Graduate TA Teaching Behaviors Impact Student Achievement in a Research-Based Undergraduate Science Course

By Ariana S. Huffmyer and Judith D. Lemus

Student experience and achievement in undergraduate science courses is important for degree retention and generating interest in careers in STEM. Graduate student teaching assistants (GTAs) provide a majority of instruction in STEM laboratory courses at universities and play a critical role in the educational experience of students. Many GTAs receive limited pedagogical training, and there is limited understanding on the influence of instructional methods used by GTAs on student academic achievement. This study examines the influence of GTA teaching behaviors on student achievement in a research-based undergraduate science course. Several aspects of GTA teaching behavior significantly influence student scores in homework, quiz, and research project areas. Behaviors including pacing, wait time, and discussion of homework are significant positive predictors of student scores, whereas teaching experience is a negative predictor. The influence of GTA questioning strategy is different for homework and research project areas, underscoring the complex relationship between teaching behaviors and student achievement. Results suggest that universities should provide support for GTAs to gain skills in teaching and pedagogical knowledge prior to instructing science laboratory courses.

Increasing the number of students in science majors and fostering pursuit of science careers is important to national interests (President's Council of Advisors on Science and Technology, 2012). In critical points along the education pathway, many students lose interest in science and chose not to pursue careers in science, technology, engineering and math (STEM). In the first 2 years of undergraduate education, students complete lower division laboratory courses in which they learn scientific research skills, especially if exposed to inquiry-based experiences. Experience and achievement in these early science courses is important for retention and interest in STEM (Benjamin, 2002; Jaeger, 2008). At research universities, graduate teaching assistants (GTAs) provide instruction to a majority of these courses, leading over 90% of science laboratory sections (Sundberg, Armstrong, & Wischusen, 2005). GTAs often have more direct contact with undergraduates than faculty, as they manage the laboratory environment for a small group of students on a regular basis. This interaction allows GTAs to play a critical role in student education and have a strong influence on their experience (Bond-Robinson & Bernard Rodriquez, 2006).

This influence may be positive when GTAs are perceived as approachable, relatable, enthusiastic, and knowledgeable of the learning needs of diverse student groups (O'Neal, Wright, Cook, Perorazio, & Purkiss, 2007; Park, 2002). There is also a potential for inexperienced GTAs to negatively influence student experience when they lack confidence and content knowledge or have limited teaching skills (Muzaka, 2009). Although GTAs are responsible for a large portion of instruction, there is a deficiency in resources for GTA professional development at research universities (Luft, Kurdziel, Roehrig, & Turner, 2004). Training programs range from half-day orientation sessions to multiday trainings (Roehrig & Luft, 2003) but are often generalized and may not address learning theories, instructional techniques, or pedagogical preparation (Luft et al., 2004; Schussler, Read, Marbach-Ad, Miller, & Ferzli, 2015). Without sufficient training in these areas, GTAs are likely to construct their instructional methods based on cultural influences, past experience, or untested beliefs about teaching (Kagan, 1992; Roehrig & Luft, 2003; Young & Bippus, 2008). These experiences and implicit biases may lead to differences in GTA teaching behaviors.

Past research has characterized the effectiveness of several teaching behaviors on student achievement. However, there is little knowledge on the teaching behaviors and strategies used by GTAs and which behaviors either positively or negatively influence student achievement. For example, a recent study (Fisher & DeChenne, 2015) found that the way in which questions were asked in the classroom had an impact on student scores and overall discussion in the classroom. In addition, Turpen and Finkelstein (2009) found that when physics professors provided longer wait times, there was an increase in student discussion. Longer wait times have also been associated with higher quality responses, increased student questions, and higher confidence (Rowe, 1986). Differences in the amount of time instructors spend discussing incorrect student answers has also been positively associated with student achievement (Elawar & Corno, 1985; Turpen & Finkelstein, 2009). The extent to which instructors establish relevance between material and real-life concepts impacts learning and increases student interest (Kember, Ho, & Hong, 2008). However, much remains unknown about the use of these instructional techniques by GTAs in university classrooms and the influence on undergraduate student achievement. An understanding of GTA teaching behaviors is necessary for improvement of undergraduate instruction through GTA training and professional development.

The purpose of our study is to investigate how GTA teaching behaviors influence student achievement in a research-based undergraduate science laboratory course. We characterized GTA teaching behaviors through observations and explored the relationship between these behaviors and student achievement measured as student scores.

