, but they also appear to have more negative views of their abilities to relate to undergraduates and understand them. Conversely, undergraduates have some negative perceptions of GTAs' abilities to convey information instructionally, but they have very positive feelings about GTAs' abilities to interact with them. It is important to point out that these results do not seem to indicate that GTAs are less favored as instructors compared with professors; however, the stereotypes articulated by undergraduates almost surely impact the teaching and learning environment in undergraduate classrooms. Even if an individual GTA or professor does not adhere to the identified stereotypes, they are likely being compared to this typical perception which may impact how students react to them in the classroom.

Our results are similar to those obtained by Park (2002) and Muzaka (2009) in which GTAs, staff, and undergraduates indicated that GTAs were understanding, approachable, laid-back and nervous; have limited control and authority; and lack content knowledge and experience. Yet, Park (2002) and Muzaka (2009) found that GTAs exhibit a youthful enthusiasm, while undergraduates from our study indicated that GTAs are less enthusiastic than professors. This may be a result of this study asking students to directly compare GTAs with professors, or because these data were collected solely from

undergraduates. These differences may also be attributable to culture, since the work done by Park (2002) and Muzaka (2009) was completed in the United Kingdom while our study took place in the United States. Our study also focused specifically on introductory courses in which laboratories function somewhat separately from lecture (taught by different instructors in different locations with different class sizes) whereas other institutions may have a different course structure. There are also different selection processes for GTAs at different universities, with some GTAs volunteering or being more enthusiastic about teaching while GTAs at a research university may not be as enthusiastic or have the time or encouragement to embrace their teaching duties.

It is also now possible to compare the perceptions of undergraduates in this study with reflections of GTAs on their own teaching characteristics. Dudley (2009) documented instructor variables of GTAs from the perspective of GTAs. These GTAs expressed ideas such as: confusion with expectations, difficulty establishing boundaries, dealing with nervousness, expectation to be knowledgeable, and being uncertain where to start. The variables of nervousness and being uncertain about where to start were also identified as characteristics of GTAs by undergraduates in our study. The variables of confusion with expectations, difficulty establishing boundaries, and expectation to be knowledgeable identified by Dudley (2009) are similar to the lack of experience, control, and knowledge that undergraduates identified with GTAs in this study. These similarities in the perceptions of GTAs by GTAs themselves and undergraduates may be explained by the fact that GTAs see themselves as students who merely have teaching responsibilities (Park, 2002; Muzaka, 2009).

Proposed Explanations of Instructor Differences

The aspects that undergraduates perceive as different between GTAs and professors may have their origin in factors that are specific to the academic context, such as who has control over the curriculum and the status of the instructor (faculty or not faculty). When using these factors to explain the themes (delivery technique, classroom atmosphere and relationship) it should be acknowledged that one factor may influence several aspects of instructor perception. Some of these factors will be highlighted below by providing literature that explains differences between student perceptions of GTAs and professors for each of the themes identified in this study.

One factor that may contribute to undergraduate perception of the instructor's delivery technique is the curriculum. Professors typically have more control over the organization of the curriculum and classroom policies, while GTAs are typically given specific assignments within the curriculum to enact with oftentimes little opportunity to alter or modify it. This was articulated in a study by Park and Ramos (2002) in which GTAs expressed they had little autonomy or ownership over what they taught, but were merely "carrying out the job," and a study by Muzaka (2009) in which GTA lack of control and authority over the curriculum led to a perception that students "see no point to us." The perception of "control", however, could also be due to professors having a greater student-teacher distance due to status, age, and possibly greater confidence in the subject matter (Roach, 1997). For undergraduates in our study, these factors may have contributed to their feelings of their GTAs being hesitant and uncertain about what they were teaching and their professors being confident, knowledgeable and organized.

The classroom atmosphere theme appears to be influenced by the perception of instructor behavior in the classroom. For instance, GTAs are perceived as being engaging and laid-back while professors are more distant and formal. Even though the classroom itself was not perceived by students as a distinguishing variable, it could be that the size of the classroom affects classroom teaching behavior. GTAs often teach smaller sections (laboratories and discussion sections) in which students are able to interact with them on a one-on-one basis, while professors often have larger classes in which it is more difficult to interact with students individually (Dudley, 2009). Professors may also appear more distant due to knowledge level and age differences, which may limit the interactions they have with students (Anderson and Carta-Falsa, 2002). Undergraduates may also perceive that since GTAs are typically similar in age to them, they can better relate to their classroom experiences and explain things to them (Muzaka, 2009).

Certainly, this perception of engagement is a positive aspect of GTA instruction. Darby (2005) concluded that when GTAs are enthusiastic about the subject matter, students are more comfortable with the subject matter and their learning is better supported. Similarly, O'Neal et al. (2007) determined that GTA enthusiasm positively impacts student retention in the sciences. This may be attributed to engaged students being more likely to learn and retain knowledge (Umbach and Wawrzynski, 2005).

For the theme of relationship, instructor age may be a factor influencing student perception. GTAs are typically younger than faculty members, and thus could be perceived by the undergraduates as closer to their own age and therefore more approachable (Muzaka, 2009). GTAs and undergraduates also have similar experiences

because GTAs are often still taking graduate classes while they are teaching (Park and Ramos, 2002; Dudley, 2009). Thus, undergraduates may think that since GTAs are more familiar with academic demands, the pressures of deadlines, and workload in their own courses and research, they may be more understanding, approachable, and relatable when undergraduates express workload issues (Muzaka, 2009). Greater age, status and confidence (Roach, 1997) may also be why undergraduates in this study afforded professors more respect than GTAs.

Limitations

One limitation of this study is the small sample size and the fact that it was conducted on a limited subset of volunteer participants in one discipline at one university. The sample was also greatly overrepresented by first year female non-biology major undergraduates. This particular sample may have had limited experience with instructors of various titles (GTA vs. professor) and the respondents therefore may have been thinking of the one GTA/professor they had in college when responding. This study is also limited in that it was performed at a large southern research university with its own culture of instruction and curriculum that may not be present at the vast majority of other schools. Further, student demographics for each of the quantitative scenarios were not identical. As with many single institution studies, the results cannot be generalized to all academic institutions; however, they can and should be used as the basis for additional investigations of GTA and professor instructional characteristics.

Recommendations for GTA Professional Development

The findings of this study suggest that undergraduates may have different perceptions of GTAs and professors in regard to several important instructional aspects. While additional investigations are certainly needed into why these factors are perceived differently by undergraduates, and whether these hold true at different institutions, these factors can be used in the meantime to help shape GTA professional development.

GTAs could be made aware of the results so they can better understand how undergraduates perceive them and come to know that most undergraduates do not understand the academic context of graduate students. GTA professional development should focus on keeping the positive aspects of GTA instructors (such as relatable, engaging, and approachable) while finding ways to decrease the perceived nervousness, uncertainty and hesitancy of GTAs. For instance, universities can work to better prepare GTAs for the specific curriculum content they will be teaching. Marincovich, et al. (1998) noted that GTA assigments are often made just prior to the semester, and suggested that if these assignments were known earlier GTAs might be able to better prepare for the courses they will be teaching. Faculty often know their course assignments months in advance and invest considerable amounts of time preparing course materials and syllabi; GTAs should be given the same advance preparation time if possible.

Shannon et al. (1998) and Luft et al. (2004) stated that GTAs should know more about curriculum delivery in general, and that providing professional development sessions in which GTAs are given background to deliver specific curricula may make

them feel more confident and knowledgeable in the classroom. In addition to training in delivering specific curricula, GTAs should be encouraged to take time to reflect on their teaching for the purposes of self-evaluation; this will allow them to more quickly identify their personal teaching style which should better promote student learning (Schussler et al., 2008).

GTAs could also be coached on behaviors that help them strike a balance between informal and relaxed and being strict and having high standards, because the former may be beneficial in terms of engaging students, but it can also be problematic in terms of authority and respect (Muzaka, 2009). Roach (1997) suggested that attire impacts student perception of the instructor, instructors dressed in professional attire may be afforded more classroom control and authority. Thus, GTAs can be informed of the importance of proper attire when instructing students.

This study identified positive and negative aspects of both GTA and professor instruction from the perspective of undergraduates. Often, where one instructor was weaker in an aspect, the other was strong. This suggests that another mechanism of professional development would be giving more opportunities for GTAs and professors to teach collaboratively, which could help each gain in aspects that are perceived as weaknesses by undergraduates. For instance, a professor working with a GTA on a course could help the GTA develop a better understanding of how to be more confident and organized about the curriculum, and how to set higher standards. To achieve this, lecture and laboratory courses could be cotaught by professors and GTAs, versus

separating those duties, or in cases in which this is not possible, GTAs could help coordinate the curriculum for a course they have been assigned to teach.

These recommendations for GTA professional development focus on maintaining as many of the positive aspects of GTA teaching as possible, while simultaneously finding ways to decrease the negative perceptions expressed by undergraduates. The overall goal of these new programs would be to increase GTA confidence in their teaching assignments and undergraduate perception of their teaching abilities, which should result in a better learning environment for everyone.

Conclusions and Future Directions

There were several factors that undergraduates in this study perceived as being different between professors and GTAs; this study documented these differences and made suggestions for how to use this information to potentially improve teaching and learning at universities. Additional research, however, could further refine these results. For instance, exploration into why undergraduates perceive these differences may clarify why first year students in this study had similar viewpoints compared with upper level students with more exposure to different instructor types. Other research could study actual GTAs and professors to see whether the differences identified in this study hold true in practice, or at other institutions. It would also be interesting to study undergraduate student perception of instructors who deviate from stereotypical instructor variables and how this impacts student perception of the instructor. For example, a professor who is nervous or a GTA who is unapproachable may be perceived more negatively, because they do not adhere to student expectations for that type of instructor.

Future studies should also clarify the terms undergraduate students used in the study, for example, what exactly students mean when they say “strict” or “uncertain.” Careful studies such as these will shed additional light on the complicated instructional relationships among undergraduates, graduate teaching assistants, and professors which may help each group to better understand how to maximize teaching and learning in undergraduate courses in the future.

Literature Cited

Anderson LE, Carta-Falsa J (2002). Factors that make faculty and students relationships effective. Coll Teach *50*, 134–138.

Austin AE (2002). Preparing the next generation of faculty: graduate school as socialization to the academic career. J High Educ *73*, 94–122.

Baldwin RG, Wawrzynski MR (2011). Contingent faculty as teachers: what we know; what we need to know. Am Behav Sci *55*, 1485–1509.

Benjamin E (2002).How over reliance upon contingent appointments diminishes faculty involvement in student learning. Peer Review *5*, 4–10.

Bolge RD (1995). Examination of Student Learning as a Function of Instructor Status (Full-Time vs. Part-Time) at Mercer County Community College.WestWindsor, NJ: Mercer County Community College.

Bond-Robinson J, Rodriques RAB (2006). Catalyzing graduate teaching assistants’ laboratory teaching through design. Res J Chem Educ *83*, 313–323.

Bos RR, Zakrajsek DB,Wolf V, Stoll S (1980). Teaching assistant traits: their influence on student ratings. Improving Coll Univ Teach *28*, 179–185.

Braxton JM, Lambert LM, Clark SC (1995). Anticipatory socialization of undergraduate college teaching norms by entering graduate teaching assistants. Res High Educ *36*, 671–686.

Coll RK, Taylor N, Fisher DL (2002). An application of the Questionnaire on Teacher Interaction and college and university classroom environment inventory in a multicultural tertiary context. Res Sci Technol Educ *20*, 165–183.

Corbin J, Strauss A (1990). Grounded theory research: procedures, canons, and evaluative criteria. Qual Sociol *13*, 3–21.

Darby L (2005). Science students' perceptions of engaging pedagogy. Res Sci Educ *35*, 425–445.

Dudley M (2009). Jumping out of an airplane: a TA's perspective on teaching effectiveness. Eastern Educ J *38*, 1–10.

Flora BH(2007).Graduate assistants: students or staff, policy or practice? The current legal employment status of graduate assistants. J High Educ Pol Manag *29*, 315–322.

Golde CM (1998). Beginning graduate school: explaining first-year doctoral attrition. In: The Experience of Being in Graduate School: An Exploration, New Directions for Higher Education, ed. M.S. Anderson, San Francisco: Jossey-Bass.

GoldeCM, DoreTM(2001).At Cross Purposes: What the Experiences of Doctoral Students Reveal about Doctoral Education, Philadelphia, PA: Pew Charitable Trusts.

Huck S (2008). Reading Statistics and Research, Boston: Pearson Education.

Jaeger AJ (2008). Contingent faculty and student outcomes. Academe *94*, 42–43.

Johnson IY (2011). Contingent instructors and student outcomes: an artifact or fact? Res High Educ *52*, 761–785.

LeCompte MD (2000). Analyzing qualitative data. Theory Pract *39*, 146–154.

Lowman J, Mathie V (1993). What should graduate teaching assistants know about teaching? Teach Psychol *20*, 84–88.

Luft JA, Kurdziel JP, Roehrig GH, Turner J (2004). Growing a garden without water: graduate teaching assistants in introductory science laboratories at a doctoral/research university. J Res Sci Teach *41*, 211–233.

Marincovich M, Prostko J, Stout F (eds.) (1998). The Professional Development of Graduate Teaching Assistants, Balton, MA: Anker Publishing.

Moos R (1979). Evaluating Educational Environments, San Francisco: Jossey-Bass.

Muzaka V (2009). The niche of graduate teaching assistant (GTAs): perceptions and reflections. Teach High Educ *14*, 1–12.

Nyquist LM, Wulff DH, Austin AE, Sprague J, Fraser PK, Calcagno C, Woodford B (1999). On the road to becoming a professor: the graduate student experience. Change *31*, 18–27.

O'Neal C, Wright M, Cook C, Perorazio T, Purkiss J (2007). The impact of teaching assistants on student retention in the sciences: lessons for TA training. J Coll Sci Teach *36*, 24–29.

Park C (2002). Neither fish nor fowl? The perceived benefits and problems of using graduate teaching assistants (GTAs) to teach undergraduate students. High Educ Rev *35*, 50–62.

Park C, Ramos M (2002). The donkey in the department? Insights into the graduate teaching assistant (GTA) experience in the UK. J Grad Educ *3*, 47–53.

Roach K(1997).Effects of graduate teaching assistant attire on student learning, misbehaviors, and ratings of instruction. Commun Quart *45*, 125–141.

Rushin JW, Saix JD, Lumsden A, Streubel DP, Summers G, Bernson C (1997). Graduate teaching assistant training: a basis for improvement of college biology teaching and faculty development? Am Biol Teach *59*, 86–90.

Schussler E, Torres LE, Rybczynski S, Gerald GW, Monroe E, Sarkar P, Shahi D, Osman MA (2008). Transforming the teaching of science graduate students through reflection. J Coll Sci Teach *38*, 32–36.

Shannon DM, Twale DJ, Moore MS (1998). TA teaching effectiveness: the impact of training and teaching experience. J High Educ *69*, 440– 466.

Sundberg MD, Armstrong JE, Wischusen EW (2005). A reappraisal of the status of introductory biology laboratory education in U.S. colleges and universities. Am Biol Teach *67*, 525–529.

Treagust DF, Fraser BJ (1986). Validation and application of the college and university classroom environment inventory (CUCEI). Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA, April 16–20.

Umbach PD (2007). How effective are they? Exploring the impact of contingent faculty on undergraduate education. Rev High Educ *30*, 91–123.

Umbach PD, Wawrzynski MR (2005). Faculty do matter: the role of college faculty in student learning and engagement. Res High Educ *46*, 153–184.

Appendix

Table 1. The CUCEI subscales and items used for the study[a]

Sub-scale	Items
Individualization:	All students in the class are expected to do the same work, in the same way and in the same time.
	It is the instructor who decides what will be done in our class.
	Students are allowed to choose activities and how they will work.
	Students are generally allowed to work at their own pace.
	Students have a say in how class time is spent.
	Teaching approaches allow students to proceed at their own pace.
	There is little opportunity for a student to pursue his/her particular interest in this class.
Involvement:	Students "clockwatch" in this class.
	Students in this class pay attention to what others are saying.
	Students put effort into what they do in classes.
	Students seldom present their work to the class.
	The instructor dominates class discussion.
	The instructor talks rather than listens.
	There are opportunities for students to express opinions in this class.
Personalization:	The instructor considers students' feelings.
	The instructor goes out of his/her way to help students.
	The instructor helps each student who is having trouble with the work.
	The instructor is unfriendly and inconsiderate towards students.
	The instructor isn't interested in students' problems.
	The instructor seldom moves around the classroom to talk with students.
	The instructor talks individually with students.
Task Orientation:	Class assignments are clear so everyone knows what to do.
	Getting a certain amount of work done is important in this class.
	Students know exactly what has to be done in our class.
	The group often gets sidetracked instead of sticking to the point.
	This class seldom starts on time.
	This is a disorganized class.
	Activities in this class are clearly and carefully planned.

[a] Undergraduates rated whether they "strongly agree" (5), "agree" (4), "disagree" (2), or "strongly disagree" (1). The items include both positive and non-positive items (underlined) about the class.

Table 2. QTI subscales and items used in the study[a]

Sub-scale	Items
Helpful/Friendly:	This teacher can take a joke.
	This teacher has a sense of humor.
	This teacher helps us with our work.
	This teacher is friendly.
	This teacher is someone we can depend on.
	This teacher's class is pleasant.
Leadership:	This teacher acts confidently.
	This teacher explains things clearly.
	This teacher holds our attention.
	This teacher is a good leader.
	This teacher knows everything that goes on in the classroom.
	This teacher talks enthusiastically about her/his subject.
Strict:	This teacher is severe when marking papers.
	This teacher is strict.
	This teacher's standards are very high.
	This teacher's tests are hard.
	We are afraid of this teacher.
	We have to be silent in this teacher's class.
Uncertain:	It's easy to be off task with this teacher.
	This teacher acts as if she/he does not know what to do.
	This teacher is hesitant.
	This teacher is not sure what to do when we are not on task.
	This teacher lets us boss her/him around.
	This teacher seems uncertain.

[a] Undergraduates rated how often they believed an instructor would do this: 4 being "always," down to 0, which was "never."

Table 3. Demographic percentages for overall study participants, as well as by scenario[a]

	Overall	Discussion Professor	Discussion GTA	Lab Professor	Lab GTA
What is your gender?					
Male	28%	29%	30%	29%	23%
Female	72%	71%	70%	71%	77%
What is your current enrollment status?					
First year	53%	58%	49%	39%	61%
Second year	21%	22%	28%	26%	9%
Third year	18%	15%	14%	21%	23%
Fourth year	6%	3%	5%	11%	7%
Fifth year and/or beyond	2%	2%	2%	3%	0%
What is your major?					
Biology	9%	7%	12%	3%	14%
Ecology and Evolutionary Biology	4%	2%	2%	8%	5%
Biochemistry, Cellular, and Molecular Biology	10%	10%	5%	13%	11%
Microbiology	3%	0%	2%	3%	7%
Other	74%	81%	79%	74%	64%
Is English your native language?					
Yes	95%	93%	98%	100%	98%
No	4%	7%	2%	0%	2%
What other Core courses have you completed?					
First semester Non-majors Biology	11%	15%	5%	13%	11%
Second semester Non-majors Biology	8%	8%	7%	11%	7%
First semester Plant Biology	2%	2%	2%	5%	0%
Second semester Plant Biology	2%	2%	2%	3%	0%
Biodiversity	17%	14%	14%	24%	23%
Honors Biodiversity	1%	3%	0%	0%	0%
Cell Biology	13%	7%	12%	18%	18%
General Genetics	6%	3%	7%	8%	7%
General Ecology	1%	0%	0%	3%	2%
No others	63%	66%	69%	50%	66%
Current course enrollment					
First semester Non-majors Biology	38%	49%	38%	29%	30%
First semester Plant Biology	4%	2%	10%	3%	2%
Biodiversity	34%	31%	38%	29%	41%
Honors Biodiversity	3%	5%	0%	3%	5%
Cell Biology	8%	7%	5%	18%	5%
General Genetics	7%	5%	7%	3%	11%
General Ecology	7%	2%	5%	16%	7%

[a] Overall n = 184; Discussion professor n = 59, Discussion GTA n = 43; Lab professor n = 38; Lab GTA n = 44.

Table 4. Significant results for the Kruskal-Wallis ($\alpha < 0.05$) and Mann-Whitney U ($\alpha < 0.013$ with the Bonferroni correction) nonparametric tests[a]

Item	Kruskal-Wallis one-way ANOVA		Mann-Whitney U, p-value			
			Classroom Variables		Instructor Variables	
	χ^2	P-value	DP vs LP	DG vs LG	DP vs DG	LP vs LG
Students are allowed to choose activities and how they will work. (CUCEI)	9.004	0.029	0.076	0.067	0.462	0.325
Students know exactly what has to be done in our class. (CUCEI)	7.95	0.047	0.920	0.663	0.085	0.034*
Students have a say in how class time is spent. (CUCEI)	7.821	0.050	0.100	0.077	0.574	0.366
This is a disorganized class. (CUCEI)	8.264	0.041	0.717	0.301	0.150	0.026*
There are opportunities for students to express their opinions in this class. (CUCEI)	10.976	0.012	0.166	0.039*	0.173	0.091
Activities in this class are clearly and carefully planned. (CUCEI)	13.09	0.004	0.696	0.344	0.052	**0.004** (P)**
This teacher is hesitant. (QTI)	10.628	0.014	0.091	0.679	0.302	**0.002** (G)**
This teacher talks enthusiastically about her/his subject. (QTI)	24.098	0.000	0.118	0.017*	0.338	**0.000** (P)**
This teacher seems uncertain. (QTI)	16.206	0.001	0.359	0.166	0.032*	**0.005** (G)**
This teacher acts as if she/he does not know what to do. (QTI)	17.319	0.001	0.103	0.554	0.086	**0.000** (G)**

[a] Significant items from the CUCEI (6) and QTI (4) shown in the order they were presented to students. Significant instructor variables are bolded in the bottom right-hand corner. Letters in parentheses indicate who had the higher values for each item. D = Discussion, L = Lab, P = Professor, G = GTA. *Mann-Whitney U, p value significant before Bonferroni correction ($\alpha < 0.05$); **Mann-Whitney U, p value significant with Bonferroni correction ($\alpha < 0.013$).

Table 5. Descriptive statistics for significant items for each scenario, including mean ± standard deviation, median, maximum, and minimum.

