

STATISTICS GRADUATE TEACHING ASSISTANTS' BELIEFS, PRACTICES AND PREPARATION FOR TEACHING INTRODUCTORY STATISTICS

NICOLA JUSTICE
University of Minnesota
njustice@umn.edu

ANDREW ZIEFFLER
University of Minnesota
zief0002@umn.edu

JOAN GARFIELD
University of Minnesota
jbg@umn.edu

ABSTRACT

Graduate teaching assistants (GTAs) are responsible for the instruction of many statistics courses offered at the university level, yet little is known about these students' preparation for teaching, their beliefs about how introductory statistics should be taught, or the pedagogical practices of the courses they teach. An online survey to examine these characteristics was developed and administered as part of an NSF-funded project. The results, based on responses from 213 GTAs representing 38 Ph.D.-granting statistics departments in the United States, suggest that many GTAs have not experienced the types of professional development related to teaching supported in the literature. Evidence was also found to suggest that, in general, GTAs teach in ways that are not aligned with their own beliefs. Furthermore, their teaching practices are not aligned with professionally-endorsed recommendations for teaching and learning statistics.

Keywords: *Statistics education research; graduate teaching assistant development; teacher practice; teacher beliefs*

1. INTRODUCTION

Graduate students are hired to teach or assist courses in many higher education institutions across the world. Their titles and roles vary across countries, institutions, and departments. In Canada, *graduate teaching fellows* have the primary responsibility for designing and administering courses, whereas *graduate teaching assistants* typically lead tutorials or lab sections, grade assigned work, hold office hours, and monitor course websites (Hoessler & Godden, 2015). At Lancaster University in the United Kingdom, *graduate teaching assistants* (GTAs) in science departments were found primarily to facilitate labs and fieldwork, whereas those in social sciences and humanities typically led discussion groups (Park & Ramos, 2002). According to Barrington (2001), GTAs in New Zealand and Australia are primarily referred to as *graduate tutors*, although Legrand (2001) describes training programs in France for doctoral candidates called *moniteurs* whose funding is provided in exchange for teaching services. In the United States, as far back as 1968, Nowlis, Clark, and Rock described GTAs roles, which may be

as extensive as instructors (with more or less complete responsibility for an entire course), or as minimal as graders or clerical workers who interact with students hardly at all. For the purposes of this study, we will use the term *graduate teaching assistant* (GTA) to refer to graduate students who are hired to teach or assist with the teaching of courses in any of the capacities described above.

Concerns have been raised about the preparation of GTAs to fulfill their roles. In a study at the University of Sheffield in the UK, Muzaka (2009) found that students and professors have concerns that GTAs lack content knowledge and confidence for teaching. Barrington (2001) cites a study reporting student discontent with the diverse quality and attitudes of GTAs at the University of Melbourne. GTAs across the United States have voiced concern about their own lack of preparation and support for teaching (e.g., Fagen & Suedkamp Wells, 2004; Golde & Dore, 2001). There is evidence to suggest that students enrolled in courses taught by GTAs are less likely to enroll in subsequent courses or to choose a major in a field related to the course matter than students enrolled in courses taught by full-time faculty members (Bettinger & Long, 2004).

Despite such concerns, authors from many countries have suggested that financial pressures on institutions will continue to promote sustained reliance on GTAs for teaching responsibilities (e.g., Barrington, 2001; Muzaka, 2009; Birch & Morgan, 2005). In the United States, GTAs carry a heavy load of the introductory statistics courses. A report by Blair, Kirkman, and Maxwell (2013) found that GTAs are responsible for teaching about 25% of introductory courses in Ph.D.–granting statistics departments in the United States. GTAs will likely continue to have a large influence on undergraduate statistics courses in the United States and abroad. Also, as pointed out by Barrington (2001), not all GTAs will become professors, but the future statistics professorate is likely largely contained in the set of all GTAs. Their current and potential influence on the discipline leads to questions about how statistics GTAs are prepared and supported for their teaching roles, as well as about their teaching practices and beliefs.

This paper describes the design and results of the *Graduate Student Statistics Teaching Inventory* (*GSSTI*), a survey developed in order to begin to establish some empirical basis regarding the preparation, teaching beliefs, and teaching practices of statistics GTAs in the United States. The *GSSTI* includes items about the types of development opportunities GTAs have experienced, the topics they have studied related to teaching introductory statistics, and the teaching pedagogies of the courses they currently teach or assist. The instrument also contains items regarding GTAs' beliefs about ideal pedagogy, assessment, and content for introductory statistics courses.

The *GSSTI*, which is the first survey of its kind and a work-in-progress, was administered to 213 statistics GTAs from 38 Ph.D.–granting statistics departments in the United States. The results were used to begin to answer questions about statistics GTAs' preparation for teaching introductory statistics and to explore the extent to which GTAs' beliefs and teaching practices are aligned with professionally endorsed recommendations for teaching introductory statistics.

2. BACKGROUND

Strategies for GTA training and development have been well documented. Conferences (e.g., Chism & Warner, 1987), edited books (e.g., Marincovich, Prostko, & Stout, 1998; Nyquist, Abbott, Wulff, & Sprague, 1991; Wulff & Austin, 2004), and entire journals (e.g., *Journal of Graduate and Professional Student Development*, established in 1993 under a different name) have been devoted to this topic. The literature is primarily descriptive and often includes examples that have been employed for training GTAs at a particular university or department (e.g., Boman, 2013; Davis & Minnis, 1993; Holmes, Martinuk, Ives, & Warren, 2013; Nyquist & Wulff, 1987; Speer, 2004).

Based on these descriptions, it is evident that there is a good deal of variation in the strategies used to develop GTAs as teachers. There is variation in the duration, scope, and content of GTA training and development (Parrett, 1987; Weimer, Svinicki, & Bauer, 1989). For example, some institutions provide GTAs with only a day- or week-long training, whereas others require their GTAs to attend weekly meetings, or enroll in semester- or year-long courses. GTA development programs may be intended for GTAs in any discipline (e.g., Wulff, Nyquist, & Abbott, 1991), or may be intended for GTAs in a specific discipline (e.g., Speer, 2004; Wyse, 2010). In addition, the content may be more generally related to future faculty preparation (e.g., Pruitt-Logan & Gaff, 2004), or may be more specifically related to preparation for a particular course. Also available are programs specifically designed to prepare international GTAs for their roles as assistants in universities outside of their native country (e.g., Constantinides, 1987).

To complement this descriptive literature, little empirical research has been conducted regarding the professional development of GTAs for teaching. The relatively few empirical studies generally fall into three areas: (1) survey studies, (2) research on GTA socialization, and (3) research on GTA development strategies. These areas will be reviewed in the next three sections.

2.1. SURVEY STUDIES

The first area of empirical research explores GTAs' preparatory experiences primarily through the use of survey studies. Some of the surveys collect data from GTAs at a single institution. These tend to investigate the types of preparatory opportunities experienced by GTAs and self-reported factors affecting their participation (e.g., Barrington, 2001). Other surveys are conducted across many institutions. Many are directed toward department chairs in specific disciplines in order to investigate the status and nature of GTA preparation in the particular field, such as mathematics (Belnap & Allred, 2009) or engineering (Torvi, 1994).

A few additional large-scale surveys in the United States have been commissioned to explore areas of need for GTA development across many institutions and disciplines (e.g., Fagen & Seudkamp Wells, 2004; Golde & Dore, 2001). Reports from these surveys tend to identify dissatisfaction among GTAs and department personnel regarding resources for GTA preparation.

2.2. RESEARCH ON GTA SOCIALIZATION

The second area of empirical research investigates the broader question of how graduate students acquire the values, skills, and knowledge—in short the culture—they need to serve as GTAs. This enculturation process is referred to as *socialization* (Merton, Reader, & Kendall, 1957).

One line of the literature on GTA socialization examines how and from whom GTAs acquire information. Studies in this area have found that GTAs look primarily to peers for information (Darling & Staton, 1989; Myers, 1998). In fact, interactions with peers have been found to be viewed as more helpful than interactions with professors, administrative personnel, mentors, campus wide training programs, department orientations, or weekly meetings (Darling, 1987; Myers, 1994; Williams & Roach, 1992; Wulff, Austin, Nyquist, & Sprague, 2004).

These findings suggest that it may be appropriate to study GTA professional development using approaches that account for peer interactions. One model for professional development that accounts for peer interactions is called *communities of practice*, first introduced by Lave and Wenger (1991). A few GTA development programs have begun to use community of practice models for GTA development (e.g., Crede, Borrego, & McNair, 2010; Holmes et al., 2013), but there are still many open questions regarding the extent to which the model is useful for establishing and studying GTA professional development related to teaching.

Another line of literature on GTA socialization suggests that the socialization process occurs in stages. For example, based on qualitative studies, Darling and Dewey (1990) describe three GTA socialization stages defined by GTAs' concerns: self, impact, and task. According to their theory, initial GTA concerns are often primarily at the self-level, when GTAs are too concerned with survival in the program (e.g., "Will I fail and drop out of the program?") to engage in deeper questions about teaching. As GTAs progress through their graduate programs, their concerns may evolve to questions about management of tasks and time (indicating GTAs are at the task-level), and eventually to questions about student learning and retention (impact-level). Based on similar observations of GTAs, Nyquist and Sprague (1998) suggest that training programs in earlier stages should include fairly prescriptive orientations, while professional development programs for GTAs in latter stages could be more reflective (see also Sprague & Nyquist, 1991). Ferzli et al. (2012) found evidence to suggest that well-designed training programs can help advance GTAs to higher stages.