Methods

Study participants

GTAs and undergraduate students in a lower division biology laboratory course at the University of Hawaii at Manoa (UHM) participated in this study. All research activities were approved under the UHM Internal Review Board for Human Subjects Research as exempt under the Code of Federal Regulations, 45CFR 46.101(b) and included appropriate informed consent. The laboratory course was held for three hours each week with 11 course sections and was accompanied by a parallel lecture course. The purpose of the laboratory course is to introduce students to the fundamentals of ecology and evolution while allowing students to develop critical thinking skills. The course is required for life science majors and is offered to nonmajors with a lower level general biology course required as a prerequisite. Each GTA served in a primary and secondary role as part of their teaching duties. While in the primary role, the GTA led all classroom activities, discussions, and grading. In the secondary role, GTAs assisted with classroom management but did not lead the class in activities.

The course curriculum included weekly homework assignments, short in-lab quizzes, and a written and oral research project presentation. Homework assignments (including short answer and problem-solving questions) were designed by the laboratory coordinator and were uniform across sections. GTAs were provided standard grading keys. GTAs had the flexibility to write their own weekly in-lab quizzes using multiple-choice, true/false, short-answer, essay, or problem-solving questions. In the second half of the semester, students developed their own original research questions and projects using collected data on the distribution and abundance of plants and invertebrates on Waahila Ridge near UHM. Under GTA guidance, students generated hypotheses, analyzed data, wrote a 10- to 12-page scientific report, and presented results in a 15-minute oral presentation. Research projects were graded by GTAs using a standard rubric developed by the laboratory coordinator. The inquiry-based nature of this research project allows for measurement of achievement in an area that requires higher order thinking skills compared with lower order thinking skills required in quiz and homework assignments. Cumulative scores for this course were weighted as 50% homework assignments, 15% in-lab quizzes, 25% research projects, and 10% participation.

Of the 10 GTAs participating in this study, three were pursuing master's degrees and seven were pursuing doctoral degrees at UHM. Two held a master's degree and eight held bachelor's degrees. Previous experience ranged from zero to nine semesters of teaching prior to the semester in which our study took place. GTAs completed the UHM teaching assistant training as required by UHM. This training takes place over two days and includes a choice of informational sessions on teaching (teaching a science lab, syllabus design, pedagogy), student perspectives, grading, and classroom management. GTAs attended a one-day orientation led by the laboratory coordinator that covered information on course objectives and grading practices. Each week, the coordinator held 1-hour meetings to discuss assignments and activity procedures.

Of the 153 students that participated in this study, 106 were in a science-degree track, 35 in health disciplines, five in social sciences, and one undecided. Six students did not report their degree track. Forty-four students reported a GPA between 3.5 and 4.0, 65 reported 3 to 3.5, 37 reported 2.5 to 3, six reported 2 to 2.5, and one reported 0.0 to 2.0. Eighty-one students were in their second year of undergraduate study, 46 in their third year, 19 in their fourth year, and seven in their fifth year.

GTA teaching behaviors rubric

We developed a GTA observation rubric following preliminary observation of GTAs and discussion with the laboratory course coordinator about observed differences in teaching behavior. Following this discussion, items were prioritized for further exploration based on a published rubric for teacher observation (Colorado Department of Education, 2016). GTA observation rubric items were scored on a five-point Likert scale or by counting the number of actions performed by the GTA based on behaviors documented in previous literature that impact student learning and/or achievement (Table 1). A subset of GTA observation rubric items were chosen for further analysis in this study. Each GTA in this course was observed twice during the semester in one activity-based and one lecture-based laboratory session. During the activity-based session, the GTA led students through an activity but did not present a lecture. In the lecture-based session, the GTA gave a 15- to 25-minute lecture followed by a short activity or discussion. Observations of each primary GTA were completed by the first author and one other graduate student (not involved in this course) or faculty member (including the second author), all of whom participated in observation calibration training. In this calibration, observers were instructed to independently watch a video of a GTA leading a laboratory course and scored their behaviors on the observation rubric. Discrepancies in scores greater than one point on the Likert scale were discussed to standardize scoring. During GTA observations, the difference in observer responses was no greater than one point on the Likert scale for each item. GTA observation scores were calculated as the average of observer scores on each item.

Quiz question analysis

Each week, primary GTAs were instructed by the laboratory coordinator to write short quizzes on any

TABLE 1

Description of graduate teaching assistant (GTA) teaching observation rubric items. Scoring type specifies mechanism of measuring rubric items.