	Professor Discussion				GTA Discussion				Professor Lab				GTA Lab			
	Mean ± Standard deviation	**Med**	**Max**	**Min**	**Mean ± Standard deviation**	**Med**	**Max**	**Min**	**Mean ± Standard deviation**	**Med**	**Max**	**Min**	**Mean ± Standard deviation**	**Med**	**Max**	**Min**
Students know exactly what has to be done in our class.	3.19 ± 1.03	4	5	1	3.47 ± 1.26	4	5	1	3.89 ± 1.16	4	5	1	3.39 ± 1.17	4	5	1
Students have a say in how class time is spent.	3.03 ± 1.17	4	5	1	2.91 ± 1.29	2	5	1	2.63 ± 1.17	2	5	1	2.41 ± 1.11	2	5	1
This is a disorganized class.	4.12 ± 1.02	4	5	1	3.84 ± 1.11	4	5	1	4.11 ± 0.95	4	5	1	3.68 ± 1.05	4	5	1
Students are allowed to choose activities and how they will work.	2.88 ± 1.10	2	5	1	2.72 ± 1.20	2	5	1	2.47 ± 1.03	2	4	1	2.23 ± 0.84	2	5	1
There are opportunities for students to express opinions in this class.	4.00 ± 1.07	4	5	1	3.86 ± 0.91	4	5	1	3.79 ± 1.02	4	5	1	3.48 ± 1.00	4	5	1
Activities in this class are clearly and carefully planned.	3.90 ± 0.84	4	5	2	3.51 ± 1.05	4	5	1	3.95 ± 0.87	4	5	2	3.30 ± 1.13	4	5	1
This teacher talks enthusiastically about her/his subject.	3.27 ± 0.83	3	4	0	3.09 ± 0.92	3	4	1	3.53 ± 0.65	4	4	2	2.64 ± 0.92	3	4	0
This teacher seems uncertain.	0.81 ± 0.94	1	4	0	1.19 ± 0.98	1	3	0	0.92 ± 0.88	1	4	0	1.45 ± 0.95	1	4	0
This teacher is hesitant.	1.02 ± 0.99	1	4	0	1.19 ± 0.96	1	3	0	0.66 ± 0.63	1	3	0	1.26 ± 0.93	1	3	0
This teacher acts as if she/he does not know what to do.	0.61 ± 1.02	0	4	0	0.98 ± 1.16	1	4	0	0.34 ± 0.78	0	3	0	1.02 ± 1.00	1	4	0

Table 6. Summary of themes (delivery technique, classroom atmosphere, and relationship) obtained from qualitative data in both the teaching and personal realms for GTAs and professors[a]

Theme	GTA	Professor
Delivery Technique	H**esitant**, **nervous**, **uncertain** (11), and unsure how to begin teaching.	**Organized** and **structured** (17), **confident** (10), in **control** (10), prepared for questions, with previous teaching **experience** (15), and greater **knowledge** (21).
Classroom Atmosphere	**Relaxed** and **laid back** (9), **interactive**, **engaging** (5), personalized, and having open **student-instructor interactions** (3).	**Distant** and **formal** (9), **strict** (13), serious, **harder** (2), with higher expectations and standards.
Relationship	**Comfortable approaching** GTAs (8) and that GTAs are **relatable** (19) and **understanding** (5).	Intimidating and **boring** (2), and **out of touch** (6), yet undergraduates **respect** (14) professors.

[a] Numbers in parentheses are how often the bolded key words emerged from the data set.

CHAPTER TWO

Evolving impressions: Undergraduate perceptions of graduate teaching assistants and faculty members over a semester

The following chapter is in press:

Kendall KD, Schussler EE (2013). Evolving impressions: Undergraduate perceptions of graduate teaching assistants and faculty members over a semester. CBE-Life Sciences Educ *12*.

Keywords: Teaching, University, Evaluations, Professional Development, Graduate Teaching Assistants, Faculty Members

Abstract

Undergraduate experiences in lower-division science courses are important factors in student retention in science majors. These courses often include a lecture taught by faculty, supplemented by smaller sections, such as discussions and laboratories, taught by graduate teaching assistants (GTAs). Given that portions of these courses are taught by different instructor types, this study explored student ratings of instruction by GTAs and faculty members to see whether perceptions differed by instructor type, whether they changed over a semester, and whether certain instructor traits were associated with student perceptions of their instructors' teaching effectiveness or how much students learned from their instructors. Students rated their faculty instructors and GTAs for 13 instructor descriptors at the beginning and near the end of the semester in eight biology classes. Analyses of these data identified differences between instructor types; moreover, student perception changed over the semester. Specifically, GTAs ratings increased in perception of positive instructional descriptors while faculty ratings declined for positive instructional descriptors. The relationship of these perception changes with student experience and retention should be further explored, but the findings also suggest the need to differentiate professional development by the different instructor types teaching lower-division science courses to optimize teaching effectiveness and student learning in these important gateway courses.

Introduction

Increased employment of, and reliance on, contingent instructors is a new reality at many institutions of higher education. The term "contingent instructor" typically refers to part-time, non-tenure track faculty, or graduate teaching assistants (GTAs; Johnson, 2011). Contingent instructors may constitute nearly half of the instructors used by undergraduate institutions (Jaeger, 2008; Baldwin and Wawrzynski, 2011), revealing the great dependence higher education can have on contingent instructor employment.

Despite this increased reliance on contingent instructors, research regarding their impact on teaching and student learning is relatively sparse and inconclusive (Benjamin, 2002, Umbach, 2007). One study documented that contingent instructors typically give higher grades (Johnson, 2011), yet another found that contingent instructors do not differentially impact final student grades (Bolge, 1995). It has been documented that out-of-class interactions are beneficial for student learning, despite these benefits contingent instructors being generally less available to students than their noncontingent colleagues (Benjamin, 2002; Jaeger, 2008), correspondingly students are less likely to complete their degree when taught by contingent instructors (Jaeger, 2008).

The issue of degree completion is important because science departments in higher education typically suffer high attrition rates from the majors (Seymour and Hewitt, 1997; President's Council of Advisors on Science and Technology, 2012). Factors such as student loss of interest in the major, course demands, instructors, and inadequate educational background and preparation of the student have been cited as impacting student attrition (Seymour and Hewitt, 1997). Seymour and Hewitt (1997) also documented attrition effects that were specific to instructors, including poor teaching, lack of approachability, lack of availability for help,

language barriers, and the high burden of instruction placed on TAs. However, studies directly exploring instructor impact on student retention are rare, and ones investigating differential effects of faculty members and GTAs on retention even rarer. Johnson (2011) found that instructor type does not differentially impact student retention, yet O'Neal et al. (2007) documented that GTAs can positively impact student retention by creating a positive laboratory atmosphere. The lack of research on this topic indicates that academia still does not fully understand the impact of instructor type on student learning (Jaeger, 2008), which lends credence to calls for further research.

Also in need of research are strategies for targeted professional development for different instructor types (Baldwin and Wawrzynski, 2011), and this study specifically addresses this need for one subgroup of contingent instructors, GTAs. A survey of 4114 graduate students by Golde and Dore (2001) found that many graduate programs (53.6%) require teaching by their graduate students. A survey in the biological sciences discipline found that 71% of laboratories at comprehensive universities (n = 15), and 91% of laboratories at research universities (n = 34) were taught by GTAs (Sundberg et al., 2005). Similarly, Rushin et al. (1997) surveyed 153 graduate school programs and found that 97% of these programs reported using GTAs to teach laboratories and/or lecture class sections. While the dependence may vary from institution to institution, there is substantial reliance in academia on GTAs instructing undergraduates (Rushin et al., 1997; Sundberg et al., 2005).

GTA Use: Benefits and Concerns

Several studies have explored perceptions of GTAs, particularly with regard to GTA preparation for, and quality of, instruction (Park, 2002; Dudley, 2009; Kendall and Schussler, 2012; Muzaka, 2009). These studies have captured the perceptions of GTA instruction from the

viewpoints of undergraduate students, GTAs, and professors from both the United States and the United Kingdom, providing multiple perspectives within higher education systems about the impact of GTA instruction on the educational environment and student learning. Throughout this article, we use the term “descriptor” (or its derivatives) to refer to words that are often used to describe perceptions of instructors (e.g., “boring,” “engaging,” or “organized”). These words have also been referred to as instructor characteristics, traits, or qualities in other studies.

In studies of student perception of instructors, GTAs have been identified with positive descriptors of being engaging, approachable, informal, relaxed, interactive, relatable, understanding, and able to personalize teaching (Park, 2002; Dudley, 2009; Muzaka, 2009; Kendall and Schussler, 2012). Yet, they have also been identified with negative descriptors including lacking confidence, knowledge, experience, and authority (Park, 2002; Dudley, 2009; Muzaka, 2009; Kendall and Schussler, 2012). Meanwhile, professors have been described with positive descriptors such as respected and organized, but negative descriptors such as strict, boring, and distant (Kendall and Schussler, 2012). Interestingly, these studies have reported contradictory results for enthusiasm; Kendall and Schussler (2012) found that professors are considered more enthusiastic, while Park (2002) and Muzaka (2009) found that GTAs are thought of as more enthusiastic.

The existence of negative descriptors for GTAs has raised concerns about using GTAs as instructors at universities. Specifically, undergraduate students report being concerned about a lack of overall subject knowledge demonstrated by their GTAs, limited communication between GTAs and professors teaching the same course, variability of standards used by GTAs to assess undergraduates, and GTAs appearing nervous and not confident (Park, 2002; Muzaka, 2009). GTAs themselves have expressed concern about contradicting professors they teach for, having

minimal teaching training, and lacking authority in the classroom (Park, 2002; Dudley, 2009; Muzaka, 2009). Similarly, professors have expressed their own concerns about GTAs' minimal instructional training and lack of confidence negatively affecting GTAs' ability to foster student learning (Park, 2002; Muzaka, 2009). Professors also noted the inability to guarantee consistency in GTA teaching and the possibility that GTAs limit undergraduate student access to "real academics" as other concerns about GTAs as instructors (Park, 2002; Muzaka, 2009).

Students' Evaluations of Teaching

Instructor teaching effectiveness is often evaluated by institutions for the purposes of instructor retention, but these evaluations also provide feedback to instructors who wish to modify their teaching to foster student learning. A key strength of these evaluations is that the feedback comes from students directly, allowing students to express their perceptions of their instructors and their instructors' ability to foster learning (Baird, 1987; Emery et al., 2003; Clayson et al., 2006; Zabaleta, 2007; Helterbran, 2008; Kogan et al., 2010). However, some question whether factors such as the grades students receive from their instructors, instructor age, or instructor gender have a stronger influence on the ratings instructors receive than actual instructional effectiveness (d'Apollonia and Abrami, 1997; Greenwald and Gillmore, 1997; Marsh and Roche, 1997; Emery et al., 2003; Clayson and Haley, 2011).

Studies exploring these concerns have provided some insight, for instance, Clayson et al. (2006) found that students systematically changed evaluations based on the grades they expected. Likewise, Baird (1987) and Greenwald and Gillmore (1997) noted a positive relationship between student evaluation ratings and course grades received. Research investigating instructor age conducted by Bos et al. (1980) noted that students rated teaching assistants in their later twenties superior to teaching assistants in their early twenties. Zabaleta

(2007) found similar age influences on student evaluations; however, Zabaleta offered the caution that age is inferred by the rater and therefore may be an erroneous factor. Inconclusive results have been found regarding the influence of instructor gender; for example, Zabaleta (2007) and Bos et al. (1980) found no significant role of instructor gender on evaluation ratings. In contrast, Feldman's (1993) meta-analysis revealed that students favor female instructors, and Basow (1995) and Centra and Gaubatz (2000) reported differences between male and female instructor ratings.

Project Rationale

Previous work has found that undergraduates have different perceptions of GTAs and faculty members with regards to several instructional descriptors, some potentially more beneficial to learning than others. Kendall and Schussler (2012) documented this in a study on perceived differences between hypothetical professors and GTAs. The current study compares student ratings of their biology instructors (faculty members and GTAs) at the same institution to test whether the differences in instructor descriptors hold true when students are asked to rate their current, non-hypothetical, instructors. The faculty members recruited for this study were all PhD-level full time lecturers or tenure-track instructors of biology lecture courses, while the GTAs were all biology graduate students pursuing either a master's or PhD and were employed part-time to teach laboratory sections associated with the lecture sections.

The goal of this study was to investigate undergraduate perceptions of faculty and GTAs at the beginning and near the end of one semester of instruction with the following questions in mind:

Q1: Will the differences between hypothetical instructor types identified by Kendall and Schussler (2012), and supported by Dudley (2009), Muzaka (2009), and Park (2002), hold true for their current, nonhypothetical, instructors?

Q2: Will student ratings of their instructors change over the course of the semester?

Q3: Will student ratings of their instructors be related to the students' perceptions of how effective their instructors are at teaching or how much they learned from their instructors?

Q4: Will demographic variables, such as instructor teaching experience, instructor gender, and instructor age affect student ratings of their instructors?

Hypotheses for this study were:

H1: Ratings of GTAs and faculty members will differ given the previous research by Park (2002), Dudley (2009), Muzaka (2009), and Kendall and Schussler (2012).

H2: Undergraduate ratings of their instructors, as well as the relative ratings between instructor types, will change over the semester as students become more aware of the behaviors and attitudes of their instructors (Nussbaum, 1992; Helterbran, 2008; Pattison et al., 2011).

H3: Student descriptor ratings for their instructors will be correlated to students' perceptions of instructors' teaching effectiveness and students' perceptions of how much they learned from their instructors (Nussbaum, 1992; Helterbran, 2008; Pattison et al., 2011).

H4: Teaching experience, gender, and age of the instructors will be correlated with student ratings of these instructors for some of the descriptors in this study (Bos et al., 1980; Baird, 1987; Feldman, 1993; Basow, 1995; Greenwald and Gillmore,

1997; Centra and Gaubatz, 2000; Dunkin, 2002; Clayson et al., 2006; Zabaleta, 2007).

The documentation of undergraduate student perceptions of teaching and learning by these two instructor types will reveal aspects students consider important for effective instruction and which should be the focal point of professional development. This in turn will lead to a more thorough understanding of how to maximize student experience and retention in gateway undergraduate science courses taught by GTAs and faculty members.

Materials and Methods

Data Collection

Data were collected by means of "initial" and "final" course surveys administered to undergraduate students in Spring 2011 at a large southern research university in the United States. For the purposes of this study, "initial" refers to surveys students completed at the beginning of the second class meeting of the semester, while "final" refers to surveys students completed 2-3 wks prior to the end of the semester.

Undergraduate participants were recruited from semester-long majors and non-majors genera biology courses (non-majors: Introductory Biology; majors: Biodiversity and Cellular Biology) in which the instructors had agreed to participate in the study. Each of the selected courses had a lecture (ranging in size from 32 students to 225 students) and a laboratory (maximum of 25 students per section) component; the laboratories were all taught by GTAs, who typically taught two or three laboratory sections per semester, and the lecture classes were all taught by PhD-level faculty members (tenure-line [n = 3] or lecturers [n = 3]).

Although in some contexts lecturers may be considered contingent instructors, at this university they are highly-qualified (PhD-level), full-time, and permanent (some have worked

for more than 10 yrs at the university) faculty members most likely indistinguishable from a tenure-line faculty member to the majority of undergraduates. The only truly contingent instructors at this institution are part-time adjunct lecturers and post-doctoral associates (who were not a part of this study), or GTAs. Throughout the study, the verbiage used when students were asked to rate their faculty member or GTA was "instructor." This prevented us from using terms that would potentially bias student responses. Throughout the methods and results, the faculty members and GTAs will be referred to as different "instructor types."

Survey Design

The initial and final surveys consisted of 15 items that were identical except in terms of tense (Table 7). Thirteen items were created based on descriptors found to differ between instructor types in a study by Kendall and Schussler (2012) during which students rated hypothetical instructors. Of those 13 items, 12 were created by selecting two descriptors for GTAs and two descriptors for faculty from each of the three instructional themes (delivery technique, classroom atmosphere, and relationship) identified in Kendall and Schussler (2012). The descriptor "enthusiasm," which was attributed to different instructor types in different studies (Park, 2002; Muzaka, 2009; Kendall and Schussler, 2012), was also chosen as an item for this study. The wording of 11 of the items was as follows: "This instructor is …" with the appropriate descriptor inserted (e.g., "boring," "enthusiastic"). The wording for the other two descriptor items ("respect" and "relate") was "I … (to) this instructor." The final two items on the surveys asked students to rate the instructor's effectiveness in teaching the material (a statement taken directly from the university student evaluations) and how much the student learned (a modification of a statement from the university student evaluations).

At the top of each survey, undergraduate participants created a unique code (their birth date and the first four letters of their mother's maiden name), allowing the researcher to match initial and final responses as well as student responses for each instructor type. Undergraduates also responded to demographic questions, including gender, enrollment status, major, native language, current Introductory Biology course enrollment, and previous Introductory Biology courses completed. After the survey was created, the researchers asked undergraduates, graduate students, administrative staff, and faculty members to review the survey and provide feedback about its face validity; these individuals expressed that the survey was suitable for the study and that the instructions were clear.

The survey instructed undergraduate participants to rate their instructors based on how well each item described him/her by circling one choice of a six-point scale including: very poor, poor, fair, good, very good, or excellent. This six-point scale was chosen to match the rating scale on the university's student evaluation of teaching form which undergraduates complete each semester. Some undergraduates expressed confusion regarding the rating scale, so students were verbally informed during the survey administration that they could also think of "excellent" as being "strongly agree" and "very poor" as being "strongly disagree" with a continuum in-between.

Initial surveys were administered in a paper-pencil format by the researchers (authors of this paper) or a trained assistant in the lecture classes at the start of the second class meeting of the semester. Thus, students had anywhere from 50 to 75 min of exposure to the faculty member prior to completing the survey. Students were informed about the purpose of the study and were told that it was voluntary, not part of their grade, and the results would not be seen by their instructor (instructors were asked to leave the room for the administration of the surveys).

Survey completion took no more than 5 min. All surveys for the lecture instructors were completed prior to the surveys in the laboratory sections, because the laboratory surveys were administered at the beginning of the second week of lab. Thus, students may have had up to 2 hours and 50 min of exposure to a GTA prior to completing the survey. Because of the number of concurrent laboratory sections, the survey administration was done slightly differently for laboratories. GTAs were provided with a script that they read introducing the study, and then they chose a student to read the rest of the instructions while the GTA was out of the room. The surveys were collected in an envelope, sealed, and delivered to the main administrative office by one of the students in the lab section. The administration of the final surveys followed the same procedures and occurred between weeks 12 and 13 of the 15 wk semester.

Instructor demographics were collected from faculty members and GTAs by means of a questionnaire sent via email during week 11 of the semester. This questionnaire asked the instructor for his/her course number, section number(s), semesters of university teaching experience (not including the current semester), age, gender, native language, highest degree completed, and degree sought (if GTA).

An exemption from written consent was obtained for the study and no incentives were offered for participation. All procedures were reviewed and approved by the Institutional Review Board for Human Subjects.

Survey Validity and Reliability

The survey used in this study was designed utilizing descriptors from three instructional themes (delivery technique, classroom atmosphere, and relationship) identified from qualitative analysis of student responses from a previous study at the same institution (Kendall and Schussler, 2012). The validity of these descriptors rests with the strength of this prior research,

as well as research by others who have identified similar themes and descriptors in their own research (Bos et al., 1980; Park, 2002; Arnon and Reichel, 2007; Walker, 2008; Muzaka, 2009). Face validity, as previously described, was also used.

We verified the reliability of the items within each theme by calculating Cronbach's alpha for both initial and final surveys, although it is important to note that analyses of the results was for individual items and not for themes. The delivery technique theme consisted of five items including nervous, uncertain, organized, enthusiastic, and confident. The classroom atmosphere and relationship themes consisted of four items each (relaxed, engaging, distant, and strict for classroom atmosphere; relate, understanding, respect, and boring for relationship). Cronbach's alpha results for the initial survey were: 0.731 for delivery technique, 0.631 for classroom atmosphere, and 0.761 for relationship. Cronbach's alpha results for the final survey were 0.816, 0.711, and 0.831, respectively, for delivery technique, classroom atmosphere, and relationship. These values represent a valid internal consistency for the items chosen from each theme.

Data Analysis

Only undergraduate respondents (n = 255) who completed both the initial and final surveys for both their faculty member and GTA were included in the data analysis. At the beginning of the analyses, the Likert-scale ratings for each survey item for each instructor were converted to numeric responses in the following manner: excellent = 6, very good = 5, good = 4, fair = 3, poor = 2, very poor = 1.

Data from the instructor demographics were also converted to numeric responses for analysis. University teaching experience was grouped as follows: 0-1 semesters (coded as "0"), 2-4 semesters (coded as "1"), 5-8 semesters (coded as "2"), and more than 9 semesters (coded as

“3”). Because the study took place in the Spring semester, few instructors had 0 semesters of experience; thus 0-1 semesters of teaching experience indicated instructors with little to no teaching experience. The ranges of 2-4 semesters, 5-8 semesters, and more than 9 were chosen based on the time it typically takes to obtain a master’s degree (2-4 semesters), and PhD (5-8 semesters). However, not all GTAs teach while obtaining their degrees, and students complete graduate degrees at various paces. Gender was converted to a 1 for females and a 2 for males. Instructor age was grouped as follows: 25 and under (coded as “1”), 26-30 (coded as “2”), 31-40 (coded as “3”), 41-50 (coded as “4”), and 51-60 (coded as “5”). These age ranges were chosen based on results from Bos et al. (1980) in which differences in perception of teaching were found for instructors in their early twenties as compared with those in their later twenties.

The analyses were performed using nonparametric methodology because the Likert-type choices on the surveys were ordinal, and there is no guarantee that students perceive the difference between intervals on the point scale equally (e.g., “Excellent” and “Very Good” may not be the same distance to a student as the distance from “Very Good” and “Good”). This theoretical lack of equal distances between responses violates assumptions for parametric methodology, requiring the data to be analyzed via nonparametric tests (Huck, 2008). The analyses followed the main questions of the study as outlined below.

Will student ratings differ between instructor types? Do these ratings change over a semester? Wilcoxon signed rank tests were conducted to evaluate whether student ratings on the initial surveys and final surveys differed significantly ($\alpha = 0.05$) by instructor type (faculty vs. GTAs). Wilcoxon signed rank tests were also conducted for each instructor type to determine whether student ratings changed significantly ($\alpha = 0.05$) over the course of the semester (SPSS 19.0).