However, concerns-based stage theory has undergone criticism, particularly from researchers in elementary and secondary teacher preparation fields. For example, very little evidence to support the theory was found in Watzke's (2007) longitudinal study of 79 pre-service teachers in the United States. Citing several other studies with longitudinal approaches that did not support stage-based theory, Watzke suggests that the concerns of teachers are more complex than accounted for by the self, task, and impact levels.

2.3. RESEARCH ON GTA DEVELOPMENT STRATEGIES

The third area of research explores the effects of different programs and strategies for GTA professional development related to teaching. Authors who have conducted reviews of this research have noted the relative lack of empirical evidence regarding the extent to which training and development program strategies can achieve desired outcomes (Abbott, Wulff, & Szego, 1989; Carroll, 1980).

Of the few empirical studies that have been conducted, many concerns arise. Some studies suffer from the use of outcome measures that are not well-supported (e.g., self-report surveys). Still others seem to measure outcomes immediately after an intervention, but do not address questions of whether the effects are retained over a meaningful period of time. Others use rigorous data collection methods (e.g., pre-post video analysis by experts) but are not very telling because they use designs that "bundle" interventions together in one overall program. It is uncertain whether the results of these studies were due to specific key interventions, or interactions between interventions, or interactions of interventions with other factors such as the environments or departments in which they were studied. Calls have been made for studies that isolate particular components of development programs instead of bundling them all together.

Among the mixed results regarding the effects of training and development strategies, two components appear promising: mentoring (regular meetings with faculty or senior GTAs) and teaching observations with feedback. These components of development programs have been found to improve GTAs' self-efficacy, reduce teaching anxiety, and change teaching beliefs and behaviors to become more student-centered. For example, in a controlled randomized study exploring teaching practices, Bray and Howard (1980) found an effect for GTAs who received feedback via video consultations, but no effect for GTAs who participated in a teaching seminar in addition to the video consultations. In another controlled randomized study Williams (1991) found that teaching observations combined with peer mentoring were able to reduce GTAs' anxiety and improve their teaching affect. Not surprisingly, GTAs have also rated teaching observations and mentoring as the most helpful types of professional development they have experienced (e.g., Dalgaard, 1982; Jones, 1993). Further support for the use of mentoring and

teaching observations has been found in studies of GTAs in STEM disciplines (e.g., Gilmore, Maher, Feldon, & Timmerman, 2014; Volkmann & Zgagacz, 2004).

2.4. RESEARCH ON GTAS IN STEM DISCIPLINES

Similar to the previously reviewed scholarship, most of the research literature focusing on GTAs in science, technology, engineering, and mathematics (STEM) disciplines is also descriptive in nature (e.g., Gilreath & Slater, 1994). Although these descriptions provide a snapshot of what GTA development looks like in the STEM disciplines, for a more critical examination, we look to the few empirical studies that have been conducted.

Empirical studies in STEM disciplines have examined the question of how GTAs' teaching is impacted by deficient content knowledge and pedagogical content knowledge. For example, Kung, Speer, and Gucler (2006) found that GTAs in mathematics exposed to only traditional teaching experiences could typically describe only a single problem-solving strategy to students. This finding suggests that the GTAs lacked pedagogical content knowledge to enable the facilitation of class discussion or assessment of open-ended student responses. In another study, Frank and Speer (2013) observed that poor content knowledge held by GTAs in physics limited their ability to interpret and score student responses correctly.

Empirical research also suggests that STEM GTAs' teaching beliefs and teaching practices may be affected by experiences with teaching mentors. When Gilmore et al. (2014) examined the relationships between GTAs' beliefs about teaching and four potentially related factors—degree of mentoring presence, prior teaching experience, prior research experience, and experience in preparatory training programs—they found that mentoring was the most influential of the factors for predicting student-centered teaching behaviors. Mentoring relationships were also explored in a case study by Volkmann and Zgagacz (2004). This study documented changes in a physics GTA's teaching practices during the course of an in-depth mentoring relationship between a GTA and a faculty member.

Unlike mentoring, the effects of training programs and workshops, overall, on impacting STEM GTAs' teaching beliefs and practices, are much less compelling. Many of the studies suffer from the methodological considerations described earlier in this paper. Other studies have found that training programs were not effective in changing STEM GTAs' beliefs and practices toward more student-centered and active learning methods (e.g., Rodrigues & Bond-Robinson, 2006; Roehrig, Luft, Kurdziel, & Turner, 2003). The lack of compelling results regarding the effectiveness of training and development programs may also be attributable to the short duration of most training programs, which tend to last for a few days, a couple weeks, or perhaps just one semester (e.g., Roehrig et al., 2003). Wyse (2010) found that biology GTAs only began to show detectable changes in beliefs after two semesters of an intensive training program.

2.5. STATISTICS GTA PREPARATION FOR TEACHING

The literature related to GTA development in the discipline of statistics is also primarily descriptive and anecdotal. Statistics departments appear to use varied methods to train and develop their GTAs into teachers. These methods include: weekly meetings (e.g., Birch & Morgan, 2005); mentoring (e.g., Froelich, Duckworth, & Stephenson, 2005); courses on teaching (e.g., Gelman, 2005; Harkness & Rosenberger, 2005); and immersion in a departmental culture (Birch & Morgan, 2005). In addition, Rumsey (1998) suggests that departments consider a collaborative approach to GTA development. Her strategy involves the use of weekly meetings in combination with a culture that is supportive, but not controlling. Rumsey's approach is inspired in part by Hogg (1991), who stated "it is clear that all of us would profit more if professors would serve as mentors to graduate students in teaching as well as research" (p. 343).

Several recommendations for preparing GTAs for teaching have also been published. In a special section of the *American Statistician*, Moore (2005) recommends that departments model a variety of interactive pedagogical methods, guide GTAs regarding the use of technology, and provide resources for GTAs to consult after the programs have commenced. He also recommends that departments create avenues of support for GTAs by fostering community among GTAs, providing mentoring opportunities, holding regular meetings, and giving GTAs feedback after observations.

Garfield and Everson (2009) concur with Moore's (2005) recommendations, describing a GTA development course that is organized around the *Guidelines for Assessment and Instruction in Statistics Education College Report (GAISE)* (American Statistical Association [ASA], 2005). The *GAISE College Report* proposes six recommendations endorsed by the ASA for teaching introductory statistics courses at the college level: emphasize statistical literacy and develop statistical thinking; use real data; stress conceptual understanding, rather than mere knowledge of procedures; foster active learning in the classroom; use technology for developing concepts and analyzing data; use assessments to improve and evaluate student learning. Garfield and Everson suggest that statistics GTA preparatory courses should have a broad focus that will develop GTAs' ability to design courses and select technology. They also warn practitioners not to overestimate GTAs' content knowledge.

The empirical studies that have been conducted to study GTAs in statistics provide evidence to support many of the recommendations that have been made for statistics GTA preparation. Noll (2011) surveyed 68 GTAs from 18 institutions and found that lack of content knowledge hindered the GTAs' ability to design quality instruction and make judgments about the reasonableness of students' answers. She suggests that statistics GTAs be given more experience with empirical data to accompany their theoretical knowledge of statistics. In another study, Green (2010) collected qualitative evidence to suggest that statistics GTAs desire more faculty guidance with regard to pedagogical content knowledge and the use of technology.

There is a growing body of literature offering descriptive examples of development programs for GTAs and making recommendations for those programs. Although some literature exists about GTA development in statistics departments, it is primarily focused on just a handful of universities and is largely anecdotal. In general, there is little empirical information about how the majority of GTAs in statistics departments are currently being prepared to teach, about their beliefs about teaching, or their teaching practices. Research in these areas could inform the design of opportunities to develop GTAs as teachers.

To learn about GTAs' professional development experiences, teaching beliefs, and teaching practices, a study was designed to answer the following research questions for GTAs in Ph.D.–granting statistics departments in the United States:

1. What professional development experiences for teaching do statistics GTAs have?
2. What are GTAs' beliefs about the teaching of introductory statistics?
3. To what extent are GTAs' teaching beliefs aligned with their teaching practices?
4. To what extent are GTA's teaching practices aligned with professionally endorsed recommendations for teaching introductory statistics?

3. METHODS

To gather data to answer the aforementioned research questions, the *Graduate Student Statistics Teaching Inventory (GSSTI)*, an online survey instrument, was created as part of the NSF-funded *e-ATLAS Project* (NSF DUE-1044812 & 1043141). The *GSSTI* is designed to measure statistics GTAs' beliefs, practices, and preparatory experiences for teaching introductory statistics. The instrument is composed of 46 items—44 of which are forced-choice—divided into four sections (see Table 1). The instrument also includes a free-response item to allow

participants to provide any additional comments or information they wish to share, as well as an item that allows participants to provide any additional comments they wish to make. The complete instrument is presented in Appendix A.