Observation rubric item	Scoring type	Brief description	Literature reference
Questioning strategy	Likert and counting	Type/number of questions posed by GTA. Ranges between majority closed questions (one desired/correct answer) and majority open questions (multiple answers open to interpretation).	Fisher & DeChenne, 2015
Wait time	Likert	Time GTA waits after posing a question for students to answer.	Honea, 1982; Rowe, 1986; Turpen & Finkelstein, 2009
Homework discussion	Likert	Amount of discussion in class on past homework problems.	Elawar & Corno, 1985; Turpen & Finkelstein, 2009
Pacing	Likert	Pace of instruction throughout lectures and activities.	Rowe, 1986; Tincani et al., 2005
Relatable concepts	Likert	GTA discussion of ways class concepts relate to student life and community.	Kember, Ho, & Hong, 2008
Survey for understanding	Counting	Number of times GTA checks survey for understanding.	Fisher & Frey, 2014

topic(s) covered in the lab manual worth five points. GTAs constructed these quizzes independently based on their choice of material. To characterize differences in the way GTAs constructed quizzes, each question was assigned a numerical value based on the level of revised Bloom's Taxonomy (Bloom, 1956) incorporated in the question as described (Anderson & Krathwohl, 2001; IUPUI Center for Teaching and Learning, 2002). Quizzes were removed of identifiers and scored in random order. Fifteen out of 120 (12.5%) of the quiz questions were removed from the analysis because they incorporated multiple levels of the revised Bloom's Taxonomy or the appropriate score was unclear. The score given for each question increased with each level: Remembering (1 point), Understanding (2 points), Applying (3 points), Analyzing (4 points), Evaluating (5 points), and Creating (6 points). A quiz index was calculated from the proportion of questions from each level weighted by the corresponding number assigned to each level. A quiz index of 6 indicates all questions were from the "Creating" level, whereas an index of 1 indicates all questions were from the "Remembering" level.

Student grades

Student cumulative scores in homework, quizzes, and research projects were collected from the laboratory coordinator with student names and identifiers removed following the completion of the course.

Data analysis

The effects of GTA teaching behaviors on homework scores and project scores were analyzed using generalized linear model (GLM) Akaike information criterion (AIC) backwards model selection in the MASS package in R statistical programming software (version 1.0.143). Effect plots of significant predictors were generated with the effects package. Student homework and research project scores were transformed using log and square root functions when necessary to meet assumptions of GLM analyses. GLM analysis is a useful framework for comparing how several predictor variables affect a continuous response variable. GLM backwards model selection starts with fitting a model with all predictor variables and subsequently dropping the least significant variable until reaching a simplified model in which all predictor variables are significant. AIC selection compares this set of models and identifies the best fitting model for further analysis. Predictor and response variables are shown in Table 2. There were no significant correlations between predictor variables as determined by nonparametric Spearman rank correlation tests. Quizzes were not graded using a standard practice, so generating predictive modeling was not appropriate for our study, and were instead analyzed using the Pearson product-moment correlation analysis.

TABLE 2

Predictor and response variables in GLM analysis.

Predictor variables	Response variables
Questioning strategy	
Wait time	
Homework discussion	Homework scores
Pacing	Research project scores
Relatable concepts	
Survey for understanding	
GTA semesters teaching experience	

Note: GLM = generalized linear model; GTA = graduate teaching assistant.

Results

Description of student scores

Aggregated mean, standard error of aggregated mean, and minimum and maximum values of student scores in quiz, homework, and research project

TABLE 3

Mean, standard error of mean, minimum values, and maximum values of student scores in the laboratory course. Maximum possible point values were: quiz (5 points), homework (35 points), and research project (130 points).

Category	*M*	*SE*	Minimum	Maximum
Quiz	3.582	0.047	1.333	4.875
Homework	27.890	0.233	16.481	33.846
Research project	97.207	1.024	56.500	128.000

categories are reported in Table 3. Quiz and homework scores were log-transformed and project scores were square-root-transformed to meet assumptions of analyses.

Effects of GTA behaviors on scores

Homework scores

The effect of GTA teaching behaviors on student homework scores was analyzed with GLM selection (Table 4). Questioning strategy (β = –0.082, p = .018), pacing (β = –0.135, p = .022), survey for understanding (β = –0.045, p = .014), and semesters of teaching experience (β = 0.019, p = .023) were significant predictors of student homework scores (R^2 = 0.092, df = 192, p < .001), with the model predicting 9.2% of variance (Table 5). Because of log-transformation of scores, a positive beta estimate indicates a negative effect of predictor on student scores and a negative beta estimate indicates a positive effect of predictor on student scores. Effect plots show visual representation of the effects of significant predictors on untransformed student homework scores (Figure 1).