Will student ratings of their instructors be related to the students' perceptions of how effective their instructors' teaching is or how much they learned from their instructors? Will student ratings be related to demographic characteristics of their instructors? Spearman rank correlations were performed on pooled data (GTAs and faculty together) to determine whether university teaching experience, age, or gender correlated with student initial or final ratings for each of the thirteen descriptors ($\alpha = 0.05$; SPSS 19.0). Spearman rank correlations were also performed on the pooled data to determine whether student perception of their instructor's teaching effectiveness or student perception of how much they learned from their instructors, as rated on the final surveys, correlated with any of the 13 student ratings of descriptors on the final surveys ($\alpha = 0.05$; SPSS 19.0).

Results

Participants

Undergraduate Students. The surveys were administered to a potential pool of 1170 students in eight lecture sections and 43 associated laboratory sections. After removing students who did not complete both initial and final surveys for both their professor and GTA, 255 participants remained. This response rate was a result of the survey being voluntary, offering no incentives, dropping of courses by students, or absences during any of the survey administrations.

Twenty-two percent of the individuals who completed the surveys were enrolled in second semester Introductory Biology for non-majors, 36% in majors Biodiversity, 13% in majors Cellular Biology, and 29% in majors Genetics. These respondents were mostly female (62%), native English speakers (98%), first-year students (49%), and non-biology majors (67%). Second- and third-year students comprised 25 and 19%, respectively, of the respondents, while

7% were fourth-year students or beyond. Thirty-two percent of students were Biology majors, with 3% declaring a concentration in Ecology and Evolutionary Biology, 5% in Microbiology, and 15% in Biochemistry, Cellular and Molecular Biology. Thirty-one percent of respondents had not completed another biology course at the university; however, 39% had completed one other semester-long lecture/laboratory biology course. Those who completed two or three courses comprised 20 and 5%, respectively; an additional 3% had completed four or more other biology courses. Of the participants who completed other courses, 31% had completed first semester Introductory Biology, 9% second-semester Introductory Biology, 32% Biodiversity, 30% Cellular Biology, 3% Ecology, 2% first-semester Botany, and 1% second-semester Botany.

Instructors. The instructors of the students who participated in the study were six faculty members teaching eight lecture sections, and 25 GTAs teaching 43 laboratory sections. The eight lecture sections consisted of three (38%) second semester non-majors' Introductory Biology classes, two (25%) Biodiversity, one (13%) Cellular Biology, and two (25%) Genetics classes.

Faculty Members. The faculty members in this study had from 3-45 semesters of university-level teaching experience, with an average of 16 semesters. Sixty-seven percent of the faculty members were male, while 33% were female. Seventeen percent of faculty members were between the ages of 31 and 40, 50% were between 41and 50, and 33% were 51and 60. The majority of the faculty members were native English speakers (83%). All faculty members held a PhD in a life sciences field.

GTAs. Teaching experience for the GTAs in this study ranged from 0to12 semesters of university-level instruction, with an average of four semesters. Of the 25 GTAs, 28% were male, and 72% were female. The majority of the GTAs were under the age of 30, with 32% being under 25 years old and 48% between 26 and 30 years old. The remaining 20% of GTAs were

between the ages of 31and 40 (12%) and 41and 50 (8%). The majority of GTAs were native English speakers (64%). Forty percent of GTAs were teaching labs associated with second semester Introductory Biology, 28% Biodiversity, 12% Cellular Biology, and 20% Genetics. Thirty-six percent of GTAs had earned a master's and bachelor's degree already, while the remaining 64% of GTAs had only completed a bachelor's degree. Of the participating GTAs, 84% were in graduate school to pursue a PhD, while 16% were seeking a master's degree.

Will student ratings differ between instructor types? Do these ratings change over a semester?

Initial Ratings. The responses provided by students indicated that they initially perceived differences between GTAs and faculty members for the following descriptors: confident ($Z = -4.623$, $p = 0.000$), engaging ($Z = -2.998$, $p = 0.003$), enthusiastic ($Z = -4.343$, $p = 0.000$), nervous ($Z = -4.053$, $p = 0.000$), and uncertain ($Z = -3.828$, $p = 0.000$; Figure 1). These ratings indicated that students perceived GTAs to be more nervous and uncertain, and faculty members to be more confident, engaging, and enthusiastic. Descriptive statistics by instructor type for initial and final ratings are shown in Table 8.

Final Ratings. By the end of the semester, participants indicated that GTAs differed from faculty members for the following descriptors: boring ($Z = -2.918$, $p = 0.004$), engaging ($Z = -2.446$, $p = 0.014$), organized ($Z = -4.546$, $p = 0.000$), relate ($Z = -2.602$, $p = 0.009$), respect ($Z = -2.075$, $p = 0.038$), and understanding ($Z = -3.311$, $p = 0.001$; Figure 2). These final ratings indicated a student perception that their faculty members were more boring than GTAs, and that their GTAs were more engaging, organized, relatable, respected, and understanding than faculty members.

Change in Ratings. The initial to final changes in student ratings are organized by those that occurred only for GTAs (Figure 3), those that occurred only for faculty members (Figure 4), and those that occurred for both faculty and GTAs (Figure 5). GTAs had changes in ratings for the descriptors of nervous ($Z = -3.808$, $p = 0.000$) and understanding ($Z = -4.278$, $p = 0.000$). GTAs decreased in perceptions of nervousness, and increased in perceptions of understanding from the initial to final surveys (Figure 3). Meanwhile, faculty members had changes in ratings for the descriptors of boring ($Z = -3.952$, $p = 0.000$), organized ($Z = -5.138$, $p = 0.000$), relate ($Z = -2.707$, $p = 0.007$), relaxed ($Z = -2.598$, $p = 0.009$), and strict ($Z = -2.408$, $p = 0.016$; Figure 4). The student participants indicated that faculty members were more boring by the end of the semester, and less organized, relaxed, relatable, and strict than GTAs (Figure4).

Ratings that changed for both faculty and GTAs over the semester included confident, engaging, respect, and uncertain. Significance values for GTAs were: confident ($Z = -3.547$, $p = 0.000$), engaging ($Z = -1.972$, $p = 0.049$), respect ($Z = -2.035$, $p = 0.042$), and uncertain ($Z = -3.034$, $p = 0.002$); for faculty members: confident ($Z = -2.657$, $p = 0.009$), engaging ($Z = -4.245$, $p = 0.000$), respect ($Z = -2.598$, $p = 0.009$), and uncertain ($Z = -2.130$, $p = 0.033$; Figure 5). For each descriptor, GTA and faculty changes were in the opposite direction. For the descriptors of confident, engaging, and respect, GTAs increased in their ratings from the initial to final surveys, whereas faculty members decreased in their ratings. For the descriptor of uncertain, GTAs decreased in their ratings while faculty members increased (Figure 5).

Will student ratings of their instructors be related to the students' perceptions of how effective their instructors' teaching is or how much they learned from their instructors?

Teaching effectiveness. For the final survey, pooled undergraduate ratings of their instructor's effectiveness in teaching the material found that these ratings correlated significantly with all final survey ratings of instructor descriptors ($\alpha = 0.05$). All *p* values were 0.000, except strict which was 0.005. There were strong positive associations between final survey ratings of instructor effectiveness and ratings of confident ($r = 0.568$), engaging ($r = 0.696$), enthusiasm ($r = 0.588$), organized ($r = 0.667$), relate ($r = 0.653$), relaxed ($r = 0.612$), respect ($r = 0.686$), and understanding ($r = 0.619$). Contrarily, there was a strong negative association between perceptions of teaching effectiveness and ratings of boring ($r = -0.522$), moderately negative associations for distant ($r = -0.476$), nervous ($r = -0.342$), and uncertain ($r = -0.472$), and a weak negative association with strict ($r = -0.120$).

Student Learning. Pooled results of undergraduate perceptions of how much they learned in the course on the final survey correlated significantly with final student ratings of all descriptors except strict ($\alpha = 0.05$, for all significant variables $p = 0.000$). These analyses indicated strong positive associations between ratings of student learning and ratings of confident ($r = 0.515$), engaging ($r = 0.595$), enthusiasm ($r = 0.525$), organization ($r = 0.570$), relate ($r = 0.574$), relaxed ($r = 0.528$), respect ($r = 0.611$), and understanding ($r = 0.517$). Contrarily, moderate negative associations were found between ratings of student learning and perceptions of boring ($r = -0.446$), distant ($r = -0.410$), nervous ($r = -0.306$), and uncertain ($r = -0.441$).

Will student ratings be related to demographic characteristics of their instructors?

Initial Ratings. Weak positive associations were identified between university teaching experience and initial student ratings for confident ($r = 0.269$, $p = 0.000$), engaging ($r = 0.154$, $p = 0.001$), relaxed ($r = 0.145$, $p = 0.001$), and strict ($r = 0.095$, $p = 0.031$), while a moderately positive association was identified for the rating of enthusiastic ($r = 0.414$, $p = 0.001$). There

were also weak negative associations between university teaching experience and ratings of nervous ($r = -0.224$, $p = 0.000$) and uncertain ($r = -0.193$, $p = 0.000$). The highest mean student ratings for confident, relaxed, and engaging were found for the two more experienced instructor groups (5-8 and 9+ semesters). Instructors with the least teaching experience (0-1 and 2-4 semesters) had the highest means for uncertain and nervous. Interestingly, mean student ratings for strict and enthusiastic were highest for instructors with the least and most teaching experience (0-1 and 9+ semesters). For the descriptors in which significant associations were found with teaching experience, descriptive statistics are included in Table 9. In terms of gender, there were weak negative associations for the ratings of organized ($r = -0.153$, $p = 0.001$) and uncertain ($r = -0.100$, $p = 0.024$), with females being rated higher for both (descriptive statistics in Table 9). Finally, instructor age had weak positive associations with initial student ratings for confident ($r = 0.222$, $p = 0.000$), engaging ($r = 0.128$, $p = 0.004$), enthusiasm ($r = 0.202$, $p = 0.000$), and relaxed ($r = 0.103$, $p = 0.020$). There were weak negative associations between instructor age and ratings of distant ($r = -0.112$, $p = 0.012$), nervous ($r = -0.230$, $p = 0.000$), and uncertain ($r = -0.214$, $p = 0.000$). Instructors aged 26-30, 41-50, and 51-60 had the highest mean ratings for the descriptor confident. Instructors classified in age groups of 26-30, 31-40, and 51-60 had the highest mean ratings for the descriptor relaxed. The highest mean student ratings for engaging and enthusiastic were for instructors in the oldest three age groups (31-40, 41-50, and 51-60). However, instructors belonging to the three youngest age groups (under 25, 26-30, and 31-40) had the highest mean student ratings for uncertain and nervous. Lastly, instructors classified in the youngest two groups (under 25, 26-30) and the oldest age group (51-60) had the highest mean ratings for distant. Descriptive statistics for these significant associations with instructor age can be found in Table 9.

Final Ratings. At the end of the semester, there were weak positive associations between university teaching experience and student ratings of confident (r = 0.148, p = 0.001) and strict (r = 0.097, p = 0.029). There were weak negative associations between teaching experience and ratings of nervous (r = -0.148, p = 0.001), organized (r = -0.089, p = 0.044), uncertain (r = -0.122, p = 0.006), and understanding (r = -0.150, p = 0.001). Mean student ratings for the descriptor confident were highest for the two more experienced instructor groups (5-8 and 9+ semesters of teaching experience). Meanwhile, instructors with the least teaching experience (0-1 and 2-4 semesters) had the highest means for uncertain and nervous. Mean student ratings for strict were highest for instructors with the least and most teaching experience (0-1 and 9+ semesters). Meanwhile, the highest mean ratings for organized and understanding were found for instructors with 2-4 and 5-8 semesters teaching experience. Descriptive statistics can be found in Table 10 for all descriptors with significant associations with teaching experience. For gender, there were weak negative associations between gender and ratings for organized (r = -0.129, p = 0.004), and understanding (r = -0.107, p = 0.015), whereby females were perceived as more organized and understanding (descriptive statistics in Table 10). Finally, instructor age had a weak positive association with final ratings for boring (r = 0.109, p = 0.014) and weak negative associations with final ratings for engaging (r = -0.107, p = 0.015), organized (r = -0.227, p = 0.000), relate (r = -0.134, p = 0.003), and understanding (r = -0.149, p = 0.001). Instructors belonging to the three youngest age groups (under 25, 26-30, and 31-40) had the highest mean ratings for organized, engaging, and relate, while instructors classified in the groups of 26-30, 31-40 and 51-60 had the highest mean ratings for understanding. Instructors in the youngest age group (under 25) and the oldest two age groups (41-50 and 51-60) had the highest mean student

ratings for boring. Descriptive statistics can be found in Table 10 for all descriptors which had significant associations with instructor age.

Discussion

Although some descriptors associated with hypothetical instructors (Kendall and Schussler, 2012) held true in non-hypothetical educational settings, it is apparent that reality did not always match with the theoretical. Interestingly, however, student perceptions of their instructors evolved over the course of the semester, with GTAs being more likely than faculty members to gain in student perception of positive instructional characteristics by the end of the semester. Weak associations were identified between instructor demographics and student ratings, suggesting that the age, gender, or teaching experience of an instructor may not impact student ratings as much as individual instructional abilities. This is evidenced by the moderate to strong associations between student ratings of instructor teaching effectiveness and perception of learning and all student ratings of the instructor descriptors except *strict*.

Student Expectations. On the basis of previous research, we expected faculty members to be rated as more boring, confident, distant, enthusiastic, organized, respected and strict than GTAs, and we expected GTAs to be rated as more engaging, nervous, relatable, relaxed, uncertain, and understanding than faculty members at the beginning of the semester (indicating that these may be stereotypes of each instructor type) (Park, 2002; Dudley, 2009; Muzaka, 2009; Kendall and Schussler, 2012). In reality, only a few of these expectations held true for actual instructors. As expected, undergraduates rated faculty members higher initially for being *confident* and *enthusiastic* while GTAs were rated higher for being *nervous* and *uncertain*. However, undergraduate students also rated faculty members higher for *engaging*, which was contrary to previous research. No differences between instructor types were seen for initial

ratings of *boring*, *distant*, *organized*, *relate*, *relaxed*, *respect*, *strict*, and *understanding*, suggesting that reality trumps stereotypes for these characteristics.

These results indicate that either initial instructor impressions were enough to overcome any pre-existing expectations students might have about their instructors or that many of the previously identified descriptor differences are not strongly engrained stereotypes. However, for the descriptors that did hold true, faculty members benefited from more positive descriptions than GTAs. Therefore, if stereotypes do exist, they seem to differentially benefit faculty members.

Change in Ratings. Ambady and Rosenthal (1992) noted that brief exposure to behavior provides enough information for an individual to predict with significant accuracy their perception of others. This suggests that student ratings of their instructors should be stable over a semester, yet student perceptions of both GTAs and faculty evolved over the semester in this study. A possible explanation is that it may depend on the descriptor; those related to more overt qualities such as *nervous* and *uncertain* may be easier to quickly judge than qualities such as *relate* and *understanding* (Ambady and Rosenthal, 1992). These types of descriptors may only be judged when students see how the instructor deals with periodic, situational issues over the semester (such as an unfair exam questions or missing class because of a car breakdown). Birch et al. (2012) found that cues such as voice clarity, preparation, and classroom control impact student impressions and expectations of lecturers; however, these cues may represent only a subset of the behaviors that students use to judge an instructor over a semester. Thus, future work should explore what factors are influencing student perceptions of their instructors, as well as changes in these perceptions, and whether instructors can direct these changes through the

display of specific behaviors or attitudes as proposed by Helterbran (2008), Nussbaum (1992), and Pattison et al. (2011).

When undergraduate student perception changed, the trend of these changes was that GTAs gained in ratings of more positive descriptors over the course of the semester whereas faculty decreased in student perceptions of positive descriptors. This may reflect students entering the classroom with certain expectations for their instructors (Helterbran, 2008; Pattison et al., 2011), and having higher expectations for a faculty member than they do for a GTA. If this is the case, it appears that faculty members were not able to meet all of the expectations of students who participated in this study, and GTAs were more easily able to meet and exceed student expectations for them. However, these differences may also be contextual, such as differences in classroom setting (lecture vs. laboratory). For example, students may expect instructors to be more engaging in smaller laboratory settings than in lecture settings, therefore engaging lecture instructors may more easily meet or exceed student expectations. Student ratings on teaching evaluations have been found to be influenced by student grades (Baird, 1987; Greenwald and Gillmore, 1997; Clayson et al., 2006), therefore, faculty ratings may decline over the semester because lecture course work is typically worth more than laboratory course work so lower grades in lecture would disproportionally affect the student's overall course grade.

The faculty and GTA descriptors which changed in opposite directions from each other (*confident*, *engaging*, *respect*, and *uncertain*) may be the most susceptible to these instructor or contextual expectations, and also the easiest to change through classroom interactions. This is important for both GTAs and faculty to understand so they can consider why students rate them lower or higher on these aspects over the semester, and actively work either to maintain or gain

in the positive aspects. These aspects should also be a point of discussion in professional development and mentoring of university instructors.

Teaching Effectiveness and Student Learning. All of the student descriptor ratings were significantly associated (moderately or strongly, except *strict* which was weak) with undergraduate perception of instructor teaching effectiveness, and all (except *strict)* with perceived student learning. This intimate relationship between student perception of each descriptor and ratings of teaching effectiveness and student learning is an important reminder to instructors and administrators of the many facets – some of which may not be thought of as instructional skills – that influence student perception of teaching. For instance, an instructor who is not rated well in teaching effectiveness by students may simply not relate well to students. Zabaleta (2007) cautioned that factors such as lifestyle likeness or age group similarity (e.g., descriptors such as ability to *relate* or *understand*) may be influencing student ratings on evaluations. Translating instructor ratings into specific behaviors students use to judge them would make it easier for instructors to make changes to their instructional delivery that could ultimately also improve student perception of instructor teaching effectiveness.

We note that *strict* was unusual in this study because it was not viewed by students the same for both teaching and learning (student ratings of strict were negatively correlated with teaching effectiveness, whereas there was a positive, but not significant, correlation between student ratings of strict and their perceptions of how much they learned). In addition, correlations between student ratings of strict and student perceptions of teaching effectiveness and learning were not as strong as with the other descriptors. In another study (Kendall and Schussler, 2013), students explained that *strict* had both negative and positive connotations to students. For instance, a strict instructor may be unbending in their rules to the point of being unfair. However,

strict can also be characteristic of a respected instructor because the instructor enforces polite behavior in a large classroom or fairly applies the same rules to all students. This potentially explains why *strict* had lower correlations with teaching effectiveness and was not correlated to student learning; however, future work is necessary to elucidate facets underlying, and perhaps confounding, student perceptions of effective instruction and learning.

Instructor Demographic Variables. Although there were associations among student ratings of their instructors and instructor demographics, these associations were weak, with the exception of university teaching experience and enthusiasm. The results for some of the descriptors support the findings by Zabaleta (2007) and Bos et al. (1980) that instructor age may influence student ratings, and that gender can influence student ratings (Feldman, 1993; Basow, 1995; Centra and Gaubatz, 2000). However, there were also results indicating that experience (novice vs. experienced) and instructor gender (Bos et al., 1980; Dunkin, 2002; Zabaleta, 2007) had no impact on student ratings of certain descriptors.

It may be more instructive to identify which descriptor ratings and instructor demographics were correlated with each other on both the initial and final surveys (Table 11). For example, *confident*, *nervous*, *strict*, and *uncertain* were significantly associated with teaching experience for both initial and final ratings, with *confident* and *strict* being weak positive associations and *nervous* and *uncertain* being weak negative associations. *Organized* was associated with gender on both the initial and final surveys, with females being perceived as being more organized than males. Finally, *engaging* was consistently perceived as being associated with age, but the association changed (according to shifts in student ratings from the initial to final survey) from a weak positive to a weak negative association. Linking these results with the consistent differences between instructor types (faculty are *confident* and *enthusiastic*

and GTAs are *uncertain* and *nervous*) may indicate that teaching experience is one of the most important distinguishing features between faculty and GTAs for students.

The notion that student ratings are influenced by these instructor demographics, however, must be treated with caution since student perception of teaching experience and age may not be accurate (also noted in Zabaleta, 2007). For instance, a young looking faculty member may have more teaching experience and be older than students perceive. So, some of the differences identified between instructor types may result from GTAs being perceived as having less teaching experience than faculty, which may or may not always be true. A follow-up study could document actual teaching experience and student perceptions of teaching experience to relate to student descriptor ratings. Regardless, these results once again highlight the influence confounding factors, such as age, teaching experience and gender of instructors, can have on student perceptions, and that these factors should be considered when administrators and instructors review evaluations of teaching.

Also unclear is whether student ratings of the descriptors are accurate reflections of classroom practice. While we quantitatively identified positive and negative aspects of instructors (similar to Pattison et al. [2011] and Helterbran [2008] identifying effective and ineffective instructional aspects) we recommend that there be further exploration regarding student perceptions and actual instructor practice. For instance, students in this study reported that GTA *confidence* increased over the semester, but it is unknown whether GTAs were actually becoming more confident, or were merely exceeding undergraduate expectations, or undergraduate students were interpreting them to be more confident (without GTAs feeling more confident). It is also possible that discipline, or content within a discipline, impacts student perception of their instructors. For instance, students may indicate that being *engaging* is

important for biology but not chemistry. Additional studies could further clarify the types of professional development to offer to instructors of different courses and even disciplines.

Conclusion

This study has revealed the complex perceptions students have about GTAs and faculty members as instructors. While some demographic characteristics impact undergraduate student perception of GTAs and faculty members, it appears that student expectations and the nature of individual instructors themselves play a more vital role in student perception. Even when differences between instructor types were identified, GTAs increased in ratings for many descriptors viewed as positive by undergraduates while they decreased in some negative descriptors over the course of a semester. This indicates that student learning may not be as hindered by the use of contingent instructors in academia as previously suspected. While future studies are still necessary to identify whether specific instructor behaviors affect rating changes over the semester, the results of this study provide support for Baldwin and Wawrzynski's (2011) call for targeted development strategies for different instructor types.