3.1. INSTRUMENT DEVELOPMENT

To create the GSSTI, two graduate students and one faculty member reviewed the *Statistics Teaching Inventory (STI)* (Zieffler, Park, Garfield, delMas, & Bjornsdottir, 2012) for items that could be adapted for GTAs. The *STI* is an instrument developed as part of the NSF-funded *e-ATLAS Project* to measure the beliefs and practices of introductory statistics instructors at the tertiary level. Using the *GAISE College Report* (ASA, 2005) to inform the content of the *STI*, items were written and revised as the instrument was reviewed by experts, piloted by members of the Consortium for the Advancement of Undergraduate Statistics Education (CAUSE), and examined in think-aloud sessions designed to expose whether items were interpreted as intended. Table 1 shows the number of items in each section of the *GSSTI* that were adapted from the *STI*.

Table 1. The GSSTI Instrument: Distribution of Items Across Sections.

| Section title | Number of Items | | |
|---|-------------------------|-----|-------|
| | Adapted from <i>STI</i> | New | Total |
| 1. Your Teaching Experience | 2 | 6 | 8 |
| 2. Your Pedagogy Practices | 7 | 0 | 7 |
| 3. Your Beliefs about Teaching Statistics | 11 | 1 | 12 |
| 4. Your Development as a Teacher | 0 | 18 | 18 |
| (Final item for participant comments) | 1 | | |
| Total | 21 | 25 | 46 |

As can be seen from Table 1, many of the *GSSTI* items were adapted from the *STI*. These adaptations primarily consisted of minor adjustments to wording to make them appropriate for graduate-level instructors and GTAs. For example, *STI* items related to teachers' classroom practices were modified to ask GTAs whether they had learned about those teaching practices. This type of item was implemented for content related to use/learning about technology tools and pedagogical methods (e.g., methods for helping students learn together in small groups; principles of designing activity-based lessons).

After collecting and modifying existing *STI* items, a review of the *GSSTI* content suggested several opportunities for expanding the item pool. To this end, additional items were written for each section of the *GSSTI*. Table 1 shows the number of new items written for each of the sections in the *GSSTI*. These items were primarily based on content and themes identified in the reviewed literature. For example, additional items were written to obtain information about GTAs' developmental experiences, especially those identified as important in the research literature, for example, faculty mentoring (Gilmore et al., 2014) and observations with feedback (Williams, 1991). Additional items were written based on our own experiences working with GTAs and anecdotes shared with us by GTAs working in other academic programs and universities. For example, in informal conversations with GTAs, we learned that there seemed to be much variation in their roles and responsibilities. Subsequently, we wrote items to examine the scope of GTAs' responsibilities and their autonomy to make decisions about course content, teaching methods, and assessment in the courses they teach or assist.

After the initial instrument was assembled, all of the items were revised and modified according to guidelines recommended by Dillman, Smyth, and Christian (2010) to reduce potential measurement error. For example, items were examined for inclusion of balanced

response options. Some check-all-that-apply items were rewritten as forced-choice. Lastly, response options for all items related to GTAs' teaching beliefs were modified to include an "undecided" response option in order to accommodate GTAs who might be undecided about their beliefs about teaching.

Items were also modified to reduce survey respondents' cognitive load. Where possible, item stems were shortened, and items with similar structures were grouped to share a common stem. Items were also sequenced within the instrument to reduce respondents' cognitive burden. For example, items related to GTAs' responsibilities were presented in order of ascending responsibility. A free-response item was also included in order to gather additional information that participants were compelled to communicate. The *GSSTI* was administered in an online format using *Survey Gizmo* (2013).

Pilot tests and think-aloud interview The survey was piloted with three GTAs who had either served as instructors or teaching assistants for introductory statistics courses, one of whom was an international GTA. The pilot participants were GTAs at the researchers' current institution, but they were asked to take the survey based on their experiences as graduate students in other statistics departments that they attended prior to their current institutions. These think-aloud participants represented three different statistics departments in US Midwestern institutions. With the international pilot participant, a think-aloud interview was conducted as she piloted the survey. This interview helped illuminate whether the student was interpreting the survey items as intended. Modifications were made when clarifications seemed necessary.

Based on the feedback received from the pilot and think-aloud participants, many items were modified. For example, the stem of one item was broken up into shorter passages to clarify response instructions. Also, to reduce participants' confusion, items that distinguished between GTAs' teaching development opportunities specific to statistics and those that were more general were combined no longer to address teaching development opportunities specific to the discipline of statistics.

3.2. DATA COLLECTION

The target population for the survey was graduate students enrolled in a Ph.D.-granting statistics department in the United States who had taught or assisted with an introductory statistics course. Because the research was focused on describing the future statistics professorate, we also included graduate students who had the potential to teach or assist introductory statistics courses in the target population. In considering this population, we made the decision only to include students (both M.A. and Ph.D. students) enrolled in a graduate program that offered a doctoral degree; our justification was that we believed the preparatory experiences and subsequently, the attention paid to teaching responsibilities might be quite different for students in programs that only offered a Master's degree.

The initial sampling frame included statistics departments from 22 institutions selected from *US News and World Report's Best Grad Schools* (2013) ranked list of statistics programs. Statistics departments from five additional institutions were also included in the sampling frame because of their interest in statistics education and potential to participate in future planned GTA development opportunities.

A snowball sampling technique was employed in which an e-mail message was sent to faculty members in the departments included in the sampling frame. The e-mail message asked the faculty members to forward an invitation email to GTAs in their departments that included the URL for the survey. To encourage participation, the invitation e-mail also promoted the fact that any GTA who completed the survey would be given the opportunity to enter a drawing for a \$100 Amazon gift card.

GTAs were also recruited when a posting on a statistics blog made the survey link public. It is worth noting that although this publicity increased the overall sample size, it also gave survey access to GTAs from outside of the target population. Because of the latter, several steps were taken to filter the data and remove responses from GTAs who were not part of the population.

All respondents who completed the survey after the date the blog post appeared were identified using metadata from *Survey Gizmo*. In this manner, we identified 55 cases as potential candidates for removal. We were able to link several of these respondents to the secondary data that was collected for use in the Amazon gift-card drawing. Based on this, we were able to remove 13 GTAs who were enrolled in statistics programs at institutions that were not a part of the population. We also filtered out 24 respondents who did not indicate a target institution and whose IP address (metadata that *Survey Gizmo* collects by default) was located at least 60 miles from one of the target institutions.

Using these methods, we were able to remove 37 of the 55 ambiguous cases. This left 18 respondents for whom we were unable to conclude whether they were a part of the intended population. Subsequently, we performed two sets of analyses, both including, and omitting the ambiguous cases. Since the conclusions drawn from both sets of analyses were similar, we only report the results for the analysis that includes these 18 cases.

4. RESULTS

In this section, we present several analyses of the survey responses. We also include pertinent discussion along with these results. Broader findings, issues, and limitations are withheld and addressed in the “Discussion” section. The results and discussion are organized and presented based on the four major research questions:

1. What professional development experiences for teaching do statistics GTAs have?
2. What are GTAs’ beliefs about the teaching of introductory statistics?
3. To what extent are GTAs’ teaching beliefs aligned with their teaching practices?
4. To what extent are GTAs’ teaching practices aligned with professionally endorsed recommendations for teaching introductory statistics?

After filtering the data, described previously, the useable sample included responses from 213 GTAs enrolled at 38 Ph.D.-granting institutions representing all major regions across the United States. Because *Survey Gizmo* requires participants to respond to all forced-choice items, all 213 respondents had complete records. The response frequencies (and percentages) for all items are presented in Appendix A. All analyses were conducted using the R statistical program (v. 3.0.2) (R Core Team, 2013).

Before presenting the results, we acknowledge that the sample was not randomly drawn from the target population, and as such, the responses may not be generalizable. Rather than try to argue that the sample is representative or draw grandiose inferences, we instead interpret the results within this study’s particular context. As such, we hope that they might inform other researchers carrying out similar work and as such contribute to the overall understanding of the landscape of GTA development. These goals are consistent with the qualitative nature of the work (e.g., Guba, 1981; Shenton, 2004).

4.1. GTA’S PROFESSIONAL DEVELOPMENT RELATED TO TEACHING

To examine GTAs’ preparation to teach introductory statistics courses, we examined the degree to which GTAs reported they had (1) training in pedagogical topics for teaching statistics, and (2) participated in teaching-specific professional development.

Pedagogical training The survey responses suggested that many of the sampled GTAs are aware of common misconceptions and errors students make when learning statistics (62%). However, a much smaller percentage (44%) is familiar with the research focused on how students learn statistics. It is interesting, although perhaps not surprising, that GTAs show more awareness about the problems and errors students make when learning statistics, than they do about potential remedies. This may be because GTAs' own experiences with difficult probability concepts have heightened their awareness of potential misconceptions; or, GTAs' familiarity with research about common misconceptions may reflect traditions of literature that are predominant in the field. Historically, research on statistics education primarily focused on the study of heuristics and errors commonly made in probabilistic or statistical situations (e.g., Tversky & Kahneman, 1974). Perhaps it is time for GTAs, and GTA educators, to focus their attention on research related to how students learn statistics (e.g., Garfield & Ben-Zvi, 2007) rather than on how students fail to learn statistics.