Research project scores

The effect of GTA teaching behaviors on student research project scores was analyzed with GLM selection (Table 6). Questioning strategy (β = 0.487, p = .017), wait time (β = –0.558, p = .001), homework discussion (β = –0.254, p = .005), and relatable concepts (β = –0.515, p = .049) were significant predictors of student research project scores (R^2 = 0.051, df = 192, p = .007), with the model predicting 5.1% of variance (Table 7). Because of transformation of scores, a positive beta estimate indicates a negative effect of predictor on student scores, and a negative beta estimate indicates a positive effect of predictor on student scores. Effect plots show visual representation of the effects of significant predictors on untransformed student research project scores (Figure 2).

Quiz scores

GTA quiz index scores were low in this study, with an average index score of 1.784 ± 0.114 (mean ± standard

TABLE 5

Estimated predictors from AIC selected generalized linear model of log-transformed student homework scores.

Predictor	Beta estimate	*SE*	*t*-value	*p*-value
Questioning strategy	–0.082	0.034	–2.396	.018*
Pacing	–0.135	0.059	–2.305	.022*
Survey for understanding	–0.045	0.018	–2.487	.014*
Semesters of teaching experience	0.019	0.008	2.285	.023*

Note: AIC = Akaike information criterion. Positive beta estimate indicates a negative effect of predictor on student scores. Negative beta estimate indicates a positive effect of predictor on student scores.
*p < .05.

TABLE 4

AIC generalized linear model selection table of log-transformed student homework scores. Bold indicates model chosen for further analysis.

Model	Predictors	Degrees freedom	Log likelihood	AICc	Delta AIC
HW0	Questioning Strategy + Homework Discussion + Pacing + Survey Understanding + Semesters Experience + Wait Time + Relatable Concepts	9	–82.415	183.8	5.91
HW1	Questioning Strategy + Homework Discussion + Pacing + Survey Understanding + Semesters Experience + Wait Time	8	–82.476	181.7	3.84
HW2	Questioning Strategy + Pacing + Survey Understanding + Semesters Experience + Wait Time	7	–82.666	179.9	2.04
HW3	**Questioning Strategy + Pacing + Survey Understanding + Semesters Experience**	**6**	**–82.719**	**177.9**	**0.00**

Note: AIC = Akaike information criterion.

error), indicating that most quiz questions were in the "Remembering" and "Understanding" levels of the revised Bloom's taxonomy. The relationship between quiz index and quiz scores was analyzed using a Pearson product-moment correlation test (Figure 3). There was a significant positive correlation between quiz index and quiz scores, $r(101) = 0.3119$, $t = -3.2987$, $p = .0013$.

Discussion

GTA teaching behaviors had a small but significant effect on student scores in this undergraduate science laboratory course. Homework, quiz, and research project scores are affected by different aspects of GTA teaching behavior.

The unique positive predictors for homework scores were pacing and surveying for understanding. There is a positive relationship between homework scores and pace of instruction. Although GTAs delivered content at different paces, they were provided the same material to teach in each class, so students in slower paced sections were presented with the same core content. As documented in previous research, a rapid pace may increase student attention (documented in learning disabled students by Darch & Gersten, 1985), providing an

FIGURE 1

Effect plot of significant predictors of untransformed student homework scores. Shading indicates 95% confidence interval.

