Helping Graduate Teaching Assistants in Biology use Student Evaluations as Professional Development

Helping Graduate Teaching Assistants in Biology Use Student Evaluations as Professional Development

K. DENISE KENDALL, MATTHEW L. NIEMILLER, DYLAN DITTRICH-REED, ELISABETH E. SCHUSSLER

Abstract

Graduate teaching assistants (GTAs) are often used as instructors in undergraduate introductory science courses, particularly in laboratory and discussion sections associated with large lectures. These GTAs are often novice teachers with little opportunity to develop their teaching skills through formal professional development. Focused self-reflection about end-of-semester teaching evaluations may be an important informal supplement to teacher training. To inform this practice, we explored the instructional behaviors that undergraduates perceived as most important for GTAs' teaching effectiveness in laboratory courses. In spring semester 2012, 1159 undergraduates in freshman-level biology lab courses rated their GTAs on 21 instructional behaviors, the GTAs' teaching effectiveness, the amount the student learned, and their expected grade in the laboratory. Using linear mixed models, we found that instructional behaviors related to the categories of teaching techniques and interpersonal rapport best predicted student ratings of GTAs' teaching effectiveness. GTAs or other novice teachers can use this information to identify specific areas for instructional improvement when considering student feedback about their teaching.

Key Words: *Biology education; teaching effectiveness; instructional behaviors.*

○ Introduction

With the national movement to increase students' learning and retention in the sciences, it is important to consider the multiple factors that influence science classrooms, including course instructors, curricula, and student perceptions. Here, we focus on student perceptions of the teaching effectiveness of biology graduate students serving as graduate teaching assistants (GTAs), who often teach subsections (e.g., lab and discussion) of large introductory courses in postsecondary institutions (Rushin et al., 1997; Sundberg et al., 2005).

Individual instructors can have positive impacts on student retention, attitudes, and learning success (Hartnett et al., 2003; O'Neal et al., 2007; Kneipp et al., 2010). Given that GTAs in introductory science courses typically have more personal contact with students than faculty do (Bond-Robinson & Rodriques, 2006; Baumgartner, 2007), they are likely to be extremely influential in undergraduate gateway courses. The nature of these influences can vary, however, as evidenced by reports that students learn little from GTAs who teach lab sections (Sunal et al., 2004) and reports that GTAs positively affect student retention in the sciences (O'Neal et al., 2007). GTAs are clearly important factors in undergraduate teaching and learning (Dotger, 2010).

Although they are often novice instructors, GTAs – and science GTAs in particular – receive minimal pedagogical support, training, or continuous mentoring during their graduate tenure (Rushin et al., 1997; Luft et al., 2004; Tanner & Allen, 2006). Many universities offer orientations or seminars before the academic year begins (at university and/or departmental levels) or distribute training manuals (Lowman & Mathie, 1993; Rushin et al., 1997) as their only form of professional development, while some universities offer formal education courses or mentoring (Boyle & Boice, 1998; Baumgartner, 2007; Lockwood et al., 2014). However, departmental and university policies are often the focal point of much GTA training, and as a result, GTAs typically receive limited pedagogical information (Luft et al., 2004; Tanner & Allen, 2006).

GTA professional development for instruction is crucial, given that 85% of GTAs report they do not feel adequately trained for their teaching assignments (Russell, 2009). However, because science GTAs are often graduate students in research programs, the time and resources available for formal teacher preparation are often limited. The addition of informal self-reflection is one manner in which GTAs could improve their teaching without expanding existing professional development programs (Boyle & Boice, 1998). Many schools have informal peer groups of GTAs who discuss teaching as one means of support for GTA teaching efforts. Given the existence of these peer groups and the value of self-reflection as a way to improve teaching, we aimed to identify

> *GTAs – and science GTAs in particular – receive minimal pedagogical support, training, or continuous mentoring during their graduate tenure.*

a small set of instructional behaviors that could be the focus of this reflection for teaching improvement. This could promote the teaching abilities of these novice instructors even in the context of universities with little time or few resources for extensive, formal professional development for GTAs.