The data also suggested that many GTAs have not yet learned about many reform-based teaching principles (e.g., design and use of active-learning based lessons, use of cooperative groups, and formative methods for assessment). Only between 40%–55% of surveyed GTAs indicated that they had learned about these topics. It is unclear why GTAs have not been exposed to these principles. Perhaps it is reflective of the norms or teaching conventions within a department; or, it could just be that it is common for instructors to believe that the use of reform-based pedagogies do not work under certain conditions (e.g., in large classes).

GTAs reported feeling quite prepared to teach with technology tools designed to enhance student learning. Although 70% indicated preparation with technology tools, it is worth pointing out that the survey items did not allow for more nuanced information than this. For example, it may be that GTAs have only been prepared to teach using tools such as calculators. Thus, although encouraging, inferences about GTAs' preparation to teach using technology tools must be treated with caution.

Participation in professional development The other main area of GTAs' preparation for teaching that was examined was the self-reported data about participation in professional development (PD) opportunities related to teaching statistics. The *GSSTI* asked GTAs about PD opportunities within their institutions (teaching observation(s) with feedback, teaching mentor, or courses) and at the broader level (workshops and conferences). Roughly 85% of the surveyed GTAs reported that they had participated in at least one of these types of PD.

Within institutions, the most common type of PD experienced for these GTAs seems to be teaching observations with feedback (55%). A smaller percentage of GTAs (40%) had received faculty mentoring, and, fewer still have taken a course (8.5%) or courses (21.6%) that lasted more than one semester. The research literature indicates that these three types of PD experiences have the largest impact on GTAs' teaching affect and behaviors (e.g., Gilmore et al., 2014; Williams, 1991; Wyse, 2010), yet many surveyed statistics GTAs indicated they had not experienced these opportunities.

Analyses were also conducted to find out how many of the five professional development experiences offered in the survey were typically experienced by GTAs. Although roughly half of the surveyed GTAs had experienced two or three of the professional development opportunities offered by the survey, 14% of surveyed GTAs reported that they had not experienced any of them. At face value, this may indicate a large need for more opportunities and encouragement for GTAs to participate in various types of professional development related to teaching. However, several considerations should be taken into account when examining this 14% figure.

On one hand, the 14% figure may be inflated. The survey may have attracted GTAs who have not had much preparatory experience and were motivated to participate in the survey as a way to express their discontent. Comments offered by participants in the open-ended question at the end

of the survey indicated that a number of participants were extremely unsatisfied with their preparatory experiences. This may suggest that the GTAs who participated were uniquely less prepared than typical statistics GTAs, and the 14% figure is too high. Also, it is likely that many GTAs have participated in other preparatory experiences that were not included in this survey. For example, many GTAs may have participated in weekly meetings, a type of professional development that was not included in the *GSSTI* and should be included in future studies.

On the other hand, the 14% figure may be too low. GTAs participating in the survey may naturally have more interest in topics related to teaching than statistics GTAs who did not choose to participate. The GTAs who participated in the survey may also be the type of GTAs who would participate in optional workshops and courses available to them. Under this consideration, the 14% might not be inflated because of the natural appeal of the survey to a more teaching-centered GTA population. Therefore, further study could be conducted to explore the extent to which GTAs are lacking in preparatory experiences.

4.2. GTAS' BELIEFS ABOUT TEACHING INTRODUCTORY STATISTICS

We also surveyed GTAs on their beliefs about teaching an introductory statistics course. The items pertaining to GTA beliefs were related to (1) the content/topics that should be included in an introductory statistics course, (2) pedagogical methods that should be used in introductory statistics courses, and (3) assessment methods that should be used in introductory statistics courses.

The surveyed GTAs appear to agree about most of the content that should be included in an introductory statistics course. More than 70% of surveyed GTAs are in favor of teaching rules and formulae for probability and theoretical probability distributions (e.g., the binomial distribution). Over 95% agreed that students should learn connections between the characteristics of the data and the inferences that are made. The content area with the least consensus was regarding the role of mathematical formulae (e.g., the formula for calculating the standard deviation or the standard error). About 37% of GTAs agreed that mathematical formulae should play a primary role, whereas just under seven percent were undecided.

GTAs had far less consensus in their responses to the items related to their pedagogical beliefs. For example, none of the three items related to delivery of course content (via lecture, use of activities, and small group work), reached 60% percent agreement. And, over 10 percent of surveyed GTAs responded "undecided" to each of the items. The only pedagogical method for which GTAs' responses were more unified was the use of technology. More than 75% were in favor of using technology, both to illustrate abstract statistical concepts and to analyze data in introductory statistics courses.

When it comes to assessment methods, the overwhelming majority of GTAs agreed that students should be assessed on their ability to complete open-ended problems. Most also believed assessments should be used to provide formative feedback to students. GTAs were divided, however, about whether quizzes and exams should be the primary method to assess student learning, with nearly 10% of respondents undecided on the matter.

4.3. ALIGNMENT BETWEEN GTAS' TEACHING BELIEFS AND PRACTICES

To examine whether statistics GTAs' beliefs and classroom practices were aligned, we computed the association between item responses in four areas: use of technology; the role of lecture; the role of activities; and use of group work. Because of the nominal-level measurements of the items, Cramér's V is reported as a measure of the alignment between GTAs beliefs and classroom practices. Similar to a correlation coefficient, Cramér's V can take values between 0 and 1 (inclusive), where higher values indicate more alignment.

The results of the *GSSTI* indicate very little association between GTAs' use of technology and their beliefs that technology should be used to illustrate abstract ideas (Cramér's $V = .06$). Of the surveyed GTAs who believe technology should be used, more than one out of four still do not use technology to illustrate most abstract statistical concepts in the courses they teach or assist.

Very little association was also found regarding beliefs and practices about using lecture to deliver course content (Cramér's $V = .08$). In particular, 90% of GTAs who believe lecture should not be primary still teach or assist courses that mostly use lecture to deliver course content.

In contrast to results about the use of technology and lecture, moderate associations were found between beliefs and practices regarding the use of activities (Cramér's $V = .372$). Surveyed GTAs who believe activities should be used, primarily, to deliver course content are about three times more likely to teach or assist courses that do. Still, of the surveyed GTAs who believe in activities, only about half teach or assist courses that do use activities, primarily, to deliver content.

Moderate associations were also found between surveyed GTAs' beliefs and practices regarding the use of cooperative group work (Cramér's $V = .377$). GTAs who believe in the use of group work are roughly 2.5 times more likely to teach or assist courses that frequently require students to work together in small groups. It is interesting to note that a quarter of surveyed GTAs who do not believe group work should be used still frequently require group work, and associations were only moderate.

We also explored whether the disconnect between GTAs' teaching beliefs and practices was related to GTAs' lack of autonomy or responsibility to make decisions. To do this, we examined the associations between responses to items related to GTAs' perceived authority to make decisions about teaching methods and alignment of beliefs and practices. Little association was found; Cramér's V was less than 0.125 for all four areas studied (undecided GTAs omitted). The results indicated that lack of authority might not be the best explanation for disagreement between beliefs and practices. We speculate that there are many reasons why GTA beliefs and practices are not aligned, such as perceived lack of sufficient class-time to implement pedagogical beliefs, limited financial resources to support technology tools, unawareness of available tools and resources, or lack of comfort using alternative pedagogical methods or technological tools in the classroom.

4.4. ALIGNMENT BETWEEN GTAS' TEACHING PRACTICES AND PROFESSIONALLY ENDORSED RECOMMENDATIONS

Seven items on the *GSSTI* were used to examine the potential alignment between GTAs' teaching practices and professionally endorsed recommendations for teaching statistics. Each of these items was written to correspond with a particular recommendation outlined in the *GAISE College Report* (ASA, 2005). Responses suggest that GTAs' teaching practices are not particularly aligned with the *GAISE* recommendations. GTAs' teaching practices were aligned with only two of the seven recommendations; using real data (66%), and using technology for developing concepts (69%).

Survey results suggest that GTAs' teaching practices are not in alignment with recommendations such as fostering active learning and emphasizing understanding rather than knowledge of procedures. Over 90% of the GTAs surveyed indicated the courses they teach (or assist with) primarily use lecture to present the course content. Similarly, a majority of GTAs surveyed (65%) indicated a focus in their classrooms was the teaching and subsequent student practice of implementing statistical procedures using formulas. Only about one-third of the surveyed GTAs reported using more student-centered methods of teaching (e.g., active- or discovery-based learning).

Finally, it is worth noting that only 35% of surveyed GTAs indicated having learned about professionally endorsed guidelines for teaching statistics. This result indicates great need and great opportunity. For example, with the endorsement of newly updated ASA *GAISE* guidelines, leaders in statistics and statistics education could collaborate with GTA educators to introduce and familiarize them with the *GAISE* recommendations for teaching introductory statistics.

5. DISCUSSION

The purpose of this study was to explore empirically statistics GTAs' professional development related to teaching, their beliefs about how introductory statistics should be taught, and the alignment between their teaching practices and beliefs, as well as with professional recommendations for teaching introductory statistics. An online survey, the *GSSTI*, was developed and administered to 213 statistics GTAs across the United States. Before launching into a broader discussion of what was learned, we point out several limitations of the survey research.