To enhance instructional abilities of GTAs, for instance, institutions or departments could help GTAs identify their strengths and weaknesses based on the descriptors of this study and use these results to form groups of GTAs who already have, or need to develop, particular instructional capacities. GTAs can also be given the opportunity to microteach, or observe, laboratories prior to teaching their own sections in order to obtain feedback and help them plan for their own teaching. This may also decrease GTA uncertainty and enhance GTA confidence in their instruction. While this study found that GTAs increased in perceptions of descriptors positively associated with teaching effectiveness over the semester, professional development may also need to emphasize that students do see value in some less positive characteristics. For

instance, it may be instructive to review the positive and negative aspects of strict with GTAs to highlight the subtle dynamics that make a difference to students in the classroom. GTAs can be encouraged to self-reflect on the teaching they are providing to undergraduates each class and identify areas of instructional weakness. Institutions can support this self-reflection by providing GTAs with the opportunity to video-record class sessions, enabling GTAs to observe their instructional behaviors more objectively. Given that student ratings can change over the course of the semester, we also recommend that GTAs encourage students to provide them with regular feedback so GTAs can make strategic modifications to their behaviors throughout the semester (Keutzer, 1993).

Undergraduates who find their introductory sciences courses to be enjoyable and educational may be more likely to pursue a degree in the sciences (Seymour and Hewitt, 1997). Accordingly, introductory science instructors are an important component in fostering student retention and learning in the sciences and the instructional abilities of faculty and GTAs should be fostered and developed by their institutions. Through researching student perceptions, instructors and professional development coordinators can better understand what undergraduates expect from instructors, and how this relates to student perceptions of teaching effectiveness and learning in their courses. Through this research, targeted professional development opportunities for each instructor type involved in gateway science courses can be developed to maximize student retention and learning in all aspects of these courses, regardless of instructor type.

Literature Cited

Ambady N, Rosenthal R (1992). Thin slices of behavior as predictors of interpersonal consequences: a meta-analysis. Psychol Bull *2*, 256–274.

Arnon S, Reichel N. (2007). Who is the ideal teacher? Am I? Similarity and difference in perception of education regarding the qualities of a good teacher and of their own qualities as techers, Teachers and Teaching *13*, 441-464.

Baird JS (1987). Perceived learning in relation to student evaluation of university instruction. J Educ Psychol *79*, 90–91.

Baldwin RG,WawrzynskiMR (2011). Contingent faculty as teachers: what we know; what we need to know. Am Behav Sci *55*, 1485–1509.

Basow SA (1995). Student evaluations of college professors: when gender matters. J Educ Psychol *87*, 656–665.

Benjamin E (2002). How over reliance upon contingent appointments diminishes faculty involvement in student learning. Peer Review *5*, 4–10.

Birch PDJ, Batten J, Manley AJ, Smith MJ (2012). An exploratory investigation examining the cues that students use to form initial impressions and expectancies of lecturers. Teach Higher Educ *17*, 1–13.

Bolge RD (1995). Examination of Student Learning as a Function Of Instructor Status (Full-Time vs. Part-Time) at Mercer Community College,WestWindsor, NJ: Mercer Community College.

Bos RR, Zakrajsek DB,Wolf V, Stoll S (1980). Teaching assistant traits: their influence on student ratings. Improv Coll Univ Teach *28*, 179–185.

Centra JA, Gaubatz NB (2000). Is there gender bias in student evaluations of teaching? J Higher Educ *71*, 17–33.

Clayson DE, Frost TR, Sheffet MJ (2006). Grades and the student evaluation of instruction: a test of the reciprocity effect. Acad Mgmt Learn Educ *5*, 52–65.

Clayson DE, Haley DA (2011). Are students telling us the truth? A critical look at student evaluation of teaching. Marketing Educ Rev *21*, 101–112.

d'Apollonia S, Abrami RC (1997). Navigating student ratings of instruction. Am Psychol *52*, 1198–1208.

Dudley M (2009). Jumping out of an airplane: a TA's perspective on teaching effectiveness. East Educ J *38*, 1–10.

Dunkin MJ (2002). Novice and award-winning teachers' conceptsand beliefs about teaching in higher education. In: Teacher Thinking, Beliefs, and Knowledge in Higher Education, ed. N Hativa and P Goodyear, Dordrecht, The Netherlands: Kluwer Academic.

EmeryCR, Kramer TR, TianRG (2003). Return to academic standards: a critique of students' evaluations of teaching effectiveness. Qual Assur Educ *11*, 37–46.

Feldman KA (1993). College students' views of male and female college teachers: Part IIL. Evidence from students' evaluations of their classroom teachers. Res Higher Educ *34*, 151–211.

Golde CM, Dore TM (2001). At Cross Purposes: What the Experiences of Doctoral Students Reveal about Doctoral Education, Philadelphia, PA: Pew Charitable Trusts.

Greenwald AG, Gillmore GM (1997). Grading leniency is a removable contaminant of student ratings. Am Psychol *52*, 1209–1217.

Helterbran VR (2008). The ideal professor: student perceptions of effective instructor practices, attitudes, and skills. Education *129*, 125–138.

Huck S (2008). Reading Statistics and Research, Boston: Pearson Education.

Jaeger AJ (2008). Contingent faculty and student outcomes. Academe *94*, 42–43.

Johnson IY (2011). Contingent instructors and student outcomes: an artifact or fact? Res Higher Educ *52*, 761–785.

Kendall KD, Schussler EE (2012). Does instructor type matter? Undergraduate student perception of graduate teaching assistants and professors. CBE Life Sci Educ *11*, 187–199. Keutzer CS (1993). Midterm evaluation of teaching provides helpful feedback to instructors. Teach Psychol *20*, 238–240.

Kendall KD, Schussler EE. 2013. More than words: Probing the terms undergraduate students use to describe their instructors. International Journal of Teaching and Learning in Higher Education *25*(2).

Kogan LR, Schoenfeld-Tacher R, Hellyer PW (2010). Student evaluations of teaching: perceptions of faculty based on gender, position, and rank. Teach Higher Educ *15*, 623–636.

Marsh HW, Roche LA (1997). Making students' evaluations of teaching effectiveness effective: the critical issues of validity, bias, and utility. Am Psychol *52*, 1187–1197.

Muzaka V (2009). The niche of graduate teaching assistants (GTAs): perceptions and reflections. Teach Higher Educ *14*, 1–12.

Nussbaum JF (1992). Effective teacher behaviors. Commun Educ *41*,167–180.

O'Neal C,Wright M, Cook C, Perorazio T, Purkiss J (2007). The impact of teaching assistants on student retention in the sciences: lessons for TA training. J Coll Sci Teach *36*, 24–29.

Park C (2002). Neither fish nor fowl? The perceived benefits and problems of using graduate teaching assistants (GTAs) to teach undergraduate students. High Educ Rev *35*, 50–62.

Pattison P,Hale JR,Gowens P (2011).Mind and soul: connecting with students. J Legal Stud Educ *28*, 39–66.

President's Council of Advisors on Science and Technology (PCAST) (2012). Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Mathematics. Report to the President. www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-engage-to-excel-final_2–25–12.pdf. Accessed 3. April. 2012.

Rushin JW, Saix JD, Lumsden A, Streubel DP, Summers G, Bernson C (1997). Graduate teaching assistant training: a basis for improvement of college biology teaching and faculty development? AmBiol Teach *59*, 86–90.

Seymour E, Hewitt N (1997). Talking about Leaving: Why Undergraduates Leave the Sciences, Boulder, CO: Westview.

Sundberg MD, Armstrong JE, Wischusen EW (2005). A reappraisal of the status of introductory biology laboratory education in U.S. colleges & universities. Am Biol Teach *67*, 525–529.

Umbach PD (2007). How effective are they? Exploring the impact of contingent faculty on undergraduate education. Rev Higher Educ *30*, 91–123.

Walker R (2008). Twelve characteristics of an effective teacher: a longitudinal qualitative, quasi-research study of in-service and pre-service teachers' opinions. Educ Horizons *87*, 61–68.

Zabaleta F (2007). The use and misuse of student evaluations of teaching. Teach Higher Educ *12*, 55–76.

Appendix

Table 7. Survey items asking students to rate their instructor for thirteen descriptors as well as teaching effectiveness and student learning[a]

Survey Items
This instructor is confident.
This instructor is organized.
This instructor is relaxed.
This instructor is engaging.
This instructor is uncertain.
This instructor is nervous.
This instructor is strict.
This instructor is distant.
I relate to this instructor.
This instructor is understanding.
I respect this instructor.
This instructor is boring.
This instructor is enthusiastic.
The instructor's effectiveness in teaching material.
The amount you will learn in the course.

[a] Rating scale consisted of "excellent," "very good," "good," "fair," "poor," and "very poor". Initial and final survey differed only in tense (i.e., The amount you will learn [initial] and The amount you learned in the course [final]).

Table 8. Descriptive statistics including mean, standard deviation (SD), median (med), minimum (min), and maximum (max) by instructor type for initial and final ratings of descriptors (in alphabetical order).

		Graduate Teaching Assistants					Faculty Members				
	Descriptor	Mean	SD	Med	Min	Max	Mean	SD	Med	Min	Max
Initial Rating	Boring	2.72	1.18	2	1	6	2.62	1.10	2	1	6
	Confident	4.86	0.90	5	1	6	5.22	0.79	5	2	6
	Distant	2.48	1.11	2	1	6	2.35	1.12	2	1	6
	Engaging	4.49	1.11	5	1	6	4.73	1.09	5	1	6
	Enthusiastic	4.32	1.08	4	1	6	4.69	1.07	5	1	6
	Nervous	2.39	1.07	2	1	5	2.07	1.11	2	1	6
	Organized	4.84	0.96	5	1	6	4.82	0.90	5	2	6
	Relate	3.69	1.14	4	1	6	3.69	1.12	4	1	6
	Relaxed	4.69	1.08	5	2	6	4.81	1.05	5	1	6
	Respect	4.77	0.98	5	1	6	4.90	1.04	5	1	6
	Strict	3.25	1.21	3	1	6	3.31	1.16	3	1	6
	Uncertain	2.42	1.05	2	1	6	2.11	1.10	2	1	6
	Understanding	4.27	1.00	4	1	6	4.29	1.03	4	1	6
Final Rating	Boring	2.68	1.29	2	1	6	3.00	1.35	3	1	6
	Confident	5.04	0.95	5	2	6	5.04	0.93	5	1	6
	Distant	2.32	1.19	2	1	6	2.48	1.16	2	1	6
	Engaging	4.62	1.21	5	1	6	4.34	1.27	4	1	6
	Enthusiastic	4.41	1.15	4	1	6	4.54	1.08	5	1	6
	Nervous	2.10	1.09	2	1	6	2.10	1.05	2	1	6
	Organized	4.89	1.13	5	1	6	4.40	1.22	5	1	6
	Relate	3.74	1.37	4	1	6	3.42	1.33	3	1	6
	Relaxed	4.80	1.22	5	1	6	4.65	1.14	5	1	6
	Respect	4.87	1.15	5	1	6	4.67	1.26	5	1	6
	Strict	3.24	1.34	3	1	6	3.10	1.26	3	1	6
	Uncertain	2.16	1.06	2	1	6	2.29	1.16	2	1	6
	Understanding	4.56	1.21	5	1	6	4.24	1.22	4	1	6

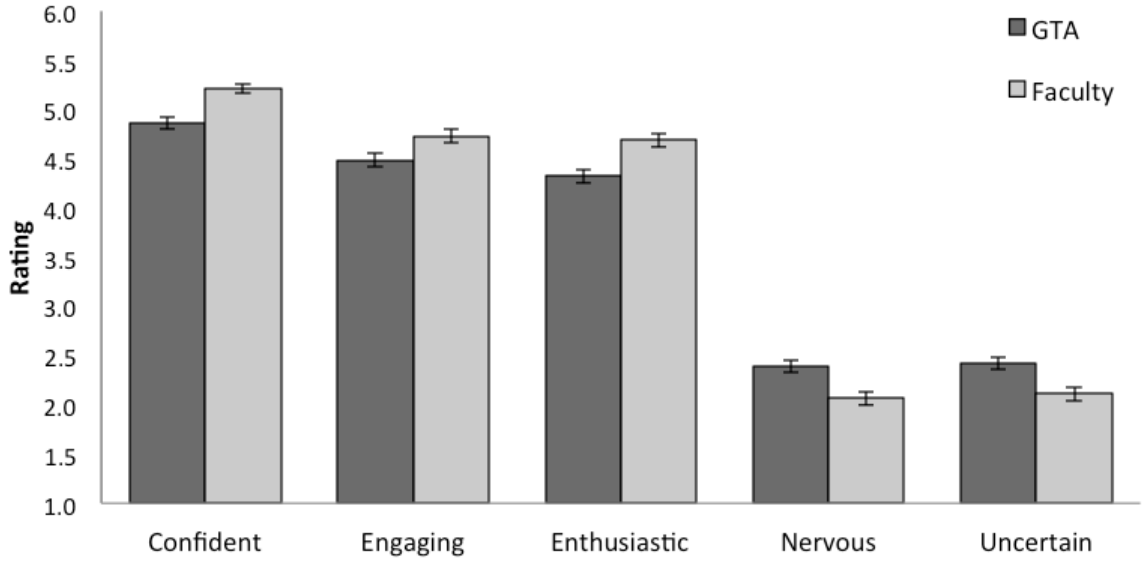

Figure 1. Mean ± standard error for descriptors (in alphabetical order) in which GTAs differed significantly from faculty members for initial ratings (Wilcoxon signed-rank tests [$\alpha = 0.005$]).

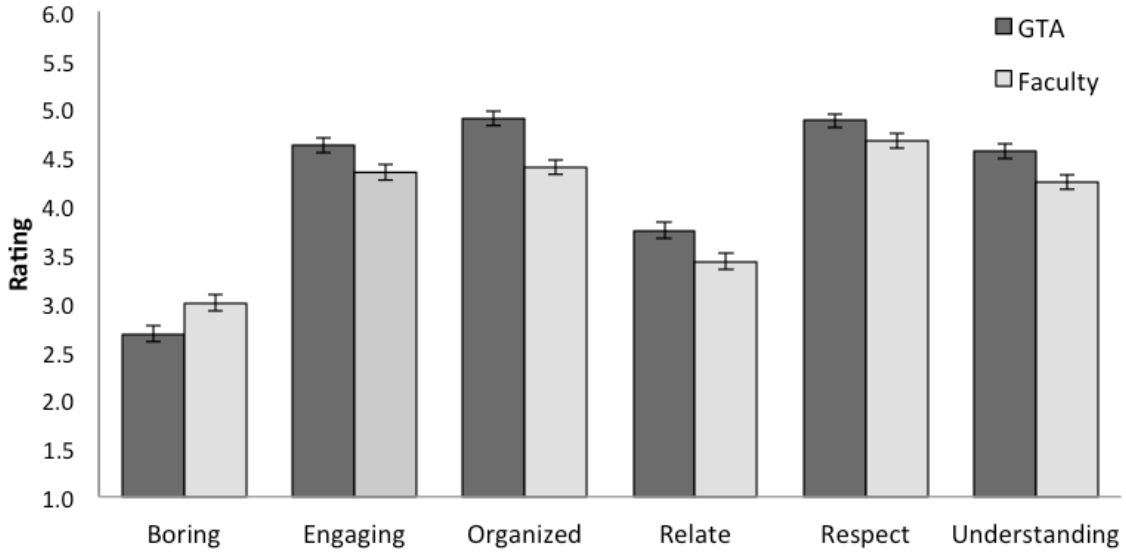

Figure 2. Mean ± standard error for descriptors (in alphabetical order) in which GTAs differed significantly from faculty members for final ratings (Wilcoxon signed-rank tests [$\alpha = 0.005$]).

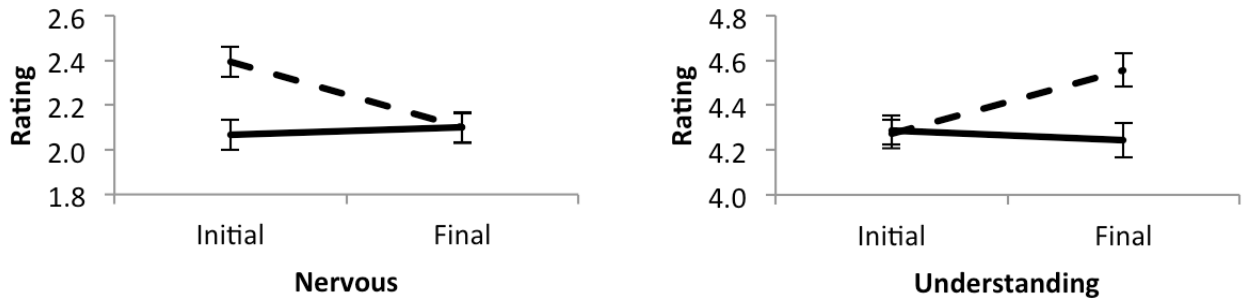

Figure 3. Descriptors (in alphabetical order) in which significant changes in mean student ratings (from 1 – 6 +/- standard error) were found over the semester for GTAs only. GTAs are depicted with dashed lines and Faculty with solid lines.

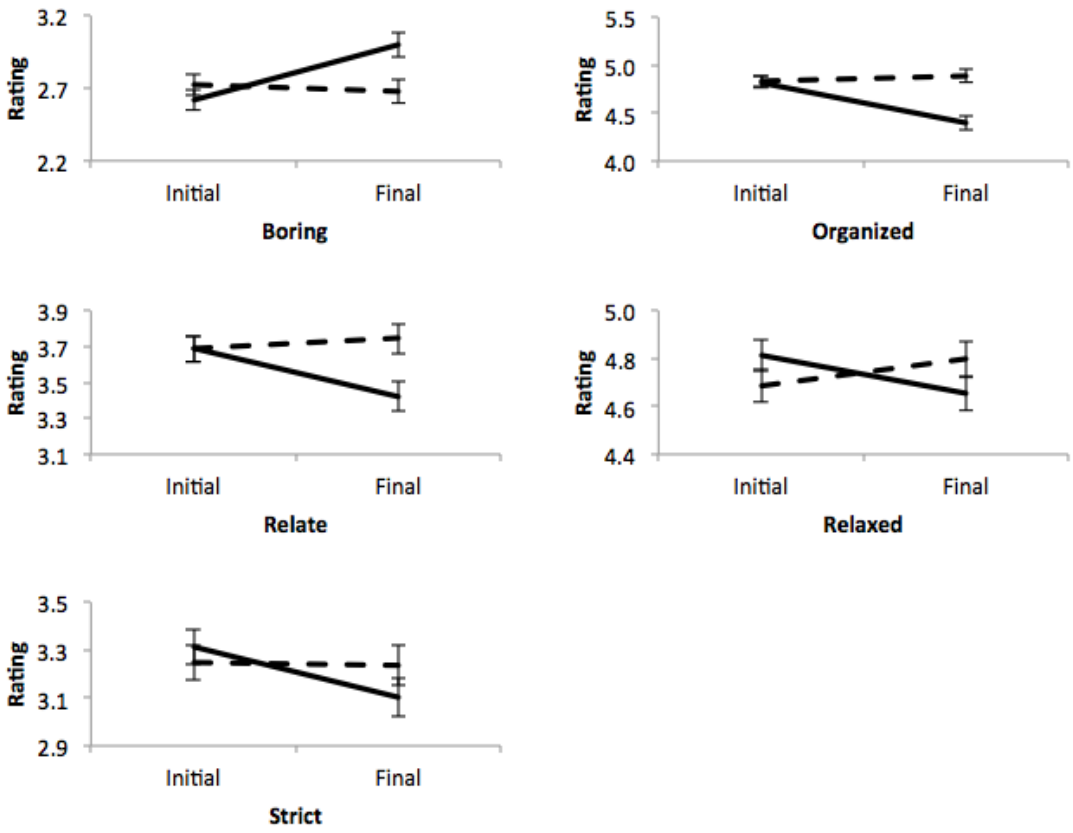

Figure 4. Descriptors (in alphabetical order) in which significant changes in mean student ratings (from 1 – 6 +/- standard error) were found over the semester for Faculty only. GTAs are depicted with dashed lines and Faculty with solid lines.

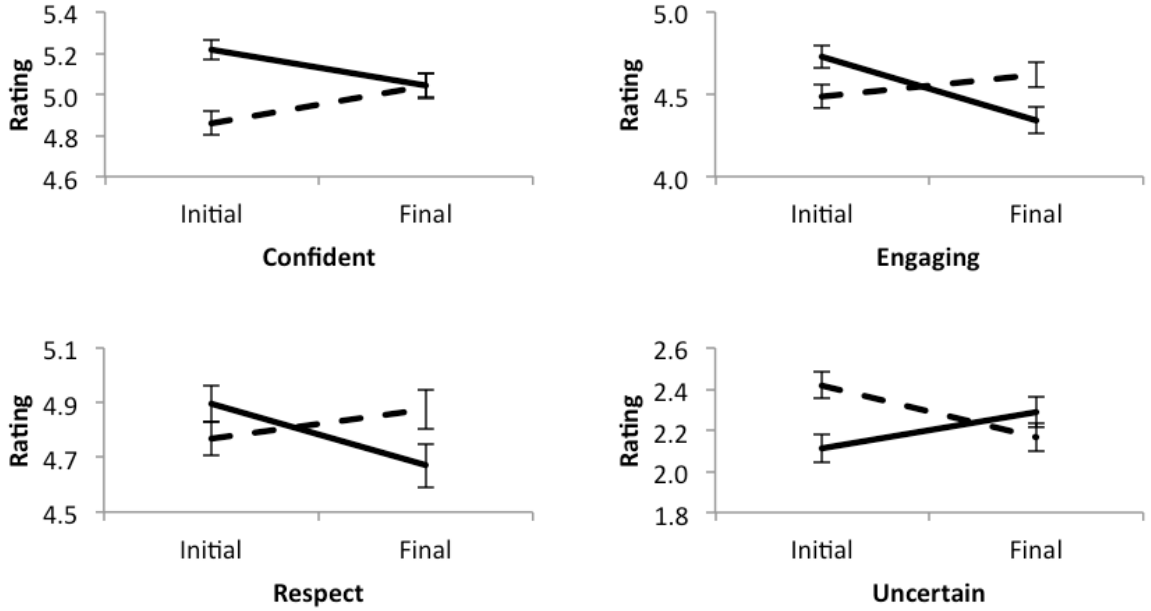

Figure 5. Descriptors (in alphabetical order) in which significant changes in mean student ratings (from 1 – 6 +/- standard error) were found over the semester for GTAs and Faculty. GTAs are depicted with dashed lines and Faculty with solid lines.

Table 9. Mean ± standard error for all initial ratings of descriptors (in alphabetical order) that had significant associations with (A) teaching experience, (B) instructor gender, or (C) instructor age (Spearman rank correlations, $\alpha = 0.05$).

