The National Academies of
SCIENCES • ENGINEERING • MEDICINE

On What Basis? Seeking Effective Practices in Graduate STEM Education

Prepared by Margaret E. Blume-Kohout, Ph.D.
Visiting Assistant Professor, Colgate University

Final Version Presented on November 6, 2017

Commissioned by the
Committee on Revitalizing Graduate STEM Education for the 21st Century
Board on Higher Education and Workforce
Policy and Global Affairs Division

Final report available at
https://www.nap.edu/catalog/25038/graduate-stem-education-for-the-21st-century

Context for *On What Basis? Seeking Effective Practices in Graduate STEM Education*

The Committee on Revitalizing Graduate STEM Education for the 21st Century requested the following paper as one part of the execution of the landscape analysis described in the Statement of Task (below). The goal of the paper was to conduct a literature review, synthesis, and analysis on the academic research on how graduate students learn and which conditions can improve retention, persistence, career outcomes, and other indicators of student success. Foci may include the differential effects of graduate education programs and models, funding mechanisms, policies and practices, pedagogy, curriculum, the use of modules or specialized tracks, and other key components.

The author was asked to cover masters and PhD level education, with a focus on the STEM disciplines, which include the social, behavioral, and education sciences. The intended scope of the research paper began with the inclusion of the following components:

- Peer-reviewed articles that focus specifically on graduate STEM education and/or articles with a broad focus on graduate education, including STEM fields.
- Research on discipline specific programs (i.e. effective models in chemistry education.)
- Research on factors/experiences common across STEM fields.
- Research on the efficacy of programs designed to improve the retention and persistence of students, including those from populations historically underrepresented in STEM fields[1].
- Research on how the graduate student experience compares and contrasts for students of different backgrounds, including, but not limited to women, students from historically underrepresented minorities (URMs), women URMs, and international students[2].
- The impact of different graduate education models including, but not limited to, use of cohorts or lab rotations, internships, additional certificate programs (such as science communication or policy) or program pathways (academic track vs. industry).
- The impact of funding mechanisms (fellowship, traineeship, research assistantship, teaching assistantship).
- The impact of institutional support structures (multiple mentors, mandatory individual development plans).
- Program evaluations of major programs, including federally-funded by the National Science Foundation and the National Institutes for Health, or those funded by private institutions.

The primary focus of the paper was limited to books, papers, articles, and studies that include rigorous research methodology. In the event there is not enough rigorous research to constitute an analysis, the author was allowed to reference the following types of information in order to establish an understanding of the current state of graduate education:

- Reviews of initiatives or case –studies that contain primarily descriptive information.
- Best practices or practical field guides related to the implementation of graduate education programs.
- Rubrics, frameworks, or other competency guides that document the knowledge, skills, and competencies graduate students may receive over the course of their studies.
- Research that establishes the drivers of change in graduate education including federal policies, funding, and other macro-level elements.
- Frameworks on educational theory to help frame and contextualize the research and results.

[1] While this research area was requested in the initial statement of work, given the constraints on time and resources on the author, the committee elected to undertake additional research on diversity, equity, and inclusion beyond what was covered in this commissioned paper.
[2] See footnote 1.

The following, which will inform the Committee in other ways, falls outside the scope of this paper:

- Reports on graduate STEM education that do not include original research findings backed by quantitative and/or qualitative data.
- Editorials, position papers, and other writing on the general state of graduate STEM education.
- Research on graduate education that only focuses on the arts and humanities.
- Research on STEM education that focuses only on undergraduate or K-12 learners.

The author was asked to identify important gaps in the body of research that may suggest areas for future research as time permitted. After the author conducted an initial scan of the literature, the committee identified certain areas which required additional research, such as the literature on diversity, equity, and inclusion. Due to the restraints on the author's time and resources, the committee conducted those reviews external to this paper. This paper was one of many components that the committee referenced its evidence-gathering activities.

Statement of Task

An ad hoc committee, under the auspices of BHEW (Board on Higher Education and Workforce) and COSEPUP (Committee on Science, Engineering, and Public Policy), and liaising with GUIRR (Government-University-Industry Research Roundtable) and TAC (Teacher Advisory Council), will lead a study of STEM graduate-level education in the U.S., revisiting and updating a similar COSEPUP study completed 20 years ago.

Specific tasks will include:

- Conduct a systems analysis of graduate education, with the aim of identifying policies, programs and practices that could better meet the diverse education and career needs of graduate students in coming years (at both the master's and Ph.D. levels—understanding the commonalities and distinctions between the two levels), and also aimed at identifying deficiencies and gaps in the system that could improve graduate education programs.
- Identify strategies to improve the alignment of graduate education courses, curricula, labs and fellowship/traineeship experiences for students with the needs of prospective employers--and the reality of the workforce landscape--which include not only colleges and universities but also industry, government at all levels, non-profit organizations, and others. A key task will be to learn from employers how graduate education can continue to evolve to anticipate future workforce needs.
- Identify possible changes to federal and state programs and funding priorities and structures that would better reflect the research and training needs of graduate students.
- Identify policies and effective practices that provide students and faculty with information about career paths for graduates holding master's and Ph.D. degrees and provide ongoing and high quality counseling and mentoring for graduate students.
- Identify the implications of the increasingly international nature of graduate education and career pathways, reflecting both the numbers of foreign students who enroll in U.S. graduate schools and the increasing global migration of U.S. STEM graduates.
- Investigate the many new initiatives and models that are influencing graduate education, including MOOCs, other digital learning programs, increasing numbers of alternative providers of master's and Ph.D. degrees, and opportunities to secure credentials through multiple sources.
- Create a set of national goals for graduate STEM education that can be used by research universities, Congress, federal agencies, state governments and the private sector to guide graduate level programs, policies and investments over the next decade, and ensure that this "blueprint" for graduate education reform is revisited and updated on a periodic basis to reflect changing realities.