5.1. LIMITATIONS

One set of limitations to the study is related to sampling complexities. For example, the benefits of participation, namely, contribution to research on GTA professional development and entry in a random drawing for a \$100 Amazon gift card, are likely to have appealed to a select sample of GTAs, potentially biasing the results. In addition, the process used to eliminate participants who were not explicitly included in the initial sampling frame may not have completely filtered out those responses, again potentially biasing the results. We also note that the use of IP addresses as a proxy for respondent location may be problematic. First, we cannot account for respondents who may have been traveling. In such a case, we may have inadvertently eliminated participants who were part of the target population. Secondly, we acknowledge that geolocation from IP addresses can be quite inaccurate depending on network conditions (e.g. ISP, mobile use). This could have also led to bias in the results.

Another limitation of the study is that the conclusions and inferences drawn are based on a fixed set of response options we included in the survey. As an example, if a GTA had experienced professional development in the form of a weekly meeting, they may not have thought to include this in the first free response item of the *GSSTI* that solicited "other" types of professional development related to teaching. This would lead us to categorize incorrectly that person as having no experience related to professional development, despite the fact that Rumsey (1998) points out that weekly meetings are a valuable development experience.

Another possible limitation arises from ambiguity in the survey items. For example, participants were asked to indicate their beliefs about whether several topics should be included in an introductory statistics course. It is feasible that participants' responses to these items might vary depending on the student population they considered when responding to the survey (e.g., statistics majors, students in science related fields, or liberal arts students seeking to fulfill a quantitative requirement).

Related to this, GTAs may have not interpreted things in the same way that they were intended. For example, inconsistency in response patterns indicate that GTAs may be confusing *activities* and *demonstrations* (e.g., although 41% of GTAs surveyed disagreed that lectures should be the primary method of delivering course content, 44% agreed that content should be presented mostly through activities). This discrepancy may also be due to the group of GTAs who selected the *undecided* response option, or it is also possible that GTAs actually agree with both items.

Unfortunately, there is not a clear choice about which interpretation, and subsequent inferences, should be drawn in these cases. As such, we try to temper the conclusions we draw in the upcoming sections, especially those where the data are ambiguous. However, even though the study has limitations, we believe that much of the data collected is still useful in terms of thinking about how statistics departments and faculty might improve GTAs' preparation to teach introductory statistics courses. We turn there next.

5.2. IMPROVING GTAS' PREPARATION TO TEACH INTRODUCTORY STATISTICS

The results of the survey indicate that many GTAs may not have experience with the types of professional development associated with positive impact on teaching (e.g., having a faculty mentor). In statistics departments where GTAs are teaching courses, faculty could be asked to take on a mentoring role, at the very least, observing GTAs' teaching and providing them with feedback to help them improve. Faculty in these departments could also encourage GTAs to participate in professional development opportunities outside of their departments (e.g., see Barrington, 2001; Pruitt-Logan & Gaff, 2004).

That being said, we understand that in many cases, limited faculty and monetary resources hinder the provision of these opportunities. When departments and faculty are unable to implement these directly, more grass roots efforts may suffice. For example, more senior GTAs might be tapped to mentor their junior colleagues. This type of peer mentorship has been found to reduce teaching anxiety and improve teaching affect (Williams, 1991). Cohorts of GTAs could also observe each other, providing feedback, while at the same time learning from what they see in the classroom (e.g., see Speer, 2004).

Surveyed GTAs also seem to hold somewhat rigid beliefs about teaching introductory statistics. Many of these beliefs regarding the content that should be included in a course, or the assessment methods that should be used to evaluate students, are not aligned with professionally endorsed recommendations in the *GAISE College Report* (ASA, 2005). It is unclear why GTAs' beliefs are unaligned with these recommendations. It may be that without appropriate professional development, GTAs are not being exposed to recommended pedagogical or assessment methods. Or, it may be that even if they have been exposed to these methods, institutional constraints, such as large class sizes, temper their beliefs about these methods' effectiveness. We speculate that the apprenticeship of observation phenomenon may be at play (Lortie, 1975). That is, in their many years of prior study while observing their own teachers using traditional methods, GTAs may have developed rigid, deeply-held beliefs about teaching. These beliefs may prevent GTAs from appropriately considering or adopting alternate teaching strategies (e.g., Pajares, 1992).

Interestingly, even for the GTAs whose beliefs aligned with the GAISE recommendations, their classroom practices often did not. For example, GTAs predominantly reported that lecturing was the primary method of delivering content, and this was true even among those GTAs who believed that lectures should not be the primary way to deliver content. Similarly, GTAs' beliefs about how technology should be used in the classroom do not appear to influence their classroom practices. Similar discrepancies between beliefs and practices have been found among elementary and secondary pre-service teachers (e.g., Raymond, 1997; Skott, 2001), for which researchers have encouraged pre-service teacher development programs to include opportunities for reflection upon one's beliefs as they relate to teaching (e.g., Cooney, Shealy, & Arvold, 1998).

These responses suggest that there may also be a need to provide GTAs more opportunities to learn about pedagogical and assessment methods that have been identified as successful for teaching statistics. Where feasible, departments could offer a course or seminar related to teaching statistics—although the research literature suggests that these should last longer than one semester for more optimal improvement. At many universities, there may be institutional-level support available to facilitate these efforts, for example, Centers for Teaching and Learning, or

Preparing Future Faculty programs. We encourage faculty to learn about these opportunities and communicate them to GTAs.

Again, we recognize that not every university can (nor should) create a course on teaching statistics. In these institutions, having GTAs participate in a focused scholarly community might be beneficial. In such a cohort, GTAs could read about, discuss, and reflect on research related to how students learn statistics, methods for using technology to enhance delivery of course content, designing and facilitating cooperative group work as an instructional technique, different forms of assessment, and the differences between demonstrations and activity-based lessons. GTAs could also be asked to consider their own classroom practices, perhaps in a teaching journal, and point out mismatches between these practices and those that are recommended.

Lastly, we point toward conferences such as the *Joint Statistics Meetings* or the *Joint Mathematics Meetings*, as potential opportunities for preparing GTAs. These venues, which many statistics GTAs already attend, seem like natural settings for preparing GTAs through specific short courses or workshops. Using these meetings, institutions could pool resources to make professional development opportunities available more broadly.

5.3. IMPLICATIONS FOR FUTURE RESEARCH

Although the *GSSTI* may provide some useful information about GTAs' development as teachers of statistics, we hope that other researchers interested in further exploring the developmental experiences, beliefs, and teaching practices of statistics GTAs can learn from our experiences and mistakes. In retrospect, we would have collected data about other common types of professional development that GTAs experience, including weekly GTA meetings and peer mentoring (e.g., mentoring by senior GTAs). We also would have asked GTAs to specify the student population for their introductory statistics course. Lastly, we would have gathered information about the GTAs' previous teaching experience, their degree status (Master's or Ph.D.), and their international student status. These factors might have helped us better understand why GTAs hold their teaching beliefs and practices, and discern reasons for the disconnect between the two.

At the same time that we consider how GTAs might be prepared to teach introductory statistics, it is also important to study the efficacy, design, and implementation of various professional development opportunities. For example, research could be conducted to explore whether there are differences in GTAs' teaching under alternative mentoring strategies. Factors such as the frequency and duration of mentoring, the role of the person doing the mentoring—faculty or peer—and the amount and level of feedback provided could all be manipulated and examined.

There are also open questions about the strategies for engaging GTAs in learning about the pedagogical and assessment methods promoted in documents such as the *GAISE College Report* (ASA, 2005). Research here could focus on the identification and exploration of effective characteristics of courses, seminars, and meetings that are designed to facilitate learning about these topics.

In addition to studying GTAs' professional development in more formal settings (e.g., courses, workshops), it is also important to understand how GTAs develop in more informal settings. For instance, Wulff et al. (2004) suggest that GTAs informally engage in professional development as they socialize with one another. Information is needed about the extent to which these communities of practice naturally emerge in statistics departments, and the degree to which they help GTAs develop as teachers. Understanding the nature and characteristics of these communities could ultimately help faculty and departments cultivate these relationships to optimize GTAs' development.

Lastly, in order to carry out this type of research, there is a need for several high-quality instruments that could be used to measure different aspects of GTAs' development as teachers. Such instruments could be used in conjunction with existing instruments designed to measure student outcomes—for example, statistical literacy (Ziegler, 2014), statistical reasoning (Sabbag & Zieffler, 2015), or students' attitudes towards statistics (Schau, Stevens, Dauphinee, & Vecchio, 1995)—to identify whether GTA development actually results in these desired student outcomes.

Ultimately, there are many opportunities for conducting research on statistics GTAs, including their professional development, given their increasing presence in teaching or assisting with statistics courses. Research on statistics GTA professional development can play a vital role in shaping the future statistics professorate, and can equip future statistics faculty to teach the diverse set of students enrolled in statistics courses to navigate our data-driven society.