A. Teaching Experience

	0-1 semesters	2-4 semesters	5-8 semesters	9+ semesters
Confident	4.83±0.13	4.68±0.09	5.00±0.07	5.34±0.05
Engaging	4.49±0.14	4.40±0.10	4.54±0.09	4.82±0.08
Enthusiastic	4.46±0.14	4.31±0.10	4.39±0.09	4.72±0.08
Nervous	2.55±0.15	2.50±0.10	2.19±0.09	1.99±0.08
Relaxed	4.69±0.14	4.49±0.10	4.72±0.09	4.94±0.08
Strict	3.26±0.15	3.17±0.12	3.17±0.09	3.44±0.09
Uncertain	2.45±0.14	2.47±0.10	2.31±0.09	2.05±0.08

B. Instructor Gender

	Female	Male
Organized	4.94±0.06	4.70±0.06
Uncertain	2.36±0.06	2.16±0.07

C. Instructor Age

	Under 25	26-30	31-40	41-50	51-60
Confident	4.73±0.11	4.91±0.08	4.86±0.10	5.32±0.06	5.22±0.08
Distant	2.85±0.15	2.43±0.09	2.23±0.11	2.27±0.08	2.45±0.14
Engaging	4.37±0.13	4.47±0.10	4.72±0.11	4.76±0.09	4.68±0.13
Enthusiastic	4.00±0.13	4.36±0.09	4.66±0.11	4.77±0.09	4.57±0.13
Nervous	2.79±0.14	2.23±0.08	2.31±0.11	2.09±0.09	1.90±0.12
Relaxed	4.63±0.13	4.66±0.09	4.79±0.11	4.66±0.09	5.10±0.10
Uncertain	2.74±0.14	2.32±0.09	2.26±0.10	2.08±0.09	2.07±0.13

Table 10. Mean ± standard error for all final ratings of descriptors (in alphabetical order) that had significant associations with (A) teaching experience, (B) instructor gender, or (C) instructor age in final ratings (Spearman rank correlations, $\alpha = 0.05$).

A. Teaching Experience

	0-1 semesters	2-4 semesters	5-8 semesters	9+ semesters
Confident	4.68±0.12	4.85±0.10	5.21±0.66	5.14±0.06
Nervous	2.38±0.13	2.36±0.11	1.91±0.08	1.99±0.07
Organized	4.63±0.15	4.73±0.10	4.88±0.09	4.41±0.09
Strict	3.38±0.17	2.89±0.13	2.98±0.10	3.41±0.10
Uncertain	2.66±0.14	2.39±0.12	1.94±0.08	2.20±0.08
Understanding	4.35±0.16	4.67±0.11	4.59±0.09	4.11±0.09

B. Instructor Gender

	Female	Male
Organized	4.79±0.07	4.47±0.08
Understanding	4.51±0.08	4.28±0.08

C. Instructor Age

	Under 25	26-30	31-40	41-50	51-60
Boring	2.91±0.14	2.58±0.11	2.67±0.16	2.98±0.11	3.12±0.15
Engaging	4.40±0.14	4.77±0.10	4.60±0.14	4.25±0.11	4.32±0.13
Organized	4.99±0.10	4.99±0.09	4.68±0.12	4.18±0.12	4.51±0.12
Relate	3.61±0.15	3.79±0.12	4.00±0.16	3.19±0.11	3.43±0.14
Understanding	4.39±0.14	4.61±0.10	5.01±0.11	3.82±0.10	4.40±0.12

Table 11. Significant Spearman ($\alpha = 0.05$) associations of descriptors with instructor demographics for initial and final ratings. Non bold descriptors have weak associations, italic descriptors have moderate associations, and bold descriptors remain significant throughout the semester.

	Initial Ratings	Final Ratings
Teaching Experience	**Confident** r = 0.269 p = 0.000	**Confident** r = 0.148 p = 0.001
	Engaging r = 0.154 p = 0.001	**Nervous** r = -0.148 p = 0.001
	Enthusiastic r = 0.414 p = 0.001	Organized r = -0.089 p = 0.044
	Nervous r = -0.224 p = 0.000	**Strict** r = 0.097 p = 0.029
	Relaxed r = 0.145 p = 0.001	**Uncertain** r = -0.122 p = 0.006
	Strict r = 0.095 p = 0.031	Understanding r = -0.150 p = 0.001
	Uncertain r = -0.193 p = 0.000	

	Initial Ratings	Final Ratings
Age	Confident r = 0.222 p = 0.000	Boring r = 0.109 p = 0.014
	Distant r = -0.112 p = 0.012	**Engaging** r = -0.107 p = 0.015
	Engaging r = 0.128 p = 0.004	Organized r = -0.227 p = 0.000
	Enthusiastic r = 0.202 p = 0.000	Relate r = -0.134 p = 0.003
	Nervous r = -0.230 p = 0.000	Understanding r = -0.149 p = 0.001
	Relaxed r = 0.103 p = 0.020	
	Uncertain r = -0.214 p = 0.000	

Gender	**Organized** r = -0.153 p = 0.001	**Organized** r = -0.129 p = 0.004
	Uncertain r = -0.100 p = 0.024	Understanding r = -0.107 p = 0.015

CHAPTER THREE

More than words: Probing the terms undergraduate students use to describe their instructors

The following chapter is in press:

Kendall KD, Schussler EE (2013). More than words: Probing the terms undergraduate students use to describe their instructors. Int J Teach and Learn in Higher Educa *17*.

Abstract

Undergraduates often use single words to describe their instructors, including “boring,” “enthusiastic,” or “organized,” but what instructional behaviors cause students to use these words? This study utilized interviews and an online survey to ask students to translate commonly used instructor descriptions into their associated instructional behaviors. Explanations from undergraduates revealed that these behaviors include both the way instructors interact personally with students as well as how they convey the content to the class. The instructional behaviors identified through this process were classified into themes representing three different aspects of classroom interactions and placed into a schema depicting the positive instructional behaviors that students indicated led to respect for an instructor. The results will give instructors insight into how what they do in the classroom relates to student perception of their ability to foster learning.

Keywords: Instructor, Behavior, University, Teaching, Learning

Introduction

Instructors have a dramatic impact on student learning. This is because student-instructor interactions impact not only student achievement, but also student attitude and the overall culture of the educational environment (Halandya et al., 1982; Koballa and Crawley, 1985; Osborne et al., 2003; Umbach and Wawrzynski, 2005). Since instructor influences can be beneficial or detrimental to student learning (Koballa and Crawley, 1985), Nussbaum (1992) suggested that understanding how instructional behaviors can differentiate effective instructors from ineffective instructors will be valuable for higher education research. For the purposes of this study, instructional behaviors refer to how instructors act or conduct themselves when interacting with students. This could encompass everything from how instructors teach (e.g., they pace and speak quickly), to interactions before and after class (e.g., they smile and laugh as they wander the room), to interchanges during office hours (e.g., they act preoccupied and are intimidating). It is through the summation of these interactions that students judge the effectiveness of their instructors in fostering learning.

Instructional behaviors should be identified from the perspective of students, because students are in the most ideal position to evaluate the instruction they are receiving in relation to their perception of learning (Zabaleta, 2007). Instructors, for instance, may believe a certain behavior is beneficial to student learning, while students may perceive this behavior as either a non-factor or even counterproductive to their learning (Feldman, 1988; Helterbran, 2008). Once instructional behaviors that impact student learning have been identified, such as the impact of enthusiasm on student attitude (Halandya et al., 1982), suggestions to improve teaching can be made in terms of modifying specific instructional behaviors (Nussbaum, 1992).

Instructor Descriptions

The purpose of this study was to explore what instructional behaviors undergraduates at one institution, a large research university, associate with words used to describe their instructors. The descriptive words were identified from a study of student perceptions of graduate teaching assistants (GTAs) and professors at the same institution (Kendall and Schussler, 2012). In this study, selected sub-scales from the Questionnaire on Teacher Interaction (QTI) (Coll, Taylor, and Fisher, 2002) and College and University Classroom Environment Inventory (CUCEI) (Treagust and Fraser, 1986; Coll et al, 2002) surveys evaluating interactions between individuals and classroom environment were used to identify whether student perceptions of GTAs and faculty members differed (Kendall and Schussler, 2012). Open-ended student responses supplemented these survey results and led to sets of descriptive words, organized into themes, that delineated GTAs from professors (Table 12); these differences are also supported by Dudley (2009), Muzaka (2009), and Park (2002). The Authors chose a sub-sample of these descriptive words (shown in bold font in Table 12) to be used as probes in the current study.

While student perception of instructors may be based on comparisons of current instructors to past instructors, there may also be some instructor perceptions that are based on instructor stereotypes. This indicates that there may be certain descriptive words and associated instructional behaviors that students are more likely to ascribe to different subsets of instructors. Therefore, using the descriptive words generated from this comparison between GTAs and professors should provide a diversity of instructional behaviors that may be present in university instructors. It may be true for instance that a student would never call a professor nervous, so that descriptive word would never be investigated if the focus was only on what students said about

professors. This study asked students to explain what an instructor would have to do in a classroom to be labeled with each descriptive word, regardless of what instructor type the descriptive word was ascribed to.

Throughout the article, we use the term “descriptions” (or “descriptors”) to refer to the words generated from Kendall and Schussler (2012) used in this study (e.g., “boring,” “enthusiastic,” or “organized”). In other studies, these may also be referred to as instructor characteristics, traits, or qualities.

Instructional Behaviors

The descriptive words undergraduate students use to convey their perceptions about their instructors are more than just off-the-cuff remarks; they are summaries of complex classroom behaviors that are seldom probed in the college teaching literature. This gap was addressed by Varca and Pattison (2001) and Pattison, Hale, and Gowens (2011) who studied the instructional behaviors that contributed to student perception of teaching performance.

Varca and Pattison (2001) implemented a “critical incident technique,” previously used in business settings, to identify instructional behaviors that contributed to student perception of excellent and poor teaching performance. This study was conducted in five different courses (student n = 252) at the University of Wyoming by asking undergraduate students to list two positive incidents (excellent teacher) or two negative incidents (poor teacher).

The results identified four critical performance dimensions associated with excellent teaching (administration, classroom delivery, student interactions, and teacher motivation), with student interaction being an underlying thread of all dimensions (Varca and Pattison, 2001). Because it was possible that the student interactions result was merely a by-product of

professional caring about students, Pattison et al. (2011) performed a follow-up study to determine more precisely what instructional behaviors students were referring to as critical.

Pattison et al. (2011) again implemented a "critical incident technique" with undergraduate and graduate students enrolled in 28 business courses at a large state university. They found that students perceived almost all instructional behaviors as an indication of the instructor caring and respecting the students. The instructional behaviors displayed by outstanding instructors included being knowledgeable, making material relevant to students, being prepared and organized, having a plan, integrating material from various courses, having well organized courses, and encouraging as well as answering student questions (Pattison et al., 2011). In contrast, ineffective instructors were tougher than necessary, did not grade impartially, did not return assignments or provide constructive feedback, were unprepared and unorganized, and impatient with students (Pattison et al., 2011). Students tended to be more satisfied with the educational environment when instructors were respectful of students, saw students as individuals, and appeared to care about them (Pattison et al., 2011); likewise, Hawk and Lyons (2008) documented that when an instructor lacked caring or respect for students, students reported feeling as if the instructor had "given up" on their learning.

Garko et al. (1994) also documented instructor behaviors through open-ended response surveys asking students what they desire in student-teacher relationships. Student responses emphasized classroom behaviors such as connecting with students, being open to discussion, and providing feedback, as positive for instructional relationships. Students also expressed the importance of familiarity, respect, availability, equality, knowledge, and caring (Garko et al., 1994). Finally, indicating a desire for mutual respect, they found that if the instructor modeled

attention and energy when conducting classroom activities, students reciprocated with attention and energy (Garko et al., 1994).

Student Evaluations of Teaching

University instructors are often provided feedback regarding their instruction by means of student evaluations of teaching. These evaluations can be used to support employment decisions (e.g., promotion or termination) by the university administration; however, they also provide insightful feedback regarding effective instruction and student learning from the student perspective (Baird, 1987; Emery et al., 2003; Clayson et al., 2006; Wright, 2006; Zabaleta, 2007; Helterbran, 2008; Kogan et al., 2010). While student evaluations of teaching have been criticized due to their anonymity (Wright, 2006; Kogan et al., 2010), students are some of the most appropriate authorities when it comes to judging the classroom teaching and learning environment (Clayson et al., 2006; Zabaleta, 2007).

Student evaluations often provide instructors with descriptive statistics (e.g., mean, median, standard deviation) of student ratings for a specified list of likert-scale items. Some evaluations also include student written responses solicited via open-ended response questions. This format provides instructors and university administrators an overview of student perception regarding the instruction and learning throughout the course; however, these responses also lack context (such as the rigor of the course or student preparation), which may impact student responses (Wright, 2006). Students may also interpret the likert-rating scale options differently (Huck, 2008) or have different interpretations of the same question (Rubin and Rubin, 2012). Therefore, Wright (2006) recommended completing in-depth interviews with a sub-sample of students, in addition to student evaluations of instruction, to clarify the circumstances leading to

the rating of the instructor (also noted in Rubin and Rubin, 2012, as reason for in-depth interviews).

Project Rationale

At most universities, it is impractical to conduct interviews of students as part of the student evaluation process. Instead, student open-ended comments may provide additional insight into student perspective if the instructor can accurately infer the meaning of the student comments (Feldman, 1988; Helterbran, 2008). For example, a student saying that the instructor was "strict" may be referring to grading in the course, but could also be referring to their standard for classroom behavior. A better understanding of the meaning of student descriptive words may improve instructor interpretation of student feedback and empower them to modify the behaviors that most significantly affect instruction (Varca and Pattison, 2001; Umbach and Wawrzynski, 2005).

This study proposes to identify instructional behaviors associated with commonly used instructor descriptions from the perspective of students enrolled in introductory biology courses at a large Southeastern research university in the United States. Undergraduate volunteers were invited to be interviewed or to complete an online survey to provide explanations of descriptive words for instructors. These data were analyzed to identify the consensus instructional behaviors that students indicated would lead them to ascribe a particular word (such as "boring") to an instructor. Although it may be possible that different groups of students (e.g., genders) have different behavioral descriptions for the same word, it was not the intent of this study to parse those differences; nor is there evidence that these differences exist in this population (Kendall and Schussler, 2012). Therefore, the goal was to obtain *consensus* behaviors associated with the descriptors that could be used as general guides for interpreting student evaluations of teaching.

The instructional behaviors identified in the study were then organized by the researchers into a schema that summarizes the behaviors that students indicated supported their classroom learning.

Materials and Methods

Data Collection

Qualitative data in the form of written and oral explanations for each descriptive word were collected by means of interviews and an online survey with undergraduate students enrolled in general biology courses in spring 2011 at a large research institution in the Southeastern United States. Students were recruited from a total of nine classes: three second semester introductory biology classes, two biodiversity classes, one cellular biology class, two genetics classes, and one ecology class. Each lecture class seated anywhere from 32 – 225 students.

Interviews. Participants were recruited by means of emails distributed to the class by the lecture instructor on behalf of the researcher, as well as in class announcements by the researcher in February and March 2011 (one in-person visit per class). The recruitment notices and announcements included information regarding the purpose of the interviews, contact information for the researcher, time frame, and nature of the interviews. Students were also assured that the interviews were not related to the course and that their instructor would not be involved in, or see the data. This recruitment yielded 24 undergraduate students who participated voluntarily in interviews.

The interviews were conducted from the 21st of March until the 21st of April solely by the first author of this paper. All interviews were conducted in a research laboratory at the university and were audiotaped with the permission of the participant. Each interview ranged in length from 17 to 37 minutes and a university bookstore gift card valued at $10.00 was offered to each participant to compensate them for their time. The interviews were transcribed verbatim by the

interviewer (Express Scribe v. 5.20), and at that time each participant was assigned a pseudonym. All procedures were reviewed and approved by the Institutional Review Board for Human Subjects.

Interviewees were informed that the purpose of the interview was to explore undergraduate perception of instructors and how instructors impact their learning. The interviews explored the following specific instructor descriptions, one at a time: boring, confident, distant, engaging, enthusiastic, nervous, organized, relate, relaxed, respect, strict, uncertain, and understanding (Kendall and Schussler, 2012). After collecting demographic information, the format of the interview was the same for each descriptive word; the interviewer asked the student for the definition of the word, and then an example of how it related to an instructor. For instance, interviewees were asked, "How would you define confidence? Can you provide me with some specific examples of how your instructor is/is not confident?" The objective of the interview was to prompt students to describe classroom behaviors that would lead them to use a particular descriptive word for that instructor.

Online survey. The last week of classes in spring 2011 an online survey (hosted by surveymonkey.com) was distributed via an email from the lecture instructor on behalf of the researcher to undergraduates in the same general biology courses. Two of the original instructors, however, did not send the survey to their students (one genetics and the ecology class). The email contained information regarding the purpose of the study, which was the same as the interviews, as well as consent information and a link to the survey.

The online survey was a mechanism to allow undergraduates who did not participate in an interview the opportunity to provide their input on the meaning of the instructor descriptions. The online survey used the same descriptive words, and asked undergraduates to provide a

definition for each term based on its use in a simple sentence (e.g., “this instructor is understanding,” “this instructor is boring,” etc.). Forty-two undergraduates anonymously completed this online survey. No incentives were offered for participation in this survey, and all procedures were reviewed and approved by the Institutional Review Board for Human Subjects. Since the same student population was invited to participate in interviews and anonymous online surveys, the possibility of double respondents exists in this study; however, the researchers will present evidence in the results section as to why this likely did not impact the results of this study.

Data Analysis

Interviews. Transcribed interviews were independently analyzed by three researchers (the authors of this article as well as a trained assistant) who read and re-read the student responses for each descriptive word and took notes on the behaviors that the students were describing. Each researcher then independently decided on the participant consensus for each descriptive word. Two researchers worked on each descriptive word: the first author completed all of the words, and the second author or the trained assistant provided the second consensus for each word. After this independent work, the research pairs for each descriptive word merged their interpretations via discussion to reach a final list of consensus instructional behaviors.

Student explanations of each descriptive word were remarkably consistent, and there was little negotiation necessary between the researchers to reach a consensus definition. For instance, the research assistant determined that students defined “understanding” as having compassion about student issues, being willing to bend the rules, accepting late work if the student has a problem, being willing to listen and help students, willing to make adjustments to student feedback, and remembering what it is like to be an undergraduate. The first author listed student

definitions of “understanding” as empathetic and compassionate, willing to help students, being in the same shoes as students or having had similar experiences, not following the rules strictly, and extending deadlines. Discussion led to the final definition of “understanding” as compassionate, flexible, empathic, and knowing what it is like to be a student. The first author then pulled representative quotes from interviews to support the results. All quotes use pseudonyms for the participants. This analysis is similar to that outlined in LeCompte (2000).

Online survey. Online survey responses were analyzed in the same fashion as the responses from interviews. Responses for each descriptive word were independently analyzed by two researchers (the authors of this article) who read and re-read the responses and took notes on how undergraduates defined the descriptive words. Once again, each researcher compiled their own findings and only then did they compare results with each other. Final behavioral descriptions for each descriptive word were determined when both researchers were in agreement. The first author then chose representative quotes from the online surveys (which were also given pseudonyms) to support each descriptive word.

Combined results. Comparison of the online survey results and the interview results revealed that both methods of data collection resulted in similar findings regarding instructional behaviors students thought were associated with specific instructor descriptions. Thus, both sets of results were combined and are reported together in the results.

Compilation of the Descriptions and Schema Development. Once the analyses were completed, the authors noted that some descriptions appeared to be the opposites of each other (e.g., the descriptions of boring and engaging). Therefore, the student explanations were used to sort the descriptive words into positive and negative counterparts of the same behaviors. The positive terms were then grouped into themes that represented different aspects of instructional

proficiency. Finally, to highlight the relationship between these themes, a schema was created that represented how the positive behaviors support classroom learning.

Results

Participants

Interviews. Undergraduates (n = 24) who participated in the interviews were mostly freshman (first year; 54%), non-biology major (71%), female (75%), native English speakers (100%). Second and third year students comprised 21% and 13%, respectively, of the respondents, with 12% more being fourth year or beyond. Twenty-nine percent of the students were biology majors, with 4% concentrating in Ecology and Evolutionary Biology, 17% in Biochemistry & Cellular and Molecular Biology, and 8% in Microbiology. The majority of participants had completed another biology course (54%); 38% had completed one other biology course, while 8% had completed two, 4% had completed three, and 4% had completed four. Twenty-one percent of participants were currently enrolled in second semester introductory biology, 42% in biodiversity, 8% in cellular biology, and 29% in genetics. The majority of participants were Caucasian (75%), while the remaining participants were African-American (13%), Hispanic (4%), Asian-Caucasian (4%), and Egyptian-American (4%).

Online survey. Forty-two undergraduate students completed the online survey; however, demographics were only obtained for 40 of these undergraduates. These undergraduate students were mostly freshman (first year; 68%), non-biology major (73%), female (70%), native English speakers (98%). Second and third year students comprised 23% and 3% of the respondents, respectively, with an additional 8% being fourth year or beyond. Twenty-eight percent of the students were biology majors, with 13% having a concentration in Biochemistry, Cellular, and Molecular Biology and 8% being Microbiology. The majority of respondents had completed

another biology course with a lab (65%), with 58% having completed one other biology course, 5% two courses, and 3% three courses. At the time of the survey, 45% were enrolled in second semester introductory biology, 30% in biodiversity, 13% in cellular biology, and 13% in genetics. Comparison of undergraduate student demographics for those who completed the online survey and those who participated in interviews revealed that of these 42 students there is a possibility that four may have completed both the interview and the survey. However, since behaviors were identified only via strong consensus among participants, it is unlikely that this potential duplication influenced the results of the study.

Student Explication of Instructor Descriptions

The results are presented by the descriptive word being explicated, in alphabetical order. For each descriptive word, the behaviors that lead to that description, as identified by students, are presented, and then supported by quotes from the interviews or online surveys. A summary of the overall results are found in Tables 13 and 14, in which the definitions for each descriptive word are sorted by the authors' inferred relationships between the words. Table 13 groups the descriptors into positive and negative ends of the instructional spectrum while Table 14 depicts the descriptors which had no opposites: organized and respect. At the end of the results, a proposed schema groups the positive behavioral descriptions into three overall themes, and shows how these may be related to student respect for an instructor.

Boring. Undergraduate students indicated that boring instructors are monotone, not engaging or interactive in teaching, and do not seem personally enthusiastic about the course material. For example, Cassia stated that they "*are just monotone and I feel like I should be in Ferris Bueller's Day Off*" (Interview). Robin explained that when instructors are boring she is "*falling asleep in your class (laughs), it's not engaging or interesting to me*" (Interview). Finally,

Olivia said, "*the instructor cannot grab the attention of the students*" (Online survey) when he or she is boring.