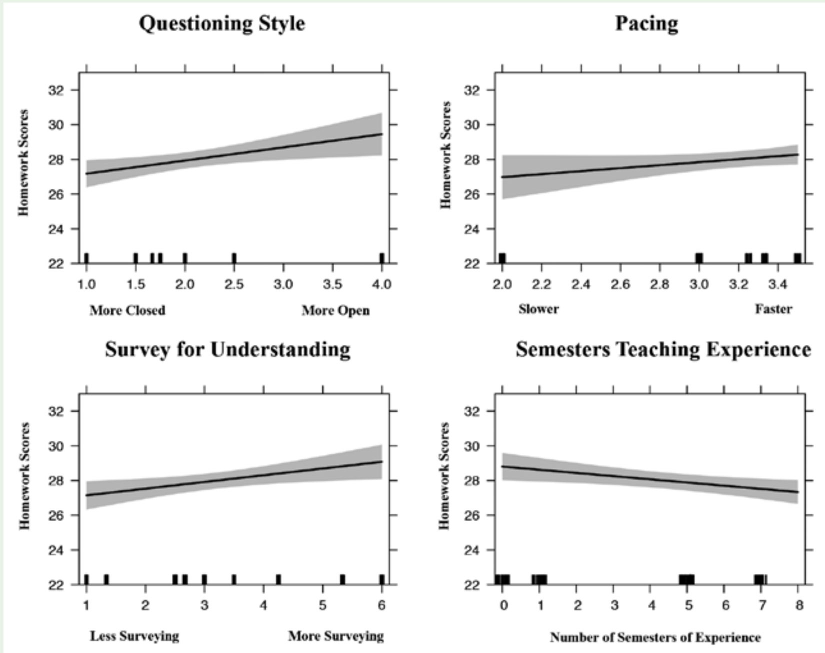


TABLE 6

AIC generalized linear model selection table of square-root-transformed student research project scores. Bold indicates model chosen for further analysis.

Model	Predictors	Degrees freedom	Log likelihood	AICc	Delta AIC
RP0	Questioning Strategy + Homework Discussion + Pacing + Survey Understanding + Semesters Experience + Wait Time + Relatable Concepts	9	–333.800	686.6	4.97
RP1	Questioning Strategy + Homework Discussion + Pacing + Semesters Experience + Wait Time + Relatable Concepts	8	–333.837	684.4	2.85
RP2	Questioning Strategy + Homework Discussion + Semesters Experience + Wait Time + Relatable Concepts	7	–333.901	682.4	0.80
RP3	**Questioning Strategy + Homework Discussion + Wait Time + Relatable Concepts**	**6**	**–334.574**	**681.6**	**0.00**

Note: AIC = Akaike information criterion.

TABLE 7

Estimated predictors from AIC selected generalized linear model of square-root-transformed student research project scores.

Predictor	Beta estimate	*SE*	*t*-value	*p*-value
Questioning strategy	0.487	0.20210	2.407	.017*
Wait time	−0.558	0.15880	−3.511	.001**
Homework discussion	−0.254	0.08922	−2.843	.005**
Relatable concepts	−0.515	0.26041	−1.979	.007**

Note: AIC = Akaike information criterion
Positive beta estimate indicates a negative effect of predictor on student scores.
Negative beta estimate indicates a positive effect of predictor on student scores.
$^*p < .05$. $^{**}p < .01$

opportunity for greater comprehension. However, a rapid pace may not accommodate every learning style, so GTAs should receive training on how to modify pace of instruction for diverse learning styles. There is also a positive relationship between the frequency of GTA surveying for understanding and homework scores. Pausing to verbally check for understanding in the classroom allows the GTA to gauge comprehension and revisit material as necessary. Checking for understanding helps instructors identify learning goals, provide feedback, and plan instruction to address student misconceptions (Fisher & Frey, 2014). The results of our study suggest that GTAs should increase surveying for understanding to increase student achievement.

GTA teaching experience is a significant negative predictor of homework scores but is not correlated to other predictor variables. This result contradicts the expectation that as an instructor gains experience, they become more effective, refine their practices, and increase student learning as documented previously (Clotfelter, Ladd, & Vigdor, 2006; Ladd & Sorenson, 2017; Priesto & Altmaier, 1994). In this study, student homework scores were lower in class sections with more experienced GTAs. Teacher burnout and emotional exhaustion has been negatively associated with student achievement (Klusmann, Richter, & Ludtke, 2016) and may be a factor influencing the decline in student scores in sections with more experienced GTAs. The lack of correlation between teaching behaviors and experience suggests that there are other behaviors associated with experience that we did not measure, including the GTAs strictness in grading. We are limited in the interpretation of the influence of teaching experience on student achievement because this value does not include detailed information on past training, both formal and informal, that GTAs received.

The unique positive predictors of research project scores were wait time, discussion of homework, and relatable concepts. Student participation in a research project requires higher level thinking skills to construct hypotheses, analyze data, and draw conclusions. Our results suggest that several aspects of GTA teaching behavior influence the development of these higher level skills. First, the positive effects of increased wait time on student scores supports previous research. Increased wait time following teacher questioning has been associated with achievement, particularly for higher level questions (Wilen & Clegg, 1986). Longer wait times are also associated with increased student discussion (Turpen & Finkelstein, 2009) and generation of student questions (Rowe, 1986). Second, the amount of discussion that GTAs provided on homework is a positive predictor for research project scores. Past work has found this type of constructive discussion increases student achievement (Elawar & Corno, 1985), which may encourage students to develop higher thinking skills in an inquiry-based learning environment. Third, the use of relatable concepts in instruction is a positive predictor of research project scores. As suggested by Kember et al. (2008), the incorporation of relevant concepts to local issues and everyday life may spark student interest in research, increasing motivation and promoting achievement as documented in our study. Our results suggest GTAs should increase wait times and feedback discussions along with incorporating relatable concepts in their classrooms to encourage higher level student thinking. These alterations in behavior may be simple for GTAs to incorporate in their teaching practices when provided with proper training opportunities.