○ Project Rationale

Kendall and Schussler (2012, 2013a) documented students' perceptions of the teaching behaviors of GTAs and faculty members, and the present study built on this work by identifying which GTA behaviors were most important for students' perceptions of biology GTAs' teaching effectiveness. Given that most student attrition occurs in the first academic year (Seymour & Hewitt, 1992; PCAST, 2012) and GTA-led class sections are often associated with introductory science courses (Rushin et al., 1997, Sundberg et al., 2005), we focused on students' perceptions of GTAs teaching first-year (100-level) biology courses. Moreover, introductory courses were chosen for study because the majority of biology GTAs at this university teach introductory lab sections. We utilized student evaluations of GTAs' teaching effectiveness (rather than peer or supervisor evaluations) because end-of-semester evaluations are the most commonly used feedback mechanism for this program, and because they are a proxy for student satisfaction about the teaching of these courses. The goal of our study was to identify instructional behaviors, specific to biology GTAs, that could be the basis for novice teachers' self-reflection about teaching practices.

○ Methods

Data Collection

Data were collected using an online survey (hosted by surveymonkey.com) administered to students in GTA-instructed biology lab sections in spring 2012 at a research-intensive university in the southeastern United States. The survey was administered in several semester-long (15 weeks) nonmajors (second-semester Introductory Biology) and majors (Biodiversity and Cellular Biology) introductory biology laboratories. These labs have predetermined syllabi and curricula managed by a central program with faculty lab coordinators. All the labs were taught by GTAs, who typically teach two (3-hour) or three (2-hour) lab sections per semester with a maximum class size of 27 students. GTAs are responsible for instructing the lab sections and for grading all of their students' lab assignments, including quizzes, handouts, reports, presentations, and exams. GTAs meet once a week with lab coordinators to prepare for the upcoming exercises, and they typically have little interaction with the lecture portion of the course.

Survey Design

The survey consisted of Likert-choice responses to questions in three categories. The questions and categories (which we refer to as "sub-scales"; Kendall & Schussler, 2012, 2013b) were identified through student surveys and interviews as being important to instructor respect and student learning: teaching techniques, interpersonal rapport, and passion for subject. Each sub-scale consisted of seven questions, each focused on one instructional behavior (Table 1; Kendall & Schussler, 2013b). The survey instructed participants to

Table 1. The 21 instructional behavior items included in the survey, sorted into their respective subscales (Teaching Techniques Index, Interpersonal Rapport Index, and Passion for Subject Index, as defined in Kendall & Schussler, 2013b), as well as the teaching-effectiveness and student-learning survey items. Students responded to each item using a Likert response scale described in the methods section.

A. Teaching Techniques Index (TTI)
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.
B. Interpersonal Rapport Index (IRI)
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.
C. Passion for Subject Index (PSI)
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.
D. Teaching Effectiveness and Student Learning
The instructors' effectiveness in teaching the material.
The amount you learned in the course.

indicate their agreement with each item using a six-point response scale: (1) strongly agree, (2) agree, (3) slightly agree, (4) slightly disagree, (5) disagree, and (6) 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: (1) very poor, (2) poor, (3) fair, (4) good, (5) very good, and (6) excellent; the amount they learned on a six-point response scale of 1–6, with 1 representing "little" and 6 representing "a lot"; and their

performance by recording the letter grade they expected to receive in the lab (A, B, C, D, F). 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, held in the last full week of the semester. If the survey was given as part of the final exam, students were offered one point of the exam's 55 points to complete it. This was accomplished by having the GTA check off the question after the student visited the computer with the online survey. If the student chose not to take the survey or to skip any questions, they still received the point. Therefore, this survey was voluntary for all students, it did not have a measureable impact on their grade, and the GTAs did not see any student responses. All procedures were reviewed and approved by the institutional review board.

Survey Reliability

To determine the reliability of the survey, Cronbach's alpha was calculated for each of the sub-scales (SPSS version 19.0). The Cronbach's alpha values for teaching-techniques, interpersonal-rapport, and passion-for-subject sub-scales were 0.943, 0.948, and 0.948, respectively. Cronbach's alpha was also calculated for each GTA, and these results ranged from 0.794 to 0.989 for teaching techniques, from 0.831 to 0.987 for interpersonal rapport, and from 0.860 to 0.990 for passion for subject. From these results it was judged that the survey was reliable.