REFERENCES

- Abbott, R. D., Wulff, D. H., & Szego, C. K. (1989). Review of research on TA training. In J. D. Nyquist, R. D. Abbott, & D. H. Wulff (Eds.), *Teaching assistant training in the 1990s. New Directions for Teaching and Learning* (No. 39, pp. 111–123). San Francisco, CA: Jossey-Bass.
- American Statistical Association (2005). *Guidelines for assessment and instruction in statistics education college report*. Alexandria, VA: Author. Retrieved from <http://www.amstat.org/education/gaise/GAISECollege.htm>
- Barrington, E. (2001). Encouraging professional training of graduate teaching assistants. *The Journal of Graduate Teaching Assistant Development*, 8, 107–113.
- Belnap, J. K., & Allred, K. (2009). Mathematics teaching assistants: Their instructional involvement and preparation opportunities. *Studies in Graduate and Professional Student Development*, 12, 11–38.
- Bettinger, E., & Long, B. T. (2004). *Do college instructors matter? The effects of adjuncts and graduate assistants on students' interests and success* (No. 10370, p. 42). Cambridge, MA: National Bureau of Economic Research. Retrieved from <http://www.nber.org/papers/w10370>
- Birch, J. B., & Morgan, J. (2005). TA training at Virginia Tech: A stepwise progression. *American Statistician*, 59(1), 14–18.
- Blair, R., Kirkman, E., & Maxwell, J. (2013). *Statistical abstract of undergraduate programs in the mathematical sciences in the United States: Fall 2010 CBMS survey*. Washington, DC: American Mathematical Society.
- Boman, J. S. (2013). Graduate student teaching development: Evaluating the effectiveness of training in relation to graduate student characteristics. *Canadian Journal of Higher Education*, 43(1), 100–114. Retrieved from <http://files.eric.ed.gov/fulltext/EJ1007032.pdf>
- Bray, J. H., & Howard, G. S. (1980). Methodological considerations in the evaluation of a teacher–training program. *Journal of Educational Psychology*, 72(1), 62–70. doi:10.1037//0022-0663.72.1.62
- Carroll, J. G. (1980). Effects of training programs for university teaching assistants: A review of empirical research. *The Journal of Higher Education*, 51(2), 167. doi:10.2307/1981372
- Chism, N. V. N., & Warner, S. B. (Eds.). (1987). *Institutional responsibilities and responses in the employment and education of teaching assistants: Readings from a national conference*. Columbus, OH: The Ohio State University Center for Teaching Excellence. Retrieved from <http://files.eric.ed.gov/fulltext/ED292783.pdf> - page=83
- Constantinides, J. C. (1987). Designing a training program for international teaching assistants. In N. V. N. Chism & S. B. Warner (Eds.), *Institutional responsibilities and responses in the employment and education of teaching assistants: Readings from a national conference* (pp.

- 275–283). Columbus, Ohio: The Ohio State University Center for Teaching Excellence. Retrieved from <http://files.eric.ed.gov/fulltext/ED292783.pdf> - page=83
- Cooney, T. J., Shealy, B. E., & Arvold, B. (1998). Conceptualizing belief structures of preservice secondary mathematics teachers. *Journal for Research in Mathematics Education*, 29, 306–333.
- Crede, E., Borrego, M., & McNair, L. D. (2010). Application of community of practice theory to the preparation of Engineering graduate students for faculty careers. *Advances in Engineering Education, Summer* 2(2), 1–22. Retrieved from <http://advances.asee.org/wp-content/uploads/vol02/issue02/papers/aee-vol02-issue02-p04.pdf>
- Dalgaard, K. A. (1982). Some effects of training on teaching effectiveness of untrained university teaching assistants. *Research in Higher Education*, 17(1), 39–50.
- Darling, A. L. (1987). TA socialization: A communication perspective. In N. V. N. Chism & S. B. Warner (Eds.), *Preparing graduate students to teach: Past, present, and future* (pp. 91–94). Columbus, OH: The Ohio State University. Retrieved from <http://files.eric.ed.gov/fulltext/ED292783.pdf> - page=83
- Darling, A., & Dewey, M. (1990). Teaching assistant socialization: Communication with peer leaders about teaching and learning. *Teaching and Teacher Education*, 6(4), 315–326. doi:10.1016/0742-051X(90)90024-Y
- Darling, A., & Staton, A. (1989). Socialization of graduate teaching assistants: A case study in an American university. *International Journal of Qualitative Studies in Qualitative Education*, 2(3), 221–235.
- Davis, W. E., & Minnis, D. L. (1993). Designing a program to prepare graduate students for careers as college teachers. *Innovative Higher Education*, 17(3), 211–224. doi:10.1007/BF00915602
- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2009). *Internet, mail, and mixed-mode surveys: The tailored design method* (3rd ed.). Hoboken, NJ: John Wiley & Sons, Inc.
- Fagen, A. P., & Suedkamp Wells, K. M. (2004). The 2000 national doctoral program survey: An on-line study of students' voices. In D. H. Wulff & A. E. Austin (Eds.), *Paths to the professoriate: Strategies for enriching the preparation of future faculty* (pp. 74–91). San Francisco, CA: Jossey-Bass.
- Ferzli, M., Morant, T., Honeycutt, B., Warren, S. E., Fenn, M., & Burns, B. (2012). Conceptualizing graduate teaching assistant development through stages of concern. In G. Gorsuch (Ed.), *Working theories for teaching assistant development* (pp. 231–275). Stillwater, OK: New Forums Press.
- Frank, B. W., & Speer, N. M. (2013). Building knowledge for teaching: Three cases of physics graduate students. In Engelhardt, P. V., Churukian, A. D., & Rebello, N. S. (Eds.), *AIP Conference Proceedings* (Vol. 1513, pp. 126–129). Melville, NY: American Institute of Physics. Retrieved from http://umaine.edu/merg/files/2012/07/Frank_Speer_PERC_2012.pdf
- Froelich, A. G., Duckworth, W. M., & Stephenson, W. R. (2005). Training statistics teachers at Iowa State University. *The American Statistician*, 59(1), 8–10. Retrieved from <http://wrstephe.public.iastate.edu/TrainingTAs.pdf>
- Garfield, J., & Ben-Zvi, D. (2007). How students learn statistics revisited: A current review of research on teaching and learning statistics. *International Statistical Review*, 75(3), 372–396.
- Garfield, J., & Everson, M. (2009). Preparing teachers of statistics: A graduate course for future teachers. *Journal of Statistics Education*, 17(2), 223–237. Retrieved from <https://ww2.amstat.org/publications/jse/v17n2/garfield.pdf>
- Gelman, A. (2005). A course on teaching statistics at the university level. *The American Statistician*, 59(1), 4–7.

- Gilmore, J. A., Maher, M. A., Feldon, D. F., & Timmerman, B. E. (2014). Exploration of factors related to the development of science, technology, engineering and mathematics graduate teaching assistants' teaching orientations. *Studies in Higher Education, 39*(10), 1910–1928.
- Gilreath, J. A., & Slater, T. F. (1994). Training graduate teaching assistants to be better undergraduate physics educators. *Physics Education, 29*(4), 200–203. doi:10.1088/0031-9120/29/4/003
- Golde, C. M., & Dore, T. M. (2001). *At cross purposes: What the experiences of today's doctoral students reveal about doctoral education* (pp. 1–65). Philadelphia, PA: Pew Charitable Trusts. Retrieved from <http://files.eric.ed.gov/fulltext/ED450628.pdf>
- Green, J. L. (2010). Teaching highs and lows: Exploring university teaching assistants' experiences. *Statistics Education Research Journal, 9*(2), 108–122. Retrieved from [http://iase-web.org/documents/SERJ/SERJ9\(2\)_Green.pdf](http://iase-web.org/documents/SERJ/SERJ9(2)_Green.pdf)
- Guba, E. G. (1981). Criteria for assessing the trustworthiness of naturalistic inquiries. *Educational Technology Research and Development, 29*(2), 75–91.
- Harkness, W. L., & Rosenberger, J. L. (2005). Training graduate students at Penn State University in teaching statistics. *The American Statistician, 51*(1), 11–13.
- Hoessler, C., & Godden, L. (2015). The visioning of policy and the hope of implementation: Support for graduate students' teaching at a Canadian institution. *Canadian Journal of Higher Education, 45*(1), 83–101. Retrieved from <http://files.eric.ed.gov/fulltext/EJ1061061.pdf>
- Hogg, R. V. (1991). Statistical education: Improvements are badly needed. *The American Statistician, 45*(4), 342–343.
- Holmes, N.G., Martinuk, M., Ives, J., & Warren, M. (2013). Teaching assistant professional development by and for TAs. *The Physics Teacher 51*(4), 218–219. Retrieved from http://cwsei.ubc.ca/SEI_research/files/Physics/Holmes-et-al_TA-ProfDevelop_PT2013.pdf
- Jones, J. L. (1993). TA training: From the TA's point of view. *Innovative Higher Education, 18*(2), 147–161.
- Kung, D., Speer, N. M., & Gucler, B. (2006). Teaching as learning: Mathematics graduate students' development of knowledge of student thinking about limits. In S. Alatorre, J. L. Cortina, M. Sáiz, & A. Méndez (Eds.), *Proceedings of the 28th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 835–836). Mérida, México: Universidad Pedagógica Nacional.
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge, United Kingdom: Cambridge University Press.
- Legrand, M. (2001). On the training of French prospective university teachers. In D. Holton (Ed.), *The teaching and learning of mathematics at university level: An ICMI study* (pp. 519–528). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Lortie, D. (1975). *Schoolteacher: A Sociological Study*. Chicago, IL: University of Chicago Press.
- Marincovich, M., Prostko, J., & Stout, F. (Eds.). (1998). *The professional development of graduate teaching assistants*. Bolton, MA: Anker Publishing Company.
- Merton, R. K., Reader, G., & Kendall, P. L. (1957). The student-physician. *The American Journal of the Medical Sciences, 235*(5), 613.
- Moore, D. (2005). Preparing graduate students to teach statistics: Introduction. *The American Statistician, 59*(1), 1–3. Retrieved from <http://www.stat.purdue.edu/~dsmoore/articles/TeachPrep.pdf>
- Muzaka, V. (2009). The niche of graduate teaching assistants (GTAs): Perceptions and reflections. *Teaching in Higher Education, 14*(1), 1–12. doi:10.1080/13562510802602400
- Myers, S. A. (1994). The availability and helpfulness of graduate teaching assistant socialization activities. *Communication Research Reports, 11*(2), 221–228.