Questioning strategy was the only predictor in common for homework and research project scores, but contrary to our expectations, it shows an opposing relationship. Previous work has found that an instructor's questioning strategy plays a key role in shaping classroom discussion and can influence student learning and engagement (Almeida & Neri de Souza, 2010). Open, high-level questions used in the classroom are more likely to increase uptake in students' ideas (Wells & Arauz, 2006), to increase student achievement (Redfield & Rousseau, 1981), and to facilitate productive discussion (Nystrand, Wu, Gamoran, Zeiser, & Long, 2003; Wells & Arauz, 2006; Zhang, Lundeberg, McConnell, Koehler, & Eberhardt, 2010) than low-level questions with predetermined answers. These findings are supported by the positive relationship between student homework scores and use of open questions in our study. However, we documented a negative relationship between use of open questions and research project scores. Other studies have also found a positive correlation between low-level questions and greater student achievement (Wilen & Clegg, 1986), and even a negative learning effect related to high-level questions (Soar & Soar, 1979), illustrating the complex relationship between questioning strategy and student achievement. Lower level and closed questions may be more effective in conveying factual knowledge and information to students. It is unclear whether a similar mechanism may have increased student memorization of specific scientific writing procedures, statistics, or other factual information that led to higher scores on research projects. However, based on these combined results, we suggest GTAs receive pedagogical training on the effective and appropriate use of a combination of open- and closed-questioning strategies in classroom instruction.

FIGURE 2

Effect plot of significant predictors of untransformed student research project scores. Shading indicates 95% confidence interval.

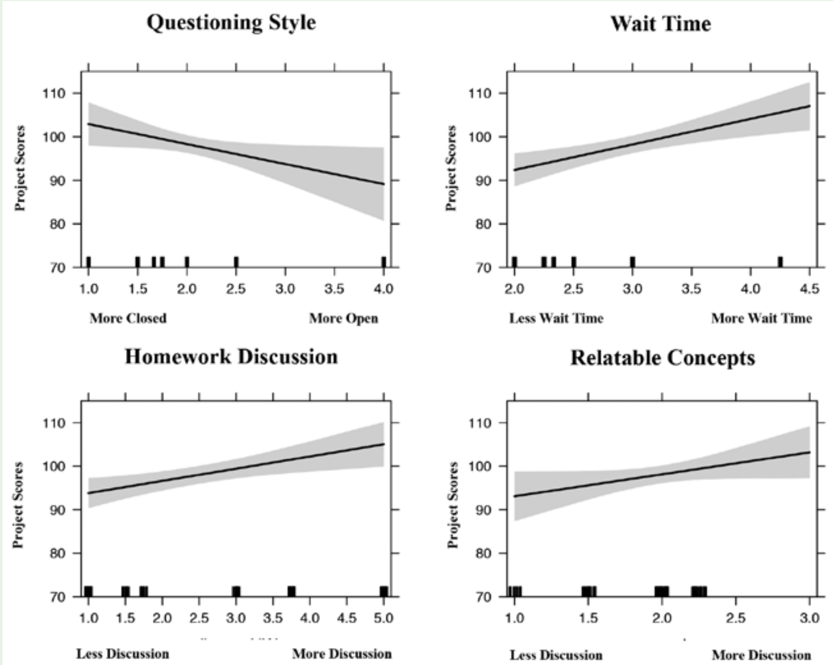
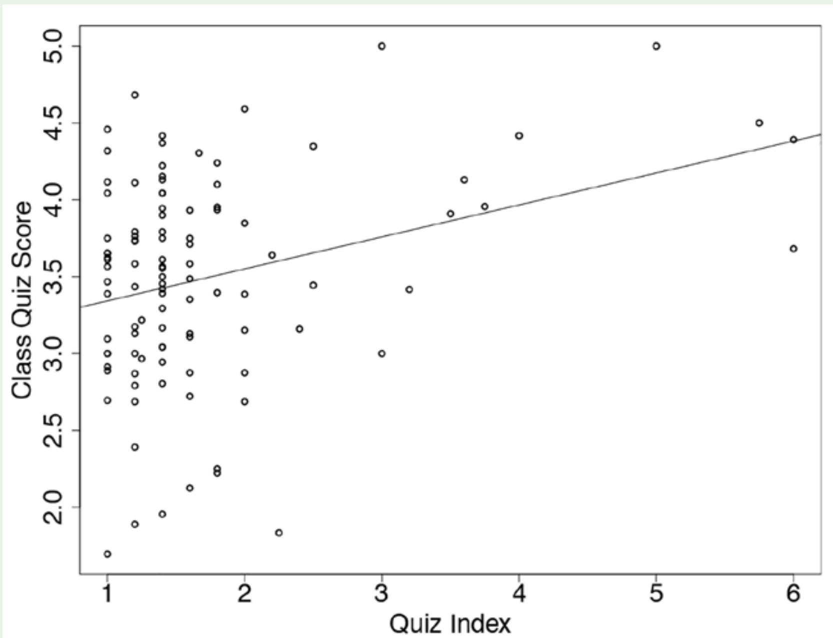