Interviewees further indicated that these instructors would not come down to the student level or adequately explain material, thus students are unable to understand or keep up with the information being taught. For example, Hannah affirmed, "*they leave out big chunks of knowledge that they know already but the students do not know it*" (Interview). Further, George said, "*they'll just read the slides, not really do any actual teaching because they just read to you,*" (Interview) which he indicated was boring.

Confident. Undergraduates indicated that confident instructors know the material and information. This was exemplified by Tristan who said, "*the instructor knows what he or she is talking about and is knowledgeable on course content*" (Online survey). Due to their knowledge, confident instructors are able to answer student questions and are sure of their teaching. This was explained by Hannah who stated, "*they can answer kind of obscure questions that students have, sometimes you just come up with crazy things and when they are able to answer them you are like 'wow they have a really in depth knowledge of it'*" (Interview). Not only do confident instructors know the material, but Megan said the instructor can "*clearly teach the subject*" (Online survey) and Rachel said they are able to "*look at the information from all different angles and multiple different ways*" (Interview), thus indicating that confident instructors are able to explain concepts to students and are good instructors.

Interviewees also indicated that confident instructors are typically prepared for class and are calm. This is exemplified by Kim who stated that confidence is "*being prepared, ready to talk and knowing what to say and not stuttering through everything and just flowing through it*

and making it in a way that's understandable for us" (Interview). This conveys that confident instructors are effective instructors.

Distant. Distant instructors are not approachable, nor are they engaging or interactive in the classroom. Melanie explained this by stating that distance is "*not approachable, not engaging*" (Interview), Kayla noted that the instructor "*does not relate to the students at all*," (Online Survey), and Donna added, "*you are really intimidated*" (Interview). Distant instructors do not care about the students as shown by their unavailability for help, and their appearance that their mind is not on their teaching. Grace explained that they are "*not really in the classroom, thinking about something else and not the material and lecture*" (Online Survey). Lola indicated that, "*she is not available to answer questions... she does not answer the emails quickly*" (Online Survey). These definitions indicate that distant is the opposite of approachability and availability.

Engaging. Undergraduates state that engaging instructors are able to keep the attention of students because they are interactive and involved in the classroom. For example, Emma said, "*he interacts with the class. He asks questions and expects a response*" (Online survey). Engaging instructors typically have a passion for teaching and bring in examples from real life for students. This is explained by Donna who said, "*she brings her samples into class from her own research*" (Interview). Taylor described an engaging teacher as one who can "*grasp the attention of all students regardless of the subject matter. Even if the subject matter is deemed boring, the instructor will be able to engage students in learning and discussion*" (Online survey). Students indicated that being interactive and involved in the classroom are important instructional behaviors for being engaging.

Interviewees further indicated that engaging instructors communicate on the student level and speak to students as individuals. For instance, Sally said, "*being able to have that kind of*

communication with students so it doesn't just go over their heads all the time" (Interview). Melanie agreed, saying, "*being able to answer one-on-one questions, ... and asking feedback from the class and things like that*" (Interview). In these cases, students indicated that engaging can be at the classroom teaching level, but also at the personal level.

Enthusiastic. Instructors who are enthusiastic are excited to teach, have a passion for the subject, and as a result they enjoy teaching. Noah described it as "*very passionate about what they teach and genuinely interested in subject and whether or not students learn*" (Online survey). These instructors are interested in and care about their students and they are happy to be in the classroom. For example, Gabriella said, "*the tone of voice the teacher uses while teaching. He or she sounds happy to be there, happy to be teaching the material and gladly answers questions*" (Online survey). Undergraduates believe that enthusiastic instructors want their students to learn and do well in the course. Interviewees such as Jessica stated that her instructor "*uses examples and you can tell how interested she is in the topic*" (Interview). Thus, enthusiastic instructors engage their students in additional examples that make the material relevant to students while showing them their passion for the subject.

Nervous. When undergraduates use the term nervous they typically mean the instructor is not confident in his or her teaching or with the material. Samantha explained, "*the teacher is not confident in their portrayal of class material and subject matter*" (Online survey). Undergraduate students also indicated that a nervous instructor will act uncomfortable and may be shaky and sweaty. Emma described this as "*he is probably sweating profusely, shaking, and stutters or has a hard time saying something in the front of the class*" (Online survey). Undergraduates expressed that nervous instructors are not good teachers. As stated by George, "*A nervous person should not be a teacher I feel*" (Interview). When prompted to further explain how nervous

instructors impact his learning, George explained that, "*it would be a negative impact because you think to yourself does this person really know what they are talking about*" (Interview), thus indicating that nervous teachers make students doubt their subject matter knowledge.

Organized. Instructors who are organized have everything ready to go for class. This is explained by Hannah who said, "c*oming in and having everything ready to go when class starts and not fiddling with technology and corrupted files and stuff, they have already taken care of that before*" (Interview). Amy explained organized as, "*knowing exactly what you are going to teach and exactly how you are going to explain it instead of say flipping back and forth between slides because you forgot which one was next"* (Interview), indicating that the instructor has thought through the lesson and have their courses planned out. Amy said, "*she has a plan every week when we go in there and we follow the plan just exactly*" (Interview). Additionally, Kaitlyn said, "*they grade assignments and tests within a timely manner and do not lose any of them*" (Online survey) indicating that organized instructors do not lose student work and promptly grade assignments.

Relate. Relatable instructors are able to connect with students because they have common interests with students and thus students feel as if these instructors are approachable, unintimidating, and accommodating. Ella said, "*there is an understanding between the instructor and the student*" (Online survey), and Jenna explained, "*relating to the instructor means that you can find some common ground and a way to communicate with each other in a comfortable manner.*" (Online survey). As a result of being able to relate, Lindsey states, "*I can easily go and talk to her about my concerns*" (Online survey). By being relatable, these instructors understand how to present the material. Kelsey explains this as, "*the stories or anecdotes they use to make*

the material engaging is something I can understand or relate to in some way" (Online survey), implying a positive effect on learning.

Relaxed. Undergraduate students indicated that relaxed instructors are not nervous or anxious; instead they are comfortable and laid-back because they are prepared for class and confident. Megan revealed, "*the teacher is confident and not afraid to be in front of the class*" (Online survey). Alina commented that, "*the instructor is just talking to you about things that you need to know. They aren't fidgety or walking around and pacing constantly*" (Online survey). Elijah merely suggested they are "*well prepared*" (Online survey). Relaxed instructors are typically lenient about rules and students feel they are approachable. Jenna elaborated on this by saying, "*the instructor is not intimidating or overwhelming, but is 'laid back,' approachable, and willing to help students*" (Online survey), and Chloe added, "*they are understanding if a student needs to turn something in late and has a valid reason*" (Online survey), indicating that relaxed instructors are flexible.

Respect. Instructors earn respect from students through being able to answer questions, and making sure students understand the material. Ethan stated that the "*instructor has taken time to answer questions outside of class and [is] willing to help*" (Online survey). Chloe elaborated that, "*respect means that the student feels that the instructor is a person who obviously loves their job and is more concerned with the students than themselves*" (Online survey). When students respect instructors they typically trust them and will follow their rules and listen to them. Jessica said that with an instructor she respects she is "*polite in class by not talking and making sure that you get your homework done. Don't cause any extra stress, that kind of thing*" (Interview). Respect can be earned, as explained by Megan:

he/she gives you respect, they know what they're talking about and they help you throughout the semester with questions. They earn your respect by their actions in front of the class and to you personally. I completely lose respect when the professors completely disregard your feelings maybe about a bad grade or they use curse words in class. (Online survey)

While respect may be earned, Ted explained, "*I'm just the kind of person that feels like [they are] authority figures, I'm going to treat them with the same respect*" (Interview), so some students will respect an instructor merely due to their status as an authority figure.

Strict. Strict instructors adhere to the rules and policies without flexibility and as a result are not tolerant of bad behavior or distractions in class. For example, Arya said the instructor "*won't allow for talking, foolery, joking"* (Online survey). Mia explained, "*she is not understanding to the various situations and dilemmas a student may have*" (Online survey), while Michael commented, "*the instructor will not budge on the schedule or work assignments*" (Online survey). Caleb explained, "*they harshly enforce assignments, have little toleration for distractions and grades on a difficult level*" (Online survey), indicating that undergraduates also consider strict instructors to be tough graders.

Uncertain. Similar to nervous instructors, undergraduates express that uncertain instructors are not confident about the material. Noah described it as, *"the opposite of confident, they are unsure of their abilities as a teacher or they are unsure of their knowledge in the subject*" (Online survey). As a result, uncertain instructors are not able to answer student questions nor are they sure how to teach. Logan explained that they are "*Indecisive. The instructor does not have a full enough understanding of the subject matter to convey a certainty in the information they are presenting*" (Online survey). Julia added that the instructor "*doesn't*

know if what she is teaching is exactly correct, can't answer all of the students questions" (Online survey), revealing a perceived lack of content knowledge.

Understanding. Understanding instructors are compassionate when dealing with student issues (personal as well as academic); they are typically empathetic. As explained by Rachel, "*she takes each students situation into account when they come to her with a problem and is willing to work with them based on their individual needs*" (Online survey). Adrianna said, "*the instructor has empathy with the students*" (Online survey). Marie said, "*the instructor knows what it is like to be in the students shoes and not every student catches on quickly to the subject*" (Online survey), indicating that understanding often stems from the instructor knowing what it is like to be a student. Specifically, interviewees noted that understanding instructors will extend deadlines for students. Wendy said, "*I'd say they are more likely to let you turn stuff in late,*" indicating they are flexible.

Themes and Schema

The researchers first organized the positive explanations shown in Table 13 into three themes: "teaching techniques," "interpersonal rapport," and "passion for subject." The teaching techniques theme reflects the behaviors of the instructors as teachers, specifically how they deliver the course material to students (e.g., use of examples, being interactive, and being calm). The interpersonal rapport theme features person to person behaviors, such as the comfort level and understanding between students and instructor (e.g., compassion, ability to relate, and approachability). Lastly, the passion for subject theme incorporates factors related to instructor knowledge and desire to teach that subject, such as their confidence in the material, enjoyment teaching that subject, and ability to answer questions about the topic. Based on student descriptions, and the lack of an opposite word for the term "respect," it was hypothesized that

respect is an outcome of good instructional practices. Therefore, the three themes are represented as leading to the overarching description of "respect," which encompasses the instructor behaviors that students indicate earns their trust, fosters their learning, and makes them more willing to engage in the class (Figure 6).

Discussion

This study identified the instructional behaviors, linked to specific descriptive words, which undergraduates at one university indicated led to respect for their instructors. Undergraduates in this study emphasized that instructors must earn their respect through three themes: teaching techniques, interpersonal rapport, and passion for subject. This mirrors the previous findings of Varca and Pattison (2001) that there are several dimensions to good teaching. Although the themes identified in this study are slightly different from those identified in Kendall and Schussler (2012), the schema from this study also reflects the findings of several studies emphasizing the importance of classroom instructional behaviors such as open communication, respect for students, caring about students as individuals, and being knowledgeable (Garko et al., 1994; Teven, 2007; Hawk and Lyons, 2008; Pattison et al., 2011).

When it came to excellent instructors, students participating in this study demanded more than just good teaching; they also put a premium on interpersonal behaviors that indicated care and concern for students. In fact, students in this study often indicated that there were two layers of behaviors for each descriptive word: personal and instructional (also seen in Arnon and Reichel, 2007). For example, the description for "relate" was explained as an instructor having something in common with an individual student, but also as being able to communicate information at the knowledge level of a class. The descriptive word "relaxed" indicated an

instructor's classroom persona while teaching, but students also indicated that this word meant that they were approachable on a personal level.

Changing Instructional Behaviors

The explanations provided by the undergraduates in this study have provided insight into the potential instructional behaviors that might lead students to assign certain descriptive words to their instructors. As suggested by Nussbaum (1992), this allows inferences to be made about the instructional behaviors that are associated with effective and ineffective instruction, from the perspective of students. Thus, instructors could potentially use the words that students ascribe to them on student evaluations to better understand what instructional behaviors they might be able to modify to convey different impressions (Goffman, 1959).

This awareness of student perception of instructional behaviors is particularly beneficial for instructors practicing reflective teaching. Reflective teaching brings about change in one's teaching by evaluating and considering the purpose of actions or teaching style through observation and reflection (Richards, 1991). Through reflective teaching instructors can identify what instructional behaviors they display in the classroom, get feedback about student perception of the impact of those behaviors on classroom learning, and then carefully consider making changes to their behaviors for the purpose of improving student learning. Instructors employing reflective teaching techniques will therefore achieve greater insight regarding what changes to make and why by listening to student perception of the classroom environment (Pena and Leon, 2011).

Instructors, for instance, may be able to improve student perception of their involvement, engagement, interaction, ability to keep student attention, and ability to make the material relevant to students by using active and collaborative learning practices such as case studies,

clickers, and question and answer sessions (Umbach and Wawrzynski, 2005; Martyn, 2007). Studies have also shown that instructors can show they care about student learning by employing behaviors such as answering student questions or making themselves available to students through office hours, review sessions, or open discussion sessions (Garko et al., 1994; Hawk and Lyons, 2008; Johnston et al., 2011). Social media sites can be used to foster student-instructor relationships by conveying a competent and trustworthy instructor; however, instructors must be aware that they need to maintain consistency between the personality they portray on social media versus in the classroom (Mazer et al., 2007). Instructors might appear more calm and comfortable in front of students if they spend more time preparing for their course, rehearsing, and visualizing success (Marincovich et al., 1998; Collins, 2004); this would support improved student learning of the content, and the ability to explain the content and answer questions about it. When instructors receive questions that they are uncertain how to answer they can use the moment as an opportunity to connect with students and show enthusiasm for finding the answer while fostering positive student attitude (Halandya et al., 1982), versus being uncomfortable and uncertain.

Relationships Among Behaviors

When Kendall and Schussler (2012) first identified the descriptive words used in this study, they attempted to organize them into themes without knowing the behavioral descriptions this study has revealed. Student explanations of the descriptive words have now allowed them to be re-grouped into positive and negative counterparts of the same behavioral aspect (Table 13). For instance, instructors identified as being “engaging” or “enthusiastic” share common instructional behaviors such as being interactive and having a passion for teaching. On the other hand, student explanations of the term “boring” indicate that it is the behavioral opposite of

“engaging” and “enthusiastic.” Similarly, “confident” and “relaxed” are described by positive instructional behaviors such as being comfortable and being sure about themselves, which appear to be the opposites of the descriptive words “uncertain” and “nervous.” This is also the case for the behaviors associated with the terms “relate” (approachable) and “distant” (not approachable). Lastly, “understanding” instructors are described as being flexible and empathetic, as compared to “strict” instructors, who appear to be the opposite.

However, this study also identified two descriptive words that could not be organized into the proposed themes. The descriptive words “organized” and “respect,” had no negative behavioral counterparts verbalized by students in this study. In the case of the word organized, it may be that there was no descriptive word that served as its opposite identified in the original study (Kendall and Schussler, 2012). Organization is associated with having everything ready to go, good planning, and prompt grading; it may be that no terms such as “disorganized” appeared in the original study because although GTAs may be perceived as less organized than professors, they are not perceived as disorganized.

The description “respect” appears to be a bit more complex, however, because in order for students to indicate that they respect the instructor, the instructor must be able to answer questions, have some flexibility, understand student perceptions, and make the course interesting. Students in this study indicated that if an instructor is one they respect, they believed they would learn from them and indicated that they were a good teacher. Based on the description of this term, it appears that respect is an overarching factor that emerges from an instructor displaying multiple positive instructive behaviors within each of the identified behavioral themes in this study. Pattison et al. (2011) also found that respect was an overarching theme of their study; however, in our case it is respect *from* students while Pattison et al. (2011)

examined respect *for* students. It is unknown, however, whether an instructor has to display all of the positive behavioral aspects described by students in this study to achieve "respect," or whether they may still hold some less positive behaviors and be respected for other reasons.

Future studies can use the schema generated from this study to see whether the same descriptions and behaviors hold true for students at different institutions, or for students taking courses in a discipline other than biology. Researchers can also determine which instructional behaviors instructors must display to gain student respect, or if respect is afforded at the beginning of the semester and then lost through the emergence of negative instructional behaviors. Also, studies can explore if particular themes of the schema are more important than others for earning respect. Moreover, the schema can be used to explore the impact of modifying particular behaviors on students evaluations, or if some behavioral modifications benefit one instructor type more than the other (e.g. GTAs versus faculty members).

Limitations

Since student participants of this study came from a single institution, it is unknown how these perceptions might compare to students at other institutions. It is possible that students at different institution types (e.g., public versus private) may view these descriptors and underlying behaviors differently. Also, institutions with more diversity may also impact student perception of these descriptors because of different cultural perceptions of teaching and learning behaviors. Another limitation of this study is that students were asked to describe their biology instructors; it is possible that students may emphasize different instructional behaviors when asked about instructors in different disciplines. Given this, it is cautioned that the results of this study are not likely to be generalizable to all institutions or disciplines, unless further research indicates that the results are in fact consistent across a diversity of contexts.

Conclusion

Instructor descriptions used by undergraduates are much more complex than the single words might imply due to the multiple personal- and classroom-level instructional behaviors that lead to the descriptions. These data have the potential to impart useful insights into undergraduate perceptions of instructors, particularly into how behavior affects student perception of learning. It also allows for the exploration of a caution made by Feldman (1988) indicating that while students determine specific characteristics that are important for good teaching, these may not be the same aspects they use to judge instructors in practice. Future studies can use these results to work with individual instructors to identify their instructional behaviors and study how they affect student perception of learning.

This study can also be used as a potential tool to interpret student evaluations of teaching (Zabaleta, 2007). Interpretations of descriptive words in student evaluations may give instructors more insight into how undergraduates perceive them, and help them to understand that it is more than just their teaching that students are reflecting on (Nussbaum, 1992; Helterbran, 2008). From these reflections, instructors can identify what instructional behavior modifications could be made in order to earn the respect of their students, and increase their abilities to foster student learning.

Literature Cited

Arnon S, Reichel N. (2007). Who is the ideal teacher? Am I? Similarity and difference in perception of education regarding the qualities of a good teacher and of their own qualities as techers, Teachers and Teaching *13*, 441-464.

Baird JS (1987). Perceived learning in relation to student evaluation of university instruction. J Educ Psychol *79*, 90–91.

Clayson DE, Frost TR, Sheffet MJ (2006). Grades and the student evaluation of instruction: a test of the reciprocity effect. Acad Mgmt Learn Educ *5*, 52–65.

Coll RK, Taylor N, Fisher DL (2002). An application of the Questionnaire on Teacher Interaction and college and university classroom environment inventory in a multicultural tertiary context. Res Sci Technol Educ *20*, 165–183.

Collins J (2004). Giving a powerpoint presentation: The art of communicating effectively. Radio Graphics *24*, 1185-1192.

Dudley M (2009). Jumping out of an airplane: a TA's perspective on teaching effectiveness. East Educ J *38*, 1–10.

Emery CR, Kramer TR, Tian RG (2003). Return to academic standards: a critique of students' evaluations of teaching effectiveness. Qual Assur Educ *11*, 37–46.

Feldman K A (1988). Effective college teaching from the students' and faculty's view: Matched or mismatched priorities? Res Higher Educ *28*, 291-344.

Garko MG, Kough C, Pignata G, Kimmel EB, Eison J (1994). Myths about student-faculty relationships: What do students really want? J Excellence Coll Teach *5*, 51-65.

Goffman E (1959). The Presentation of Self in Everyday Life. New York: Doubleday.

HalandyaT, Olsen R, Shaughnessy J (1982). Relations of student, teacher, and learning environment variables to attitudes toward science. Sci Educ *66*, 671-687.

Hawk TF, Lyons PR (2008). Please don't give up on me: When faculty fail to care. J Management Educ *32*, 316-338.

Helterbran VR (2008). The ideal professor: student perceptions of effective instructor practices, attitudes, and skills. Education *129*, 125–138.

Huck S (2008). Reading Statistics and Research, Boston: Pearson Education.

Johnston PH, Ivey G, Faulkner A (2011). Talking in class: Remembering what is important about classroom talk. The Reading Teach *65*, 232-237.

Kendall KD, Schussler EE (2012). Does instructor type matter? Undergraduate student perception of graduate teaching assistants and professors. CBE Life Sci Educ *11*, 187–199.

Koballa TR. Jr, Crawley FE (1985). The infuence of attitude on science teaching and learning. School Sci and Mathematics *85*, 222-232.

Kogan LR, Schoenfeld-Tacher R, Hellyer PW (2010). Student evaluations of teaching: perceptions of faculty based on gender, position, and rank. Teach Higher Educ *15*, 623–636.

LeCompte MD (2000). Analyzing qualitative data. Theory Pract *39*, 146–154.

Marincovich M, Prostko J, Stout F (eds.) (1998). The Professional Development of Graduate Teaching Assistants, Balton, MA: Anker Publishing.

Martyn M (2007). Clickers in the classroom: An active learning approach. EDUCAUSE Quarterly *30*, 71-74.

Mazer JP, Murphy RE, Simonds CJ (2007). I'll see you on "Facebook": The effects of computer-mediated teacher self-disclosure on student motivation, affective learning, and classroom climate. Comm Eduac *56*, 1-17.

Muzaka V (2009). The niche of graduate teaching assistants (GTAs): perceptions and reflections. Teach Higher Educ *14*, 1–12.

Nussbaum JF (1992). Effective teacher behaviors. Commun Educ *41*,167–180.

Osborne J, Simon S, Collins S (2003). Attitudes toward science: A review of the literature and its implications. Int J Sci Educ *25*, 1049-1079.

Park C (2002). Neither fish nor fowl? The perceived benefits and problems of using graduate teaching assistants (GTAs) to teach undergraduate students. High Educ Rev *35*, 50–62.

Pattison P, Hale JR, Gowens P (2011).Mind and soul: connecting with students. J Legal Stud Educ *28*, 39–66.

Pena C, Leon LD (2011). The use of digital video to foster reflective practice in teacher education. Int J Instructional Media *38*, 125-132.

Richards JC (1991). Towards reflective teaching. The Teacher Trainer *5*, 4-8.

Rubin H J, Rubin IS (2012). Qualitative Interviewing: The Art of Hearing Data (3rd ed.). Los Angeles, CA: Sage.

Teven JJ (2007). Teacher caring and classroom behavior: Relationships with student affect and perceptions of teacher competence and trustworthinesss. Comm Quarterly *55*, 433-450.