○ Data Analysis

Data Coding

Responses to 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, and strongly disagree = 1. The index for each sub-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, and 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, specific biology concentrations = 2–4 [Biochemistry, Cellular, and Molecular Biology = 2, Ecology and Evolutionary Biology = 3, and Microbiology = 4], or 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. Students' 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, and excellent = 6.

Statistical Analyses

To analyze the survey responses, we fit a series of generalized linear mixed models (GLMMs) to the data; GLMMs are an extension of the generalized linear model that also includes random effects in addition to fixed effects. We treated student demographics and sub-scale indices (Likert score sum for each sub-scale; teaching-technique index [TTI], interpersonal-rapport index [IRI], and passion-for-subject index [PSI]) as fixed effects and fit random intercepts for section, GTA, and course. The full model included GTAs' 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 with different combinations of significant (α = 0.05) fixed effects (Table 2). 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 on the basis of Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC) scores. Models were

Table 2. Statistical information for the full and reduced (A1–A6) models tested as part of the study. For each model, the "Fixed Effects" column shows what aspects of the data set were being tested. The model with the highest log likelihood and lowest AIC and BIC is the best model.

Model	Fixed Effects	df	Log Likelihood	AIC	BIC
Full	TTI + IRI + PSI + Gender + Major + Eng + TotalBio + Course + Year + Grade + Learned	15	−798.050	1626.11	1697.07
A1	TTI + IRI + PSI + Grade + Learned	10	−790.640	1601.29	1648.60
A2	TTI + IRI + PSI + Learned	9	−790.680	1599.37	1641.95
A3	TTI + IRI + Grade + Learned	9	−790.510	1599.01	1641.59
A4	**TTI + IRI + Learned**	8	−790.084	**1596.17**	**1634.02**
A5	TTI + Learned	7	−800.270	1614.54	1647.65
A6	TTI	6	−824.776	1661.55	1689.94

Notes: The random intercepts were always the same (course, GTA, and section), but the degrees of freedom (df) varied by the number of factors being tested.

compared by examining the difference in AIC and BIC scores between models (ΔAIC and ΔBIC, respectively). All analyses were conducted in R version 2.15.

○ Results

Participants

The survey was administered to a potential pool of 1881 undergraduates in 86 laboratory sections. After removing students who did not complete the survey in its entirety, 1159 undergraduates remained (62%). This response rate was a result of the survey being voluntary or of student absences during administration of the surveys. Participants were primarily female (60.1%) freshman (56.7%) non–biology majors (79.0%) who spoke English as their native language (96.3%). Second- and third-year students comprised 23.6% and 11.0%, respectively, of the respondents, with 8.6% being fourth year or beyond. Most of the respondents were currently enrolled in a nonmajors second-semester Introductory Biology course (56.1%), while the rest were currently enrolled in majors' Biodiversity (21.7%) or Cell Biology (22.2%) courses. The majority of respondents had completed at least one other semester-long biology lab course (71.8%), including 49.9% that had completed the nonmajors first-semester Introductory Biology course, 18.2% the Biodiversity course, and 2.9% the Cell Biology course.

Generalized Linear Mixed Models

The 1159 complete survey responses came from undergraduates taught by 39 GTAs. To determine the instructional sub-scale (TTI, IRI, or PSI) that best predicted undergraduate perception of GTAs' teaching effectiveness, we first removed 319 responses that had no variation in sub-scale indices (e.g., all were "strongly agree"). Data from one GTA were identified as outliers on the basis of Cook's D in preliminary analysis without random intercepts, and were therefore removed from the data set. The minimum number of student responses for a GTA was 5, the maximum was 53, and the mean was 30 responses with a standard deviation of 10.31.

A reduced model with TTI, IRI, and Learned as fixed effects best predicted GTAs' effectiveness, based on both AIC and BIC scores (ΔAIC = 2.84, ΔBIC = 7.57; Table 2). Although students' ratings of both GTAs' passion for subject (PSI) and estimated grade (Grade) predicted students' perceptions of GTAs' teaching effectiveness in the full model ($P = 0.007$ and $P = 0.016$, respectively), the best-fitting model, as determined by AIC and BIC, did not include these variables. 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. Students associated GTAs' teaching effectiveness with the amount they learned (Pearson's $r = 0.671$, $P < 0.001$), not their expected grade (Pearson's $r = 0.019$, $P = 0.911$).