FIGURE 3

Correlation between quiz index and average student quiz scores.



The construction of quizzes by GTAs is another reflection of their questioning strategy. The results of correlation analyses (Figure 3) indicate that students received higher scores on quizzes with higher order, open questions. We are limited in separating student learning gains from quiz difficulty, as higher order questions offer more flexibility in answers than lower order questions. The differences in question type may also influence GTA grading practice, because higher order questions allow for increased leniency. Considering these limitations, we recommend that GTAs receive training on administration of low-point value assignments to provide a consistent experience for students across course sections.

Our results highlight the complex relationship that exists between teaching behaviors and student achievement. Capturing information on students' experience in this course, including the perceived accessibility and effectiveness of their GTA and conceptual learning gains, would add valuable information in future studies. In addition, measuring GTA perceptions of their teaching ability and pedagogical views would provide insights on factors leading to differences in teaching behaviors. Future GTA training opportunities should include a discussion on effective teaching behaviors and practices to encourage achievement across a spectrum of learning styles. Trainings should also highlight the implicit biases and beliefs held by GTAs to encourage recognition of effects of their instructional methods on student learning and achievement. Increasing the quality of GTA training opportunities at universities is an important but overlooked component of improving academic experience and encouraging retention of undergraduate students in STEM. ■

Acknowledgments

We thank Dr. Anuschka Faucci of the UHM Biology Department for her help in the development and execution of this research. Dr. Joanna Philippoff in the UHM College of Education provided valuable input on the survey and observation design. We also thank the graduate teaching assistants and undergraduate students who participated in this research and the volunteer faculty and graduate students who assisted with teaching observations. Finally, we greatly appreciate the comments from the three anonymous reviewers for helping to improve the manuscript. This is HIMB contribution #1733 and SOEST contribution #10420.

References

Almeida, P., & Neri de Souza, F. (2010). Questioning profiles in secondary science classrooms. *International Journal of Learning and Change, 4,* 237–251.

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York, NY: Longman.

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

Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals*. New York, NY: Longman.

Bond-Robinson, J., & Bernard Rodriquez, R. A. (2006). Catalyzing graduate teaching assistants' laboratory teaching through designed research. *Chemical Education Research, 83,* 313–323.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2006). Teacher-student matching and the assessment of teacher effectiveness. *Journal of Human Resources*, *41,* 778–819.

Colorado Department of Education. (2016). *Rubric for evaluating Colorado's teachers*. Denver, CO: Author.

Darch, C., & Gersten, R. (1985). The effects of teacher presentation rate and praise on LD students' oral reading performance. *Educational Psychology, 55,* 295–303.

Elawar, M. C., & Corno, L. (1985). A factorial experiment in teachers' written feedback on student homework: Changing teacher behavior a little rather than a lot. *Journal of Educational Psychology, 77,* 162–173.

Fisher, D., & Frey, N. (2014). *Checking for understanding: Formative assessment techniques for your classroom*. Alexandria, VA: ASCD.

Fisher, G. R., & DeChenne, S. E. (2015). Presenting clicker questions with an open versus closed response format. *Journal of Microbiology & Biology Education, 16,* 254–255.

Honea, J. M. (1982). Wait-time as an instructional variable: An influence on teacher and student. *The Clearing House, 56,* 167–170.