- Myers, S. A. (1998). GTAs as organizational newcomers: The association between supportive communication relationships and information seeking. *Western Journal of Communication*, 62(1), 54–73. Retrieved from <http://www.tandfonline.com/doi/full/10.1080/10570319809374597>
- Noll, J. A. (2011). Graduate teaching assistants' statistical content knowledge of sampling. *Statistics Education Research Journal*, 10(2), 48–74. Retrieved from [http://iase-web.org/documents/SERJ/SERJ10\(2\)_Noll.pdf](http://iase-web.org/documents/SERJ/SERJ10(2)_Noll.pdf)
- Nowlis, V., Clark, K. E., & Rock, M. (1968). *The graduate student as teacher*. Washington, DC: American Council on Education.
- Nyquist, J. D., Abbott, R. D., Wulff, D. H., & Sprague, J. (Eds.). (1991). *Preparing the professoriate of tomorrow to teach: Selected readings in TA training*. Dubuque, IA: Kendall/Hunt Publishing Company.
- Nyquist, J. D., & Sprague, J. (1998). Thinking developmentally about TAs. In M. Marincovich, J. Prostko, & F. Stout (Eds.), *The professional development of graduate teaching assistants* (pp. 61–68). Bolton, MA: Anker Publishing Company.
- Nyquist, J. D., & Wulff, D. H. (1987). The training of graduate teaching assistants at the University of Washington. In N. V. N. Chism & S. B. Warner (Eds.), *Institutional responsibilities and responses in the employment and education of teaching assistants: Readings from a national conference* (pp. 144–154). Columbus, OH: The Ohio State University Center for Teaching Excellence. Retrieved from <http://files.eric.ed.gov/fulltext/ED292783.pdf> - page=83
- Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, 62(3), 307–332. Retrieved from <http://media.dropr.com/pdf/9gnziXcoKpR4NEygcZferbvB3HonvxCa.pdf>
- Park, C., & Ramos, M. (2002). The donkey in the department? Insights into the graduate teaching assistant (GTA) experience in the UK. *Journal of Graduate Education*, 3, 47–53. Retrieved from [http://www.lancaster.ac.uk/people/gyaccp/park and ramos \(2002\).pdf](http://www.lancaster.ac.uk/people/gyaccp/park%20and%20ramos%20(2002).pdf)
- Parrett, J. L. (1987). A ten-year review of TA training programs: Trends, patterns, and common practices. In N. V. N. Chism & S. B. Warner (Eds.), *Institutional responsibilities and responses in the employment and education of teaching assistants: Readings from a national conference* (pp. 67–79). Columbus, OH: The Ohio State University Center for Teaching Excellence. Retrieved from <http://files.eric.ed.gov/fulltext/ED292783.pdf> - page=83
- Pruitt-Logan, A. S., & Gaff, J. G. (2004). Preparing future faculty: Changing the culture of doctoral education. In D. H. Wulff & A. E. Austin (Eds.), *Paths to the professoriate: Strategies for enriching the preparation of future faculty* (pp. 177–193). San Francisco, CA: Jossey-Bass.
- R Core Team (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org/>
- Raymond, A. M. (1997). Inconsistency between a beginning elementary school teacher's mathematics beliefs and teaching practice. *Journal for Research in Mathematics Education*, 28(5), 550–576.
- Rodrigues, R. A. B., & Bond-Robinson, J. (2006). Comparing faculty and student perspectives of graduate teaching assistants' teaching. *Journal of Chemical Education*, 83(2), 305–312.
- Roehrig, G. H., Luft, J. A., Kurdziel, J. P., & Turner, J. A. (2003). Graduate teaching assistants and inquiry-based instruction: Implications for graduate teaching assistant training. *Journal of Chemical Education*, 80(10), 1206–1210.
- Rumsey, D. J. (1998). A cooperative teaching approach to introductory statistics. *Journal of Statistics Education*, 6(1), 163–171. Retrieved from <https://www2.amstat.org/publications/jse/v6n1/rumsey.html>

- Sabbag, A., & Zieffler, A. (2015). Assessing learning outcomes: An analysis of the GOALS-2 instrument. *Statistics Education Research Journal* 14(2), 93–116. Retrieved from [http://iase-web.org/documents/SERJ/SERJ14\(2\)_Sabbag.pdf](http://iase-web.org/documents/SERJ/SERJ14(2)_Sabbag.pdf)
- Schau, C., Stevens, J., Dauphinee, T. L., & Vecchio, A. D. (1995). The development and validation of the Survey of Attitudes Toward Statistics. *Educational and Psychological Measurement*, 55(5), 868–875.
- Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information*, 22(2), 63–75.
- Skott, J. (2001). The emerging practices of a novice teacher: The roles of his school mathematics images. *Journal of Mathematics Teacher Education*, 4(1), 3-28.
- Speer, N. M. (2004). Meeting the changing needs of mathematics graduate student teaching assistants: The evolution of professional development activities for instruction using collaborative groups. *Journal of Faculty Development*, 20(1), 5–10. Retrieved from http://umaine.edu/merg/files/2012/07/Speer2004_ChangingNeedsJFacDev.pdf
- Sprague, J., & Nyquist, J. D. (1991). A developmental perspective on the TA role. In J. D. Nyquist, R. D. Abbott, D. H. Wulff, & J. Sprague (Eds.), *Preparing the professorate of tomorrow to teach* (pp. 295–312). Dubuque, IA: Kendall/Hunt Publishing Company.
- SurveyGizmo. (2013). *SurveyGizmo*, computer software. Boulder, Colorado, accessed March 2013. Retrieved from <http://www.surveygizmo.com/>
- Torvi, D. A. (1994). Engineering graduate teaching assistant instructional programs: Training tomorrow's faculty members. *Journal of Engineering Education*, 83(4), 376–382.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, 185(4157), 1124-1131.
- US News and World Report: Best Grad Schools*. (2013). Retrieved from <http://grad-schools.usnews.rankingsandreviews.com/best-graduate-schools/search.result/program+top-science-schools/top-statistics-schools+y>
- Volkman, M. J., & Zgagacz, M. (2004). Learning to teach physics through inquiry: The lived experience of a graduate teaching assistant. *Journal of Research in Science Teaching*, 41(6), 584–602.
- Watzke, J. L. (2007). Longitudinal research on beginning teacher development: Complexity as a challenge to concerns-based stage theory. *Teaching and Teacher Education*, 23(1), 106–122.
- Weimer, M., Svinicki, M. D., & Bauer, G. (1989). Designing programs to prepare TAs to teach. In J. D. Nyquist, R. D. Abbott, & D. H. Wulff (Eds.), *Teaching assistant training in the 1990s. New Directions for Teaching and Learning* (No. 39, pp. 57–70). San Francisco, CA: Jossey-Bass.
- Williams, L. S. (1991). The effects of a comprehensive teaching assistant training program on teaching anxiety and effectiveness. *Research in Higher Education*, 32(5), 585–598. doi:10.1007/BF00992630
- Williams, D. E., & Roach, D. K. (1992). Graduate teaching assistant perceptions of training programs. *Communication Research Reports*, 9(2), 183–192.
- Wulff, D. H., & Austin, A. E. (Eds.). (2004). *Paths to the professoriate: Strategies for enriching the preparation of future faculty*. San Francisco, CA: Jossey-Bass.
- Wulff, D. H., Austin, A. E., Nyquist, J. D., & Sprague, J. (2004). The development of graduate students as teaching scholars: A four-year longitudinal study. In D. H. Wulff & A. E. Austin (Eds.), *Paths to the professoriate: Strategies for enriching the preparation of future faculty* (pp. 46–73). San Francisco, CA: Jossey-Bass.
- Wulff, D. H., Nyquist, J. D., & Abbott, R. D. (1991). Developing a TA program that reflects the culture of the institution: TA training at the University of Washington. In J. D. Nyquist, R. D. Abbott, D. H. Wulff, & J. Sprague (Eds.), *Preparing the Professoriate of Tomorrow to*

- Teach. Selected Readings in TA Training* (pp. 113–122). Dubuque, IA: Kendall/Hunt Publishing Company.
- Wyse, S. (2010). *Breaking the mold: Preparing graduate teaching assistants to teach as they are taught to teach*. Unpublished doctoral dissertation. Michigan State University. East Lansing, Michigan.
- Zieffler, A., Park, J., Garfield, J., delMas, R., & Bjornsdottir, A. (2012). The Statistics Teaching Inventory: A survey on statistics teachers' classroom practices and beliefs. *Journal of Statistics Education*, 20(1), 1–29. Retrieved from <http://ww2.amstat.org/publications/jse/v20n1/zieffler.pdf>
- Ziegler, L. (2014). *Reconceptualizing statistical literacy: developing an assessment for the modern introductory statistics course*. Unpublished manuscript. University of Minnesota, Minneapolis. Retrieved from <http://iase-web.org/documents/dissertations/14.LauraZiegler.Dissertation.pdf>

NICOLA JUSTICE
Education Sciences Building
56 East River Road
Minneapolis, MN 55455

**APPENDIX A: THE GRADUATE STUDENT STATISTICS TEACHING INVENTORY
ITEMS (AND PERCENTAGES FROM EACH RESPONSE CATEGORY)**

Percentages of respondents in each response category are given in parentheses.