Treagust DF, Fraser BJ (1986). Validation and application of the college and university classroom environment inventory (CUCEI). Paper presented at the Annual Meeting of the American Educational Research Association, San Francisco, CA, April 16–20.

Umbach PD, Wawrzynski MR (2005). Faculty do matter: the role of college faculty in student learning and engagement. Res High Educ *46*, 153–184.

Varca PE, Pattison P (2001). Attaining teaching excellence: A critical incident technique. J Legal Studies Educ *19*, 153-185.

Wright RE (2006). Student evaluations of faculty: Concerns raised in the literature, and possible solutions. Coll Student J *40*, 417-422.

Zabaleta F (2007). The use and misuse of student evaluations of teaching. Teach Higher Educ *12*, 55–76.

Appendix

Table 12. Student instructor descriptions grouped into three overarching themes (delivery technique, classroom atmosphere, and relationship) within the teaching and personal realms of instructor type (modified from Kendall and Schussler, 2012)[a]

Realm	Theme	GTA	Professor
Teaching	Delivery Technique	Hesitant, **nervous**, **uncertain**, and unsure how to begin teaching.	**Organized** and structured, **confident**, in control, prepared for questions, **enthusiastic,** with previous teaching experience, and greater knowledge.
	Classroom Atmosphere	**Relaxed** and laid back, interactive, **engaging**, personalized, and having open student-instructor interactions.	**Distant** and formal, **strict**, serious, harder, with higher expectations and standards.
Personal	Relationship	Comfortable approaching GTAs and GTAs are **relatable** and **understanding**.	Intimidating and **boring**, and out of touch, yet undergraduates **respect** professors.

[a] Words shown in bold were the descriptive words selected to further explore in the current study.

Table 13. Undergraduate explanations of descriptive words, sorted into paired, positive and negative aspects.

Positive Terms	Explanation	Negative Terms	Explanation
Engaging	Keep student attention by being interactive or involved, have a passion for teaching, bring in examples, and communicate on student level.	*Boring*	Monotone, not engaging or interactive, not interesting, have no personal enthusiasm, do not come to student level, and as a result are unable to make material interesting or keep student attention.
Enthusiastic	Exciting, happy, enjoy and have a passion for teaching, care about students, are interactive, have a connection to the topic, are able to interest students, and use examples.		
Confident	Know the material and information, are sure of their teaching, can answer student questions, come prepared, do not stutter or shake, are calm, and thus are effective instructors.	*Nervous*	Not confident, and thus are uncomfortable, shaky and sweaty, are a poor teacher, and do not know material.
Relaxed	Not nervous or anxious; comfortable, prepared, and laid-back. Approachable, flexible about rules, and confident.	*Uncertain*	Not confident, do not know the material, unable to answer student questions, are unsure how to teach, and often come off as being nervous.
Relate	Approachable, able to connect with students, have common interests, understand how to present the material to students.	*Distant*	Not personally approachable, not engaging or interactive in class, do not care about students, are intimidating, and not relatable.
Understanding	Compassionate when dealing with student issues, are often flexible (especially in terms of deadlines), empathetic, and they know what it is like to be a student.	*Strict*	Adhere to policies and rules, are inflexible, are tough graders, and not tolerant of bad behavior or distractions.

Table 14. Undergraduate explanations of the descriptive words "organized" and "respect," in which there were no negative terms or similar words associated with them.

Term	Explanation
Organized	Instructors who are organized have everything ready to go and orderly, have a plan for the course, start on time, provide prompt grading, do not lose assignments, and have a conscientious flow to class.
Respect	Through their actions and knowledge they are able to earn the trust of students and in turn student will follow their rules and provide undivided attention to the instructor indicating they respect the instructors. Instructors which are respected are able to answer questions and make the material interesting, listen to students' perspective, and are willing to help students. Further, some instructors are respected due to their experience and seniority as an authoritative figure.

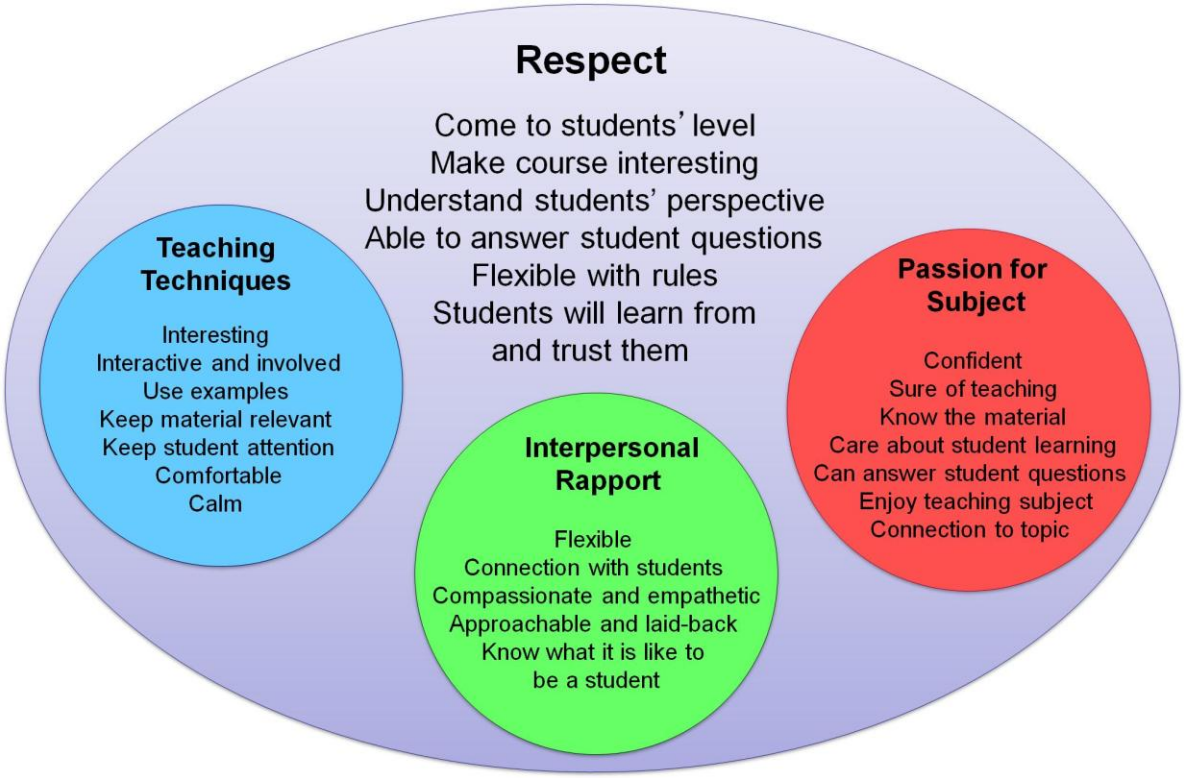

Figure 6. Schema depicting positive instructional behaviors within the three themes (teaching techniques, interpersonal rapport, and passion for subject) identified from student perspectives, which lead to the overarching description of respect.

CHAPTER FOUR

What matters most for graduate teaching assistant professional development?

K. Denise Kendall, Dylan Dittrich-Reed, Matthew L. Niemiller, Elisabeth E. Schussler

Abstract

Graduate teaching assistants (GTAs) are often used as instructors in introductory science courses, particularly in laboratory and discussion sections associated with large lectures. Introductory courses are important for student retention and learning in the sciences. Therefore, this study explored the behaviors most important for GTA teaching effectiveness from an undergraduate perspective. In Spring 2012, 1159 undergraduates in 100-level biology courses rated their GTA on 21 instructional behaviors that had previously been identified as important for student learning, the GTA's teaching effectiveness, and the amount the student learned in the laboratory course. Using linear mixed models we found that teaching techniques and interpersonal rapport best predicted teaching effectiveness of GTAs. Consequently, we make recommendations for four main themes of GTA professional development (content presentation, active learning, student relation, and policies and flexibility) that should maximize the limited professional development time for GTAs.

Introduction

With the national movement to increase student learning and retention in the sciences, it is important to consider the multiple factors that impact science education, including course instructors, curricula, and undergraduate student perceptions. This study focuses on student perception of the teaching effectiveness of graduate teaching assistants (GTAs) who often teach sub-sections (e.g., laboratories and discussion) of large

introductory science courses (Rushin et al., 1997; Sundberg et al., 2005) at higher education institutions.

Individual instructors impact student retention, attitudes, and learning success (Hartnett et al., 2003; O'Neal et al., 2007; Kneipp et al, 2010). GTAs in introductory courses typically have more personal contact with students than faculty (Bond-Robinson and Rodreques, 2006; Baumgartner, 2007); therefore, GTAs are likely extremely influential instructional staff in undergraduate gateway science courses (Sheridan, 1991; Rushin et al., 1997; McComas and Cox-Petersen, 1999; Sundberg et al., 2005). The nature of these influences can vary, however, as evidenced by reports that students learn little from GTAs who teach laboratory sections (Sunl et al., 2004) to reports that GTAs positively impact student retention in the sciences (O'Neal et al, 2007). GTAs must be recognized as key influences in science education (Sheridan, 1991; Dotger, 2010), and as such, formal pedagogical training for GTAs could be critical at institutions of higher education.

GTAs reportedly receive minimal support, training, or continuous mentoring during their graduate tenure (Rushin et al, 1997; Shannon et al., 1998; Luft et al, 2004; Penwel et al., 2004; Tanner and Allen, 2006; Dotger, 2010) despite the fact that GTAs are often novice instructors (Shannon et al., 1998; Luft et al., 2004). Some universities offer pre-academic year orientations or seminars (at university and/or departmental levels) or distribute training manuals (Lowman and Mathie, 1993; Rushin et al., 1997; Luft et al., 2004; Dotger, 2010), while other universities offer formal education courses or mentoring (Boyle and Boice, 1998; Baumgartner, 2007). However, departmental and

university policies are the focal point of much GTA training, and as a result, GTAs typically receive limited pedagogical information (Shannon et al., 1998; Luft et al., 2004).

Professional development of GTAs is crucial given that 85% of GTAs report they do not feel adequately trained for their teaching assignments (Russell, 2009). GTAs require more than yearly workshops, training manuals, and orientations to achieve awareness of teaching and learning skills (Bond-Robinson and Rodriques, 2006; Luft et al., 2004; Tanner and Allen, 2006; Trautmann, 2008). However, efforts to train and supervise GTAs may go unrewarded or unsupported (Prieto, 2002), or may encounter resistance from administration, faculty, and the graduate students themselves. GTAs often receive conflicting signals regarding professional development: GTAs are told that teaching is important, but they are also told to limit the time they devote to teaching assignments (Nyquist et al., 1999; Austin et al, 2009), particularly at research-oriented institutions (Sheridan, 1991; Flora, 2007).

GTA professional development for teaching is critical to the preparation of future faculty (Boyer, 1990; Nyquist et al, 1999; Austin et al., 2009) considering that GTAs are the faculty of tomorrow (Lowman and Mathie, 1993, Austin, 2002). Many current faculty members indicate they never participated in formal teaching or learning professional development during graduate school (Boyle and Boice, 1998; Austin, 2002; Trautmann, 2008). This could be the cause of, or contribute to, the faculty belief that no formal training is necessary (Rushin et al., 1997). Thus, administrations might limit GTA training opportunities because faculty are unsupportive and graduate students do not see

the utility. Finally, even the best GTA development programs may be limited in time, resources, and scope (Roach, 1997) and choose not to invest in extensive professional development.

Given these limitations, this study aims to identify a small set of instructional behaviors that undergraduates perceive as important for effective instruction in order to generate brief but effective professional development sessions for GTAs. Training GTAs in the identified instructional behaviors should increase undergraduate learning and promote the teaching abilities of GTAs with less time and cost.

Project Rationale

Instructors have a central role in fostering student learning and retention through effective instruction. Students expect effective instructors to display positive teaching behaviors, such as making material relevant, being prepared, being organized, encouraging student questions, and being compassionate (Helterbran, 2008; Pattison et al., 2011). If an instructor displays negative behaviors including grading inconsistently, not providing students with feedback, or being impatient, students might perceive them as less effective teachers (Pattison et al., 2011). Kendall and Schussler (2013a, 2013b) further revealed that words commonly used to describe instructors consist of underlying instructional behaviors that span both personal and classroom contexts.

Given that most student attrition occurs in the first academic year (Seymour and Hewitt, 1997) and GTA-led sections are often associated with introductory science courses (Rushin et al., 1997; Sundberg et al.., 2005), this study focuses on first-year (100 level) biology courses. To date, it is unknown whether certain GTA behaviors are more

important for student perception of effective instruction than others. Departments and universities may face constraints to prepare GTAs for their teaching assignments. Therefore, recommendations for GTA professional development, particularly in first-year laboratory courses, may be based on the results of this study.

Methods

Data Collection. Data were collected using an online survey (hosted by surveymonkey.com) administered to undergraduate students in GTA-instructed Introductory Biology laboratories in Spring 2012 at a research-intensive university in the southeastern United States. This survey was administered to undergraduates enrolled in several courses of 15-week semester-long, 100 level, non-majors (second semester Introductory Biology) and majors (Biodiversity and Cellular Biology) general Biology courses. All the laboratories were taught by GTAs, who typically teach two or three laboratory sections per semester, with a maximum class size of 27 students.

Survey Design. The survey consisted of seven items (Likert-style response scales) in each of the three dimensions (teaching techniques, interpersonal rapport, and passion for subject – these will later be referred to as "sub-scales") identified by Kendall and Schussler (2013b) as important for instructor respect and student learning (Table 15). The survey instructed undergraduate participants to indicate their agreement with each item using a six-point response scale of: strongly agree, agree, slightly agree, slightly disagree, disagree, and strongly disagree. In addition to these 21 items, the survey asked students to rate their perception of their GTA's teaching effectiveness, the amount they learned in the laboratory, and their performance (expected grade). Students rated their GTA's teaching

effectiveness on a six-point response scale of: very poor, poor, fair, good, very good, excellent; and the amount they learned on a six-point response scale from 1-6 with 1 representing “little” and 6 representing “a lot”. Students also provided demographic information, including current enrollment status, native language, major, gender, current biology course enrollment, and previous biology course enrollment.

Survey administration. The online survey was appended to the regular programmatic lab evaluations, which are often completed as part of the final lab exam. If the survey was completed as part of the lab final exam, students were offered one point of their 55-point lab final to complete the survey. This was accomplished by having the GTA check off the question after the student visited the computer. If the student chose not to take the survey or to skip the question they still received the point. Therefore, this survey was voluntary for all students, did not have a measureable impact on their grade, and GTAs did not see any student responses. All procedures were reviewed and approved by the Institutional Review Board for Human Subjects.

Survey reliability. To determine the reliability of the survey, Cronbach’s alphas was calculated for each of the sub-scales (SPSS v19.0, IBM). The Cronbach’s alpha for teaching techniques, interpersonal rapport, and passion for subject were 0.943, 0.948, and 0.948, respectively. Cronbach’s alphas was also calculated for each GTA, and these results ranged from 0.794 to 0.989 for teaching techniques, 0.831 to 0.987 for interpersonal rapport, and 0.860 to 0.990 for passion for subject. From these results it was judged that the survey was reliable.

Data Analysis

Statistical Analyses. To account for the non-independence of survey responses, we fit a series of linear mixed effects models to our data. We treated student demographics, sub-scale indices (sum of the seven responses within each sub-scale; teaching technique index [TTI], interpersonal rapport index [IRI], passion for subject index [PSI]), and demographics as fixed effects and fit random intercepts for section, GTA, and course. The full model included GTA teaching effectiveness (Effectiveness) as the response variable and TTI, IRI, PSI, current enrollment status (Year), previous biology course enrollment (TotalBio), native English speakers (Eng), biology majors (Major), gender (Gender), estimated grade (Grade), and amount learned (Learned) as fixed effects. We compared the full model to various reduced models that included different combinations of the fixed effects that were significant at the $\alpha = 0.05$ level in the full model (TTI, IRI, PSI, Grade, and Learned). The *p* values for a given fixed effect were approximated by the *p* values of a likelihood ratio test comparing the full model with a model reduced by the fixed effect. We selected the best model based on Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) scores. All analyses were conducted in R (R version 2.15).

Results

Participants. The survey was administered to a potential pool of 1,881 undergraduates in 86 laboratory sections. After removing students who did not complete the survey in its entirety, 1,159 undergraduates remained. This response rate was a result of the survey being voluntary or student absences during administration of the surveys.

Participants were primarily freshman (56.7%), non-biology majors (79.0%), female (60.1%), and native English speakers (96.3%); additional demographic information can be found in the supplementary materials.

Survey responses. The 1,159 complete survey responses stemmed from undergraduates taught by 39 GTAs. As our objective was to determine the sub-scale (TTI, IRI, or PSI) that best predicted undergraduate perception of GTA teaching effectiveness, we removed 319 responses that had no variation in sub-scale indices. In other words, we removed all student responses to sub-scale items that, for example, consisted of all "strongly agree". Excluding the data from one GTA which was identified as an outlier, the minimum number of student responses for a GTA was five, the maximum was 53 responses, and the mean was 30 responses with a standard deviation of 10.31.

Linear Mixed Effects Models. A reduced model with TTI, IRI, and Learned as fixed effects best predicted GTA effectiveness, based on both AIC and BIC scores (AIC = 1596.17, BIC = 1634.02, Table 16). Although both student ratings of GTA passion for subject (PSI) and estimated grade (Grade) predicted student perception of GTA teaching effectiveness in the full model (p = 0.007 and p = 0.016, respectively), neither was included in the best fitting model. The perceived amount that a student learned (Learned) was the only demographic or self-assessment predictor to remain in the best fitting model. This result was corroborated by a simple correlation test using mean values of Effective, Learned, and Grade for each GTA included in the mixed model. Students

associate GTA teaching effectiveness with the amount they learned (Pearson's $r = 0.671$, $p < 0.001$), not their expected grade (Pearson's $r = 0.019$, $p = 0.911$).

Conclusions

Our study provides evidence that student perceptions of GTA teaching effectiveness can be predicted by instructional behaviors representative of the teaching techniques and interpersonal rapport sub-scales and student perception of how much they learned. This association between GTA teaching effectiveness and student learning, versus grade and GTA teaching effectiveness suggests that students are able to distinguish between grades and learning when it comes to evaluating instruction. Additionally, students appear to value both classroom (teaching technique) and personal (interpersonal rapport) aspects as important to effective instruction from GTAs. As a result, GTA professional development should advance instructional behaviors in both categories to cultivate student learning. Accordingly, we recommend below four short sessions to specifically promote instructional behaviors in the teaching technique and interpersonal rapport sub-scales.

Teaching Techniques. Students in this study highlighted the importance of teaching techniques (aspects of this include being calm, being interesting, keeping student attention, using good examples, making the material relevant to students, being interactive, and being good at presenting the material) for effective instruction. We believe professional development programs could offer sessions focused in two sessions to foster these skills: 1) content presentation and 2) active learning.

Content Presentation. GTAs that are calm, use good examples, keep student attention, make the material relevant to students, and are good at presenting the material are perceived as more effective instructors; therefore graduate programs could offer a session that fosters these basic instructional skills. This session would include video recorded examples of how teachers with different instructional styles maintain student attention (modulating tone, moving around classroom, gesture; Richards, 1991) as well as examples and analogies, such as examples from pop culture or recent news, that teachers can use to support student learning for a specific discipline (Udovic et al., 2002; Pryor, 2008*)*. Participant discussion of these examples, and time to create personalized strategies should be included in the session. Extensions of this session could include the establishment of a peer mentor group, or systematic mentoring, to provide GTAs support for discussion about how to present course material to undergraduates, and solicit feedback, observations, and informal evaluations of their own or the teaching of others (Richards, 1991; Boyle and Boice, 1998; Ward and West, 2002; Kohut et al., 2007).

Active Learning. The aim of this session would be to provide GTAs with strategies for increasing classroom interaction and keeping student attention. The session would introduce GTAs to what active learning is and then present them with exemplary active learning techniques (e.g., clicker usage, asking questions, or case studies) that encourage student involvement in the classroom (Udovic et al, 2001; Prince, 2004; Allen and Tanner, 2005). GTAs should end the session with the creation of their own strategies to use in their own classes to increase student involvement (Garet et al., 2001; Allen and Tanner, 2005), including technological aspects (Kemp and O'Keefe, 2003), of

implementing these techniques. A follow-up session could be offered to allow GTAs to discuss what techniques they implemented, what difficulties they encountered, and how successful these techniques were.

Interpersonal Rapport. This study also revealed the importance of interpersonal rapport (aspects include being compassionate, being empathetic, being flexible, being laid-back, connecting with students, being approachable, and having a good relationship with students) for student perception of effective instruction from GTAs. We propose two professional development sessions: 1) student relation and 2) policies and flexibility.

Student Relation. GTAs who connect with students, have a good relationship with students, are approachable, and are laid-back create a more comfortable learning environment for students. This professional development session would begin by having GTAs reflect on their own undergraduate experiences and recall being a student in a 100 level course. This can be followed by participation in "ice-breaker" strategies that could be used by GTAs get to know students individually throughout the semester (Mendes, 2003). This session should also emphasize that GTAs should learn and use their students' names, and get to know their students' career goals, as this can promote positive interactions in the classroom (Mendes, 2003). These strategies might promote positive student perceptions regarding GTA approachability and ability to establish good relationships with students.

Policies and Flexibility. The objective of this session would be to increase GTA awareness about when to enforce policies, and when to disregard or bend them, as well as developing behaviors of empathy, compassion, and flexibility. The enforcement of

policies can support an instructor's classroom management and establishment of authority; however, adjustment of policies throughout a course may promote perceptions of being empathetic, compassionate, and flexible. This session can consist of short literature discussions (Kemp and O'Keefe, 2003) and/or role playing events (Fischer and Lann, 2002; Woodward-Young, 2008), that may make GTAs more receptive to innovative or alternative solutions to situational events that may occur when teaching (Kemp and O'Keefe, 2003; Lynd-Balta et al., 2006). Since discussion among GTAs, and particularly the inclusion of experienced GTAs, will likely provide a range of options about the need to be empathetic, compassionate, and flexible, while still maintaining course policies an extension to this session could be the establishment of GTA learning circles (Lynd-Balta et al., 2006). This will enhance GTA instructional decision-making skills so they can increase their confidence in dealing with the myriad of student issues that can occur over a semester.