○ Discussion

Our results showed that students' perceptions of GTAs' teaching effectiveness were predicted best by instructional behaviors representative of the teaching-techniques and interpersonal-rapport sub-scales and students' perceptions of how much they had learned.

Teaching-techniques instructional behaviors include being calm, being interesting, keeping students' attention, using good examples, making the material relevant to students, being interactive, and being good at presenting the material. The interpersonal-rapport sub-scale includes instructional behaviors of being compassionate, empathetic, flexible, laid-back, approachable, connecting with students, and having a good relationship with students. Consequently, these are the instructional behaviors that GTAs should focus on in informal self-reflection regarding student evaluations.

It is important to note that students in this study put equal importance on both formal classroom instructional (teaching technique) and informal relationship (interpersonal rapport) aspects when evaluating the instruction of their GTA. The importance of classroom and relationship aspects has also been identified by research exploring effective instruction at other education levels (e.g., K–12) and in other disciplines (e.g., business) (Varca & Pattison, 2001; Evans, 2002; Arnon & Reichel, 2007; Helterbran, 2008). Novice teachers such as GTAs may focus their efforts solely on teaching techniques and not realize that students put a high value on the interactions they have with their teacher outside of formal instruction. As a result, novice instructors should carefully monitor student feedback regarding their perceived rapport with students, in addition to feedback about their actual teaching practices.

To encourage informal self-reflection on student feedback, programs could organize peer-mentoring groups, consisting of novice GTAs paired with more experienced GTAs, with or without a staff or faculty leader. A successful example of peer-mentoring in biology is described by Lockwood et al. (2014). In this program, experienced GTAs participated in peer-mentoring training and then met with their novice GTA mentees throughout the first semester of their teaching to assist with reflective practice. Boyle and Boice (1998) suggested that assigned mentor pairs yield more instructional improvement than naturally formed pairs.

Based on the feedback from the present study, GTAs should work together to identify and sort student responses into those related to teaching techniques and interpersonal rapport. The initial peer-mentoring reflection should be on whether students' perceptions of each were generally positive, neutral, or negative. The next stage of reflection should break down the student feedback within each category into specific behavioral aspects or words that students used to describe them. Undergraduate interviews conducted by Kendall & Schussler (2013b) illuminated that a myriad of instructional behaviors can contribute to a single term being used to describe an instructor (e.g., "strict" can mean tough grading but also adhering to rules). The peer mentors would then discuss the interpretation of their students' ratings and remarks, using the education literature to assist in these interpretations (e.g., Kendall & Schussler, 2013b). This step should identify specific strengths and weaknesses of their instruction, as perceived by students, and the peer mentors should work together to identify ways to address the identified weaknesses.

Although our results provide compelling evidence for the factors that students perceive as important to GTAs' teaching effectiveness, this study was conducted with volunteer participants from one discipline at a single institution of higher education. How students defined "teaching effectiveness," as well as "learning," is unknown to us; future research should clarify students' definitions of these terms. Yet, even with these unknowns, this information provides a starting

point for GTAs to better analyze their end-of-semester evaluations. A follow-up study should investigate whether GTA self-reflection on these aspects results in teaching modifications and subsequent improvement in students' perceptions of their teaching effectiveness.

Fostering improved GTA instruction in science departments can promote positive classroom experiences for students, which in turn should promote student learning, retention, and positive attitude toward the sciences. However, GTAs must be provided with the necessary support to promote their teaching effectiveness and maximize their students' learning in introductory science courses. This study provides awareness of the aspects of teaching that undergraduates put a premium on, and how this can be used to improve a GTA's, or any novice teacher's, reflection on their teaching.

○ Acknowledgments

References

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 teachers. *Teachers and Teaching, 13,* 441–464.

Austin, A.E. (2002). Preparing the next generation of faculty: graduate school as socialization to the academic career. *Journal of Higher Education, 73,* 94–122.

Austin, A.E., Campa, H., III, Pfund, C., Gillian-Daniel, D.L., Mathieu, R. & Stoddart, J. (2009). Preparing STEM doctoral students for future faculty careers. *New Directions in Teaching Learning, 17,* 83–95.

Baumgartner, E. (2007). A professional development teaching course for science graduate students. *Journal of College Science Teaching, 36,* 16–21.