1. Informed Consent for Information Disclosure

By completing this survey you are consenting to allow the use of the responses you provide for research purposes. All identifying information will be removed from your record and only aggregate data will be presented in the sharing of results through publications or research presentations.

Your participation in this research is very much appreciated and is completely voluntary.

- You have my permission to use my responses in your research
 You do not have my permission to use my responses in your research

Section 1: Your Teaching Experience

2. Have you ever had any of these responsibilities with introductory statistics courses? Check all that apply.

- Grading papers (86.4%)
 Holding office hours for students (89.7%)
 TA for lab/ discussion section (83.1%)
 Instructor (30.5%)
 Course coordinator (lead teacher, in charge of multiple sections) (5.2%)

Please answer the following questions with regards to the course you currently teach or the course you currently assist with. If you are teaching more than one introductory statistics course, choose one to keep in mind as you answer these questions.

3. Please rate the extent to which you have the responsibility to make decisions about each of the following regarding this course.

| | Little or No Responsibility | Some Responsibility | Most of the Responsibility |
|---------------------------------|-----------------------------|---------------------|----------------------------|
| 1. Course content | (67.1%) | (26.8%) | (6.1%) |
| 2. Teaching methods | (38.5%) | (39.4%) | (22.1%) |
| 3. Creating student assessments | (51.6%) | (34.3%) | (14.1%) |

Section 2: Your Pedagogy Practices

4. Consider a student who is fully engaged in your introductory statistics course.

Indicate the extent to which you think that student would agree or disagree with each of the following statements.

| | Strongly Disagree | Disagree | Agree | Strongly Agree |
|---|-------------------|----------|---------|----------------|
| The content was presented mostly through the instructor or TA's lectures. | (3.3%) | (4.7%) | (41.8%) | (50.2%) |
| The course frequently required students to work together. | (12.2%) | (40.8%) | (36.2%) | (10.8%) |
| The content was presented mostly through activities. | (14.1%) | (52.6%) | (28.2%) | (5.2%) |
| This course encouraged students to discover ideas on their own. | (14.1%) | (53.1%) | (28.6%) | (4.2%) |
| This course required students to do a lot of practice of procedures using formulas. | (5.6%) | (18.3%) | (54.0%) | (22.1%) |
| This course often used technology (e.g., web applets, statistical software) to help students understand concepts. | (9.4%) | (21.6%) | (38.0%) | (31.0%) |
| A large proportion of the data sets were hypothetical data sets that the instructor made up. | (17.4%) | (48.4%) | (29.1%) | (5.2%) |

Section 3: Your Beliefs about Teaching Statistics

5. Please rate the extent to which you agree or disagree with each of the following statements as they reflect your beliefs (but not necessarily your actual teaching) of an introductory statistics course.

| | UNDECIDED | Strongly Disagree | Disagree | Agree | Strongly Agree |
|--|-----------|-------------------|----------|---------|----------------|
| Rules of probability should be included. | (3.8%) | (1.4%) | (6.1%) | (50.7%) | (38.0%) |
| The topic of theoretical probability distributions (e.g., the binomial distribution) should be included. | (6.1%) | (1.9%) | (18.8%) | (44.6%) | (28.6%) |
| Technology should be used to illustrate most abstract statistical concepts. | (4.2%) | (2.3%) | (11.3%) | (40.4%) | (41.8%) |

| | UNDECIDED | Strongly Disagree | Disagree | Agree | Strongly Agree |
|--|-----------|-------------------|----------|---------|----------------|
| Students should learn connections between the characteristics of the data and the inferences that are made. | (0.9%) | (0.5%) | (0.5%) | (34.3%) | (63.8%) |
| Lectures should be the primary way for students to learn statistical content. | (12.2%) | (7.0%) | (33.8%) | (39.9%) | (7.0%) |
| Mathematical formulas (e.g., the formula for calculating the standard deviation or the standard error) should play a primary role. | (6.6%) | (17.4%) | (38.5%) | (30.9%) | (6.6%) |
| Quizzes and exams should be used as the primary way to evaluate student learning. | (8.9%) | (5.6%) | (33.3%) | (40.4%) | (11.7%) |
| Assessments should be used to provide informative feedback to students to improve their learning. | (2.3%) | (0.9%) | (2.3%) | (57.7%) | (36.6%) |
| Students should be assessed on their ability to complete an open-ended statistical problem. | (6.1%) | (3.3%) | (9.9%) | (45.5%) | (35.2%) |
| Students should analyze data primarily using technology. | (3.8%) | (3.8%) | (14.1%) | (48.4%) | (30.0%) |
| Students should frequently be required to work together in small groups. | (12.2%) | (5.6%) | (23.9%) | (39.9%) | (18.3%) |
| The course content should be presented mostly through activities. | (16.9%) | (3.3%) | (35.7%) | (35.2%) | (8.9%) |

Section 4: Your Development as a Teacher

6. Please indicate whether or not you have had any of the following experiences with different types of professional development activities related to teaching.

| | Yes | No |
|---------------------|---------|---------|
| Conference | (24.9%) | (75.1%) |
| 1-2 day workshop | (39.4%) | (60.6%) |
| 3 – 10 day workshop | (9.4%) | (90.6%) |

| | Yes | No |
|--------------------------------|---------|---------|
| Summer or Semester-long course | (43.7%) | (56.3%) |
| Year-long course | (8.5%) | (91.5%) |
| Multiple courses | (21.6%) | (78.4%) |
| Faculty Mentor | (39.9%) | (60.1%) |
| Observation with Feedback | (52.1%) | (47.9%) |
| Other: (please specify below) | (6.1%) | (93.4%) |

7. Other (from above question, if applicable)

| |
|--|
| |
|--|

8. Please rate the extent to which you have learned about each of these topics.

| | Not at all | To some extent | To a major extent |
|---|------------|----------------|-------------------|
| Research on how students learn statistics | (56.3%) | (35.7%) | (8.0%) |
| Technology tools designed to enhance student learning | (28.2%) | (58.2%) | (13.6%) |
| Research on common misconceptions and errors students make | (38.0%) | (47.9%) | (14.1%) |
| Methods for helping students learn together in cooperative groups | (52.6%) | (38.5%) | (8.9%) |
| Professionally endorsed guidelines for teaching introductory statistics | (65.7%) | (28.2%) | (6.1%) |
| Appropriate methods for assessment of student learning | (44.6%) | (43.2%) | (12.2%) |
| Non-traditional course content (e.g., big data, randomization methods) | (52.6%) | (36.2%) | (11.3%) |
| Principles of designing activity-based lessons | (59.2%) | (35.2%) | (5.6%) |

9. If you have any other comments you would like to share about your experiences as a teacher of statistics, or any other comments about this survey, please share them in the space below.

| |
|--|
| |
|--|

APPENDIX B: INSTITUTIONS REPRESENTED

Institutions indicated by GSSTI participants that were retained in the final data set are offered here in alphabetical order.

1. Arizona State University
2. Boston University
3. Carnegie Mellon University
4. Columbia University
5. Cornell University
6. Florida State University
7. George Mason University
8. Harvard University
9. Iowa State University
10. Johns Hopkins School of Public Health
11. Michigan State University
12. North Carolina State University
13. New York University
14. Ohio State University
15. Purdue University
16. Stanford University
17. The University of Chicago
18. University of California, Berkeley
19. University of California, Davis
20. University of California, Irvine
21. University of California, Los Angeles
22. University of North Carolina, Chapel Hill
23. University of Arizona
24. University of Florida
25. University of Georgia
26. University of Illinois, Urbana-Champaign
27. University of Maryland, College Park
28. University of Michigan
29. University of Minnesota
30. University of New Hampshire
31. University of Pennsylvania
32. University of Pittsburgh
33. University of Rochester
34. University of Virginia
35. University of Washington
36. University of South Carolina
37. University of Wisconsin, Madison
38. Yale University