IUPUI Center for Teaching and Learning. (2002). *Bloom's taxonomy "revised" key words, model questions, & instructional strategies*. Indianapolis, IN: Indiana University–Purdue University. Retrieved from https://coe.uni.edu/sites/default/files/wysiwyg/BloomRevisedTaxonomy.pdf

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

Kagan, D. M. (1992). Professional growth among pre-service and beginning teachers. *Review of Educational Research*, *62*, 129–169.
Kember, D., Ho, A., & Hong, C. (2008). The importance of establishing relevance in motivating student learning. *Active Learning in Higher Education*, *9,* 249–263.
Klusmann, U., Richter, D., & Ludtke, O. (2016). Teachers' emotional exhaustion is negatively related to students' achievement: Evidence from a large-scale assessment study. *Journal of Educational Psychology, 108*, 1193–1203.
Ladd, H. F., & Sorenson, L. C. (2017). Returns to teacher experience: Student achievement and motivation in middle school. *Education Finance and Policy, 12,* 241–279.
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.
Muzaka, V. (2009). The niche of graduate teaching assistant (GTAs): Perceptions and reflections. *Teaching in Higher Education, 14,* 1–12.
Nystrand, M., Wu, L. L., Gamoran, A., Zeiser, S., & Long, D. A. (2003). Questions in time: Investigating the structure and dynamics of unfolding classroom discourse. *Discourse Processes, 35,* 135–198.
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. *Journal of College Science Teaching*, *36*(5), 24–29.
Park, C. (2002). Neither fish nor fowl? The perceived benefits and problems of using graduate teaching assistants (GTAs) to teach undergraduate students. *Higher Education Review, 35,* 50–62.
President's Council of Advisors on Science and Technology. (2012). *Report to the President: Engage to excel: Producing one million additional college graduates with degrees in science, technology, engineering, and mathematics*. Washington, DC: White House.
Priesto, L. R., & Altmaier, E. M. (1994). The relationship of prior training and previous teaching experience to self-efficacy among graduate teaching assistants. *Research in Higher Education*, *35,* 481–497.
Redfield, D. L., & Rousseau, E. W. (1981). A meta-analysis of experimental research on teacher questioning behavior. *Review of Educational Research*, *51,* 237–245.
Roehrig, G. H., & Luft, J. A. (2003). Graduate teaching assistants and inquiry-based instruction: Implications for graduate teaching assistant training. *Journal of Chemical Education*, *80*, 1206–1210.
Rowe, M. B. (1986). Wait-time: Slowing down may be a way of speeding up. *Journal of Teacher Education, 1,* 37–68.
Schussler, E. E., Read, Q., Marbach-Ad, G., Miller, K., & Ferzli, M. (2015). Preparing biology graduate teaching assistants for their roles as instructors: An assessment of institutional approaches. *CBE—Life Sciences Education*, *14*(3), ar31.
Soar, R., & Soar, R. (1979). Emotional climate and management. In P. Peterson & H. Walberg (Eds.), *Research on teaching: Concepts, findings, and implications.* Berkeley, CA: McCutchan.
Sundberg, M. D., Armstrong J. E., & Wischusen, E. W. (2005). A reappraisal of the status of introductory biology laboratory education in U.S. colleges and universities. *American Biology Teacher,* 67, 525–529.
Tincani, M., Ernsbarger, S., Harrison, T., & Heward, W. L. (2005). Effects of two instructional paces on pre-K children's participation rate, accuracy, and off-task behavior. *Journal of Direct Instruction, 5,* 97–109.
Turpen, C., & Finkelstein, N. D. (2009). Not all interactive engagement is the same: Variation in physics professors' implementation of peer instruction. *Physics Education Research*, *5,* 020101.
Wells, G., & Arauz, R. M. (2006). Dialogue in the classroom. *Journal of the Learning Sciences, 15*, 379–428.
Wilen, W. W., & Clegg, A. A., Jr. (1986). Effective questions and questioning: A research review. *Theory and Research in Social Education*, *14,* 153–161.
Young, S. L., & Bippus, A. M. (2008). Assessment of graduate teaching assistant (GTA) training: A case study of a training program and its impact on GTAs. *Communication Teacher, 22,* 116–129.
Zhang, M., Lundeberg, M., McConnell, T. J., Koehler, M. J., & Eberhardt, J. (2010). Using questioning to facilitate discussion of science teaching problems in teacher professional development. *Interdisciplinary Journal of Problem-Based Learning, 4*, 57–82.

Ariana S. Huffmyer (ashuff@hawaii.edu) *is a PhD candidate in Marine Biology and* ***Judith D. Lemus*** *is a science education specialist, both at the Hawaii Institute of Marine Biology at the University of Hawaii at Manoa.*