Bond-Robinson, J. & Rodriques R.A.B. (2006). Catalyzing graduate teaching assistants' laboratory teaching through design research. *Journal of Chemical Education, 83,* 313–323.

Boyle, P. & Boice, B. (1998). Systematic mentoring for new faculty teachers and graduate teaching assistants. *Innovative Higher Education, 22,* 157–179.

Dotger, S. (2010). Offering more than "Here is the textbook": teaching assistants' perspectives on introductory science courses. *Journal of College Science Teaching, 39,* 71–76.

Evans, J.F. (2002). Effective teachers: an investigation from the perspectives of elementary school students. *Action in Teacher Education, 24,* 51–62.

Hartnett, N., Römcke, J. & Yap, C. (2003). Recognizing the importance of instruction style on students' performance: some observations from laboratory research – a research note. *Accounting Education, 12,* 313–331.

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

Kendall, K.D. & Schussler, E.E. (2012). Does instructor type matter? Undergraduate student perception of graduate teaching assistants and professors. *CBE–Life Sciences Education, 11,* 187–199.

Kendall, K.D. & Schussler, E.E. (2013a). Evolving impressions: undergraduate perceptions of graduate teaching assistants and faculty members over a semester. *CBE–Life Sciences Education, 12,* 92–105.

Kendall, K.D. & Schussler, E.E. (2013b). More than words: probing the terms undergraduate students use to describe their instructors. *International Journal of Teaching and Learning in Higher Education, 25,* 200–212.

Kneipp, L.B., Kelly, K.E., Biscoe, J.D. & Richard, B. (2010). The impact of instructor's personality characteristics on quality of instruction. *College Student Journal, 44,* 901–905.

Lockwood, S.A., Miller, A.J. & Cromie, M.M. (2014). Preparing future biology faculty: an advanced professional development program for graduate students. *American Biology Teacher, 76,* 17–21.

Lowman, J. & Mathie, V. (1993). What should graduate teaching assistants know about teaching? *Teaching of Psychology, 20,* 84–88.

Luft, J.A., Kurdziel, J.P., Roehrig, G.H. & Turner, J. (2004). Growing a garden without water: graduate teaching assistants in introductory science laboratories at a doctoral/research university. *Journal of Research in Science Teaching, 41,* 211–233.

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 trainng. *Journal of College Science Teaching, 36,* 24–29.

Pattison, P., Hale, J.R. & Gowens, P. (2011). Mind and soul: connecting with students. *Journal Legal Studies Education, 28,* 39–66.

President's Council of Advisors on Science and Technology. (2012). Engage to excel: producing one million additional college graduates with degrees in science, technology, engineering, and mathematics. Report to the President. Available online at http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-engage-to-excel-final_2-25-12.pdf.

Rushin, J.W., De Saix, J., Lumsden, A., Streubel, D.P., Summers, G. & Bernson, C. (1997). Graduate teaching assistant training: a basis for improvement of college biology teaching & faculty development? *American Biology Teacher, 59,* 86–90.

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

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

Sunal, D.W., Wright, E.L. & Day, J.B. (Eds.) (2004). *Science Teaching for the 21st Century.* Greenwich, CT: Information Age Publishing.

Sundberg, M.D., Armstrong, J.E. & Wischusen, E.W. (2005). A reappraisal of the status of introductory biology laboratory education in U.S. colleges & universities. *American Biology Teacher, 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 Sciences Education, 5,* 1–6.

Varca, P.E. & Pattison, P. (2001). Attaining teaching excellence: a critical incident technique. *Journal of Legal Studies Education, 19,* 153–185.

K. DENISE KENDALL is a laboratory coordinator at the School of Integrative Biology, 286 Morrill Hall, MC-120, University of Illinois, 505 S. Goodwin Ave., Urbana, IL 61801; e-mail: kendalld@illinois.edu. MATTHEW L. NIEMILLER is a postdoctoral scholar at the Illinois Natural History Survey, 1816 S. Oak St., MC 652, Champaign, IL 61820. DYLAN DITTRICH-REED is an assistant professor in the College of Agriculture, Forestry, and Life Sciences, Clemson University, Clemson, SC 29634. ELISABETH E. SCHUSSLER is an associate professor in the Department of Ecology and Evolutionary Biology and director of teaching and learning in the Division of Biology, University of Tennessee, Knoxville, TN 37996.