Limitations and Future Direction

Although this study provides compelling evidence for the importance of GTA instructional behaviors for student perception of teaching effectiveness, these results must be considered with caution until confirmed by studies at other universities. This study was conducted utilizing volunteer participants from one discipline at a single higher education institution; consequently, it is unknown whether these results can be generalized to other disciplines or universities.

This study was conducted within the context of fostering student retention and learning in the sciences through instructor teaching effectiveness and student learning.

Future research exploring the implications of these behaviors, in particular with regards to student perception of their learning and likelihood of retention in the sciences, is warranted. For instance, while this study documented that student perception of teaching techniques and interpersonal rapport predict student perception of teaching effectiveness from GTAs, it is unknown whether this translates into positive student attitude toward the sciences, or motivation to pursue a science degree. While further research is warranted, the results of this study are useful to GTA professional development administrators as a guide for supporting GTAs.

Summary

If higher education science departments aspire to foster student learning and retention, then it is necessary for all aspects of science courses to promote positive learning environments (Seymour and Hewitt, 1997; President's Council of Advisors on Sciences and Technology, 2012). Fostering GTA professional development in science departments can promote positive classroom experiences for students, which in turn should promote student learning, retention, and attitude toward the sciences. However, GTAs must be provided with the necessary background and professional development to promote their teaching effectiveness and maximize their students' learning in introductory science courses. Since many faculty members note their only teaching experience stemmed from GTA assignments, commitment to GTA professional development by science departments and graduate schools could not only enhance teaching and learning by current GTAs, but could also improve instruction by future faculty members.

Literature Cited

Allen D, Tanner K (2005). Infusing active learning into the large-enrollment biology class: Seven strategies, from simple to complex. CBE-Life Sci Educ *4*, 262-268.

Austin AE (2002). Preparing the next generation of faculty: Graduate school as socialization to the academic career. J High Educ *73*, 94-122.

Austin AE, Campa III H, Pfund C, Gillian-Daniel DL, Mathieu R, Stoddart J (2009). Preparing STEM doctoral students for future faculty careers. New Directions for Teach. and Learn *117*, 83-95.

Baumgartner E (2007) A professional development teaching course for graduate students. J Coll Sci Teach *36*, 16-21.

Bond-Robinson J, Rodriques RAB (2006). Catalyzing graduate teaching assistants' laboratory teaching through design. Res J Chem Educ *83*, 313–323.

Boyer EL (1990). Scholarship Reconsidered: Priorities of the Professoriate. New York: The Carnegie Foundation for the Advancement of Teaching, Jossey-Bass.

Boyle P, Boice B (1998). Systematic mentoring for new faculty teachers and graduate teaching assistants. Innovative High Educ *22*, 157-179.

Dotger S (2010). Offering more than "Here is the Textbook": Teaching assistants'' perspectives on introductory science courses. J Coll Sci Teach *39*, 71-76.

Fischer J, Laan SV (2002). Improving approaches to multicultural education: Teaching empathy through role playing. Multicultural Educ *9*, 25-26.

Flora BH (2007).Graduate assistants: students or staff, policy or practice? The current legal employment status of graduate assistants. J High Educ Pol Manag *29*, 315–322.

Hartnett N, Roemcke J, Yap C (2003). Recognizing the importance of instruction style on student's performance: Some observations from laboratory research – a research note. Account Educ *12*, 313-331.

Helterbran VR (2008). The ideal professor: student perceptions of effective instructor practices, attitudes, and skills. Education *129*, 125–138.

Kemp PR, O'Keefe RD (2003). Improving teaching effectiveness: Some examples from a program for the enhancement of teaching. Coll Teach *51*, 111-114.

Kendall KD, Schussler EE (2013). Evolving impressions: Undergraduate perceptions of graduate teaching assistants and faculty members over a semester. CBE-Life Sci Educ *12*.

Kendall KD, Schussler EE (2013). More than words: Probing the terms undergraduate students use to describe their instructors. Int J Teach and Learn. in High Educ *25*.

Kneipp LB, Kelly KE, Biscoe JD (2010). The impact of instructor's personality characteristics on quality of instruction. Coll Student J *44*, 901-905.

Kohut GF, Burnap C, Yon MG (2007). Perceptions of the observer and the observed. Coll Teach *55*, 19-25.

Lowman J, Mathie V (1993). What should graduate teaching assistants know about teaching? Teach Psychol *20*, 84–88.

Luft JA, Kurdziel JP, Roehrig GH, Turner J (2004). Growing a garden without water: graduate teaching assistants in introductory science laboratories at a doctoral/research university. J Res Sci Teach *41*, 211–233.

Lynd-Balta E, Erklenz-Watts M, Freeman C, Westbay TD (2006). Professional development using an interdisciplinary learning circle. J Coll Sci Teach *35*, 18-24.

McComas WF, Cox-Petersen AM (1999). Enhancing undergraduate science instruction – The G-Step approach, Capitalizing on the pedagogical strengths of science educators and the content expertise of science TAs. J Coll Sci Teach *29*, 120-125.

Mendes E (2003). What empathy can do. Educ Leadership *61*, 56-59.

Nyquist LM, Wulff DH, Austin AE, Sprague J, Fraser PK, Calcagno C, Woodford B (1999). On the road to becoming a professor: the graduate student experience. Change *31*, 18–27.

O'Neal C, Wright M, Cook C, Perorazio T, Purkiss J (2007). The impact of teaching assistants on student retention in the sciences: lessons for TA training. J Coll Sci Teach *36*, 24–29.

Pattison P, Hale JR, Gowens P (2011).Mind and soul: connecting with students. J Legal Stud Educ *28*, 39–66.

Penwell RA, Elsawa SF, Pitzer T (2004). Cooperative and active learning in undergraduate biological laboratories at FIU—Implications to TA teaching and training. Bioscene *30*, 9-12.

President's Council of Advisors on Science and Technology (PCAST) (2012). Engage to Excel: Producing One Million Additional College Graduates with Degrees in

Science, Technology, Engineering, and Mathematics. Report to the President. www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-engage-to-excel-final_2–25–12.pdf. Accessed 3. April. 2012.

Prieto L (2002). Teaching assistant development: Research and impressions. In W. Buskist, V. Hevern, & G W. Hill, IV, (Eds.). Essays from e-xcellence in teaching, 2002 (Chap. 1). Retrieved 29.November.2011 from the Society for the Teaching of Psychology Web site: http://teachpsych.lemoyne.edu/teachpsych/eit/index.html

Prince M (2004). Does active learning work? A review of the research. J Engineering Educ *93*, 223-231.

Pryor GS (2008). Using pop culture to teach introductory biology. Am Biol Teach *70*, 396-399.

Richards JC (1991). Towards reflective teaching. Teach Trainer *5*, 4-8.

Roach K (1997).Effects of graduate teaching assistant attire on student learning, misbehaviors, and ratings of instruction. Commun Quart *45*, 125–141.

Rushin JW, Saix JD, Lumsden A, Streubel DP, Summers G, Bernson C (1997). Graduate teaching assistant training: a basis for improvement of college biology teaching and faculty development? Am Biol Teach *59*, 86–90.

Russell JA (2009). A survey of basic instructional program graduate teaching assistant development and support processes. Res Quarterly for Exercise and Sport *80*, 792-795.

Seymour E, Hewitt N (1997). Talking about Leaving: Why Undergraduates Leave the Sciences, Boulder, CO: Westview.

Shannon DM, Twale DJ, Moore MS (1998). TA teaching effectiveness: the impact of training and teaching experience. J High Educ *69*, 440– 466.

Sheridan JD (1991). A proactive approach to graduate teaching assistants in the research university: One graduate dean's perspective. In JD Nyquist, RD Abbot, DH Wulff, & J Sprague (Eds.), Preparing the Professorate of Tomorrow to Teach (pp. 24-28). Dubuque, IA: Kendallmunt.

Sunal DW, Wright EL, Day JB (Eds.) (2004). Science Teaching for the 21st Century. USA: Information Age Publishing Inc.

Sundberg MD, Armstrong JE, Wischusen EW (2005). A reappraisal of the status of introductory biology laboratory education in U.S. colleges and universities. Am Biol Teach *67*, 525–529.

Tanner K, Allen D (2006). Approaches to biology teaching and learning: On integrating pedagogical training into the graduate experiences of future science faculty. CBE-Life Sci Educ *5*, 1-6.

Trautmann NM (2008). Learning to teach: Alternatives to trial by fire. Change *40*, 40-45.

Udovic D, Morris D, Dickman A, Postlethwait J, Wetherwax P (2002). Learning in an introductory biology course. BioScience *52*, 272-281.

Ward JR, West LS, Isaak TJ (2002). Mentoring: A strategy for change in teacher technology education. J Tech and Teach Educ *10*, 553-569.

Woodward-Young PA (2008). A spot of otherness: Using the dot game to nurture intellectual empathy in future teachers. Democracy & Education *17*, 51-52.

Appendix

Data coding. Responses for each survey item for each instructor were converted to a numeric scale in the following manner: strongly agree = 6, agree = 5, slightly agree = 4, slightly disagree = 3, disagree = 2, strongly disagree = 1. The index for each Likert-style scale (teaching techniques, interpersonal rapport, and passion for subject) was calculated as the sum of the seven responses. Current enrollment status (number of years of undergraduate study, first year = 1, second year = 2, third year =3, fourth year = 4, fifth year and beyond = 5) and previous biology course enrollment (number of semesters of biology, coded as summation of total biology courses completed) were treated as continuous variables. Native language (English = 1 or other = 2), major (biology = 1, 2-4 account for biology concentrations [Biochemistry, Cellular, and Molecular Biology = 2, Ecology and Evolutionary Biology = 3, and Microbiology = 4], and other = 5), and gender (male = 1 or female = 2) were treated as categorical variables. Student estimates of the grade they earned were converted from letter to numeric (A-F to 4-0) and treated as a continuous variable. Student ratings of the amount they learned in the course (little = 1, a lot = 6, with continuum between) were treated as a continuous variable. Teaching effectiveness responses were converted to numeric scale in the following manner: very poor = 1, poor = 2, fair = 3, good = 4, very good = 5, excellent = 6.

Additional student demographics. Second- and third-year students comprised 23.6% and 11.0%, respectively, of the respondents, with 8.6% more being fourth year or beyond. Most of the respondents were currently enrolled in a non-majors second semester Introductory Biology course (56.1%), while the rest were currently enrolled in majors Biodiversity (21.7%) or Cellular Biology (22.2%). The majority of respondents had completed at least one other semester-long laboratory Biology course (71.8%) including 49.9% having completed non-majors first semester Introductory Biology, 18.2% Biodiversity, and 2.9% Cellular Biology.

Table 15. Twenty-one instructional behaviors included in the survey, divided into their respective sub-scales (teaching techniques, interpersonal rapport, passion for subject).

Survey Items
Teaching Techniques
My instructor is calm.
My instructor is interesting.
My instructor can keep student attention.
My instructor uses good examples.
My instructor makes the material relevant to me.
My instructor is interactive.
My instructor is good at presenting the material.
Interpersonal Rapport
My instructor is compassionate.
My instructor is empathetic.
My instructor is flexible.
My instructor is laid-back.
My instructor can connect with students.
My instructor is approachable.
My instructor has a good relationship with students.
Passion for Subject
My instructor is confident.
My instructor is sure of his/her teaching.
My instructor knows the material.
My instructor can answer student questions.
My instructor has a connection to the topic.
My instructor enjoys teaching.
My instructor cares about student learning.

Table 16. Description of fixed effects and random intercepts for full and reduced models including DF, log (Likelihood), AIC, and BIC results.

Model	Fixed Effects	Random Intercepts	DF	log (Likelihood)	AIC	BIC
Full	TTI + IRI + PSI + Gender + Major + Eng + TotalBio + Course + Year + Grade + Learned	Course + GTA + Section	15	-798.05	1626.11	1697.07
A1	TTI + IRI + PSI + Grade + Learned	Course + GTA + Section	10	-790.64	1601.29	1648.60
A2	TTI + IRI + PSI + Learned	Course + GTA + Section	9	-790.68	1599.37	1641.95
A3	TTI + IRI + Grade + Learned	Course + GTA + Section	9	-790.51	1599.01	1641.59
A4	**TTI + IRI + Learned**	Course + GTA + Section	8	-790.08	**1596.17**	**1634.02**
A5	TTI + Learned	Course + GTA + Section	7	-800.27	1614.54	1647.65
A6	TTI	Course + GTA + Section	6	-824.78	1661.55	1689.94

GENERAL CONCLUSIONS

The results presented in this dissertation indicate that undergraduate perception of instructors is influenced by instructor type and underlying behaviors displayed by individual instructors. Because many of the behaviors were perceived by undergraduates as linked to GTAs specifically, the results can be used to make targeted GTA professional development recommendations. The goal of GTA professional development should be to maintain as many of the positive aspects of GTA instruction as possible (as expressed by students), while simultaneously finding ways to decrease the negative perceptions expressed by undergraduates.

Chapter One found that undergraduate students can identify core differences due to instructor type (focusing on hypothetical differences versus specific instructors), and the qualitative data independently confirmed four of the five significant items obtained from the quantitative data. Undergraduates perceive professors to be more structured, confident, in control, organized, experienced, knowledgeable, distant, formal, strict, serious, hard, boring, out of touch, and respected than GTAs. Meanwhile, undergraduates perceive GTAs to be more uncertain, hesitant, nervous, relaxed, laid-back, engaging, interactive, relatable, understanding, and able to personalize teaching as compared to professors.

Chapter Two investigated student perception of descriptors for actual faculty members and GTAs teaching biology courses at UTK. While the results did not match all of the expectations based on the hypothetical comparisons of Chapter One (faculty members were expected to be more boring, confident, distant, enthusiastic, organized,

respected and strict than GTAs; GTAs were expected to be more engaging, nervous, relatable, relaxed, uncertain, and understanding than faculty [Kendall and Schussler, 2012; Muzaka, 2009; Park, 2002]), some expectations did hold true. As expected, undergraduates rated faculty members higher initially for being confident and enthusiastic while GTAs were rated higher for being nervous and uncertain. Interestingly though, undergraduate students rated faculty members higher for engaging, which is contrary to what was expected based on previous research. Also, no differences were seen for initial ratings of boring, strict, distant, relaxed, understanding, organized, respect, and strict.

It was also determined, however, that undergraduate student ratings of instructors change over the course of the semester, and for some characteristics these changes are significant. The trend of these changes is that GTAs gain more positive characteristics over the course of the semester (decreases in nervousness and uncertainty and gains in understanding, confidence, engaging and respect). Meanwhile, professors decrease in ratings of most positive descriptors (decrease in organization, relatedness, relaxed, confidence, engaging, respect), concurrently increasing in some negative descriptors (uncertain, strict, and boring). In four of the characteristics, faculty and GTA changes over the semester are exactly opposite from each other (confident, engaging, respect, and uncertain; Kendall and Schussler, 2013a). This is important because while students may stereotype instructors based on title at the beginning of the semester, there is room to accommodate changes in these perceptions over the course of the semester. This suggests instructors can shape student perception of them as instructors.

All of the instructor descriptors were significantly associated (moderately or strongly, except strict which was weak) with undergraduate perception of instructor teaching effectiveness, and all except strict were associated with perceived student learning. Positive associations (considered as effective teaching characteristics) included confident, organized, relaxed, engaging, relate, understanding, respect, relaxed, and enthusiasm. Negative associations (ineffective characteristics) included uncertain, nervous, distant, boring, and strict (Kendall and Schussler , 2013a). This once again indicates that these perceptions of instructors have significant meaning to undergraduates regarding the quality of instruction they feel they are receiving. Although it is unknown whether instructors with or without those characteristics are actually better or worse instructors, positive student attitude has been found to be correlated with student learning (Halandya, Olsen, and Shaughnessy, 1982).

Chapter Three asked undergraduates to explain the underlying behaviors that led to the instructor descriptions identified in Chapters One and Two. These detailed behavioral understandings allow for the identification of specific actions that instructors can take to modify their classroom teaching. For instance, they could choose to increase student learning by diminishing behaviors associated with ineffective instruction, and increasing behaviors associated with effective instruction. A schema generated from the results indicates that the descriptive words students use can generally be sorted into behavioral themes of teaching techniques, interpersonal rapport, and passion for subject. These three themes ultimately lead to an instructor being respected by their students (Kendall and Schussler, 2013b).

Lastly, **Chapter Four** determined that students find instructional behaviors related to the teaching techniques and interpersonal rapport themes as most important to GTA teaching effectiveness. Furthermore, it was documented that student perception of GTA teaching effectiveness is correlated with student perception of how much they learned in the laboratory, but not with student perception of their grade in lab. Consequently, recommendations for short professional development sessions about teaching techniques (e.g., content presentation and active learning strategies) and interpersonal rapport (e.g., relating to students and managing policies) were made in order to promote effective instruction in introductory courses by GTAs.

Recommendations for GTA Professional Development

Since the focus of this dissertation was the Division of Biology at UTK, and specifically on GTAs who teach biology courses in the Division, it is now possible to make tailored GTA professional development recommendations for this institution. Currently, GTA professional development consists of a half day pre-academic year new Biology GTA orientation, and a half day University GTA orientation covering policies. GTAs also attend weekly prep meetings for the laboratory they teach and there are optional graduate level courses in education (offered Spring semester; ranging from 1-3 credit hours).

Since each lab course that a GTA teaches currently holds a weekly prep meeting, it would be ideal to hold microteaching (short practice teaching session) and practice grading sessions (grading sample good and poor assignments as a group and discussing scores) at these meetings so that GTAs could receive feedback from their peers and lab

coordinators. In addition to these weekly meetings, GTAs and the lecture instructor associated with their course could hold meetings at least monthly to discuss the content of the course, as well as concerns or problems with students or course policies. GTAs should also be required to conduct observations of at least three instructors each semester and volunteer to conduct peer evaluations of fellow GTAs. In addition to these suggestions, each GTA can undertake self-reflection and maintain a written portfolio/journal of their teaching experience which they could discuss with an experienced GTA mentor (assigned at the beginning of each semester).

The offering of a one credit hour seminar course on teaching science led by experienced GTAs may also benefit new GTAs by providing them a venue to hear new teaching ideas and discuss difficulties they are experiencing. This seminar could also discuss articles selected from educational journals, such as *Journal of College Science Teaching*, *International Journal of Science Education*, or *CBE-Life Science Education*, chosen by fellow GTAs or faculty members to expand GTA knowledge regarding teaching techniques.

Sessions focused on specific teaching topics (i.e., grading rubrics, writing quizzes, etc.), and co-teaching opportunities could be offered through the Division of Biology and led by either experienced GTAs, faculty members, or program coordinators (as discussed in Chapter Four). Co-teaching with guidance and mentoring from an experienced faculty members or more senior GTA may help GTAs identify aspects of their teaching that they can modify over their teaching assignment, while specific workshops and/or weekend retreats could help standardize teaching from the diverse GTAs teaching a variety of

courses for the Division of Biology. GTAs wishing to enhance their curriculum development skills could also be more purposely integrated into curriculum design tasks by their departments or faculty advisors as a means to help increase GTA confidence, organization, and knowledge about what they teach. Furthermore, the formation of a GTA professional development council, consisting of faculty members and GTAs, would be beneficial in identifying both departmental needs as well as GTA development needs to help shape teaching assignments and professional development opportunities.

Instructor ability to relate well to students is also important for instruction and student learning. To foster positive relationships between GTAs and undergraduates, a "meet and greet" event could be organized so that undergraduates could meet their GTA prior to the semester. Professional development could encourage GTAs to share information regarding their educational background, personal interests and hobbies, and favorite things (such as organism, ice cream flavor, cartoon character) as a means for undergraduates to identify with GTAs. GTAs could also self organize research symposia/poster sessions during which they present their research to undergraduates who would like to gain insight into GTA research projects. Sessions could also be organized for GTAs, faculty members, and program coordinators to convene and exchange thoughts regarding course policies and accommodation options for students with difficulties that could emphasize instructor characteristics of flexibility and empathy.

GTA professional development can benefit from the schema developed in Chapter Three by using student evaluations, observations and self-reflection fro GTAs to identify their teaching strengths and weaknesses specific to the three behavioral themes. For

example, analysis may indicate that a GTA knows the content well and relates to students well, but that they are not good at explaining the content. This can then be the focus of professional development activities for this GTA. These behavioral themes could also be used in role-playing activities in which GTAs must recognize positive and negative behaviors associated with each theme.

While these suggestions may not be applicable to all institutions, they can serve as starting ideas for others to modify into their own GTA professional development programs to foster better instruction and learning. Other institutions should consider their culture as well as available resources to guide which changes will be most beneficial to their overall programs, students, and instructors. Many GTAs will be future faculty members in the higher education system (Austin, 2002); by providing GTAs with insight about the important synergy between the personal and instructional realms, they will be better prepared to educate the scholars of tomorrow.

Literature Cited

Austin AE (2002). Preparing the next generation of faculty: graduate school as socialization to the academic career. J High Educ *73*, 94–122.

Halandya T, Olsen R, Shaughnessy J (1982). Relations of student, teacher, and learning environment variables to attitudes toward science. Sci Educ *66*, 671-687.

Kendall KD, Schussler EE (2012). Does instructor type matter? Undergraduate student perception of graduate teaching assistants and professors. CBE Life Sci Educ *11*, 187–199.

Kendall KD, Schussler EE (2013a). Evolving impressions: Undergraduate perceptions of graduate teaching assistants and faculty members over a semester. CBE-Life Sciences Educ *12*.

Kendall KD, Schussler EE (2013). More than words: Probing the terms undergraduate students use to describe their instructors. Int J Teach and Learn in Higher Educa *17*.

Muzaka V (2009). The niche of graduate teaching assistant (GTAs): perceptions and reflections. Teach High Educ *14*, 1–12.

Park C (2002). Neither fish nor fowl? The perceived benefits and problems of using graduate teaching assistants (GTAs) to teach undergraduate students. High Educ Rev *35*, 50–62.

VITA

Denise Kendall is the daughter of Christina Kendall; she was born in Neu-Ulm, Germany. After graduating from Karns High School in 2005, Denise attended Pellissippi State Technical Community College pursing a major in the Biological Sciences. Upon the completion of her Associates of the Arts in 2007, she was admitted to the Bachelors program with a concentration in Ecology and Evolutionary Biology Honors at the University of Tennessee. Throughout her time in the undergraduate program at the University of Tennessee, Denise participated in numerous research projects involving cave-dwelling crayfish, plant physiology, and edible dormice. Denise completed her Bachelors of Science Degree in 2009, after which she was accepted into the Philosophy of Sciences Degree program in Ecology and Evolutionary Biology with a research program concentrated on teaching and learning.