The products of this study will be an interim report and a final report that is widely disseminated for analysis and adoption of new programs, policies, and practices that enhance STEM graduate education. This may include dissemination activities on campuses, at professional society meetings and in other venues to share the report's findings and recommendations and to engage stakeholders in discussions around implementing new strategies, programs and models.

On What Basis?
Seeking Effective Practices in Graduate STEM Education

Margaret E. Blume-Kohout, PhD
October 25, 2017

The goal of this paper was to conduct a literature review, synthesis, and analysis of academic research on how graduate students learn, and which conditions can help to improve persistence and completion, and achieve desired career outcomes. Some 1,336 original research and review articles, program evaluations, reports, book chapters, and working papers were initially identified through Web of Science search as potentially relevant to this aim, of which—following review of abstracts—approximately 300 merited closer reading. In addition to these, over 100 additional journal articles and conference papers citing focal references from the first set or relating to topics of specific interest to the Committee (e.g., individual development plans) were identified via Google Scholar search, collected, and reviewed. Unfortunately, as Brint and Clotfelter (2016) also recently concluded, case studies of single-campus interventions generally provide inadequate evidence for their effectiveness, as even a seemingly rigorous mixed-methods evaluation can suffer from selection bias, maturation bias, or the Hawthorne effect. Few studies involve implementations across multiple campuses, with those primarily reflecting Federal or Foundation-sponsored programs. Below, I summarize key findings, noting the relative strength of the evidence base for each.

Key Findings and Recommendations:

- **Regular, directive interactions between faculty mentors and graduate students are instrumental to students' research skill development.** Research assistants on faculty-led projects tend to have more frequent interaction with their faculty members, all else equal. Institutions and funding agencies should consider how to ensure similar research integration for students supported by other funding mechanisms. Evidence is strong, based on quantitative, qualitative, and mixed-methods studies.

- **For PhD students seeking academic faculty positions, extensive pedagogical training and independent full-course instruction may be better postponed to postdoctoral fellowship programs.** Graduate students should have some exposure to academic teaching as part of a larger instructional team—for example as teaching assistants, discussion section leaders, and so on—and should receive appropriate and timely training in active learning methods to support that experience, in part to evaluate their own talent and interest in teaching. However, postdoctoral fellowship programs that combine mentored research with pedagogical training, mentored course development and teaching, and iterative feedback over 2-3 year term, yield better outcomes than nearly any predoctoral intervention studied in terms of actual behavioral change. This approach would avoid potential negative effects of TA positions on PhD time-to-degree, and would better target resources to PhDs with genuine interest in the combined research and teaching mission of academia. Strong theoretical basis and empirical evidence exists for pedagogical training extending beyond one semester, to include reflective exercises, iterative expert feedback, and structured, progressive instructional responsibility.

- **Individual development plans (IDPs) can be a useful tool, but effectiveness requires institutional and faculty buy-in,** as well as establishing psychologically safe mentor-protégé relationships such that students can be honest about their career goals and preferences, and mentors provide informed and constructive feedback on students' progress. The IDP process has strong theoretical support and evokes substantial empirical literature on effectiveness of goal-setting, but very little evidence is found for use or effectiveness of IDPs in the graduate education context.

- **Evidence indicates Responsible Conduct of Research ethics training is mostly ineffective,** across multiple delivery modes. Extant studies generally find, at best, mixed results, and in some cases reflect negative impacts of training on ethical decision-making and related metacognitive processes.

- **Early lab rotations appear to yield benefits for mentoring 'fit', student persistence, and subsequent research productivity.** None of the studies reviewed focused on evaluating the practice of lab rotations per se; rather, this practice came up in several other contexts, primarily in papers discussing student satisfaction or predicting PhD students' scholarly productivity. Thus, at present, evidence remains moderate: descriptive and correlational, but with both quantitative and qualitative support.

- **Democratizing access to undergraduate research experiences (UREs), then prioritizing candidates with UREs in admissions may improve persistence, retention, and diversity**, both by ensuring admitted students have some idea what research entails, and by leveling the field for admitted students with respect to faculty mentoring. While evidence is strong for benefits of UREs, this specific conclusion represents inductive reasoning. Multi-site, controlled implementation and evaluation are needed.

- **Student-centered active learning instructional strategies are effective in graduate classrooms, and team-based learning may improve student attitudes toward teamwork.** Evidence is strong that properly implemented active learning strategies increase students' retention and understanding and improve performance on applied problem-solving, and reduce classroom achievement gaps. At the graduate level, most evidence comes from graduate medical education. More evidence is needed to evaluate effectiveness of specific instructional strategies in other disciplinary contexts.

- **Faculty advisors are sometimes poor mentors. We don't know yet if mentor training will help.** Evidence for sustained behavioral change from faculty mentor training programs is weak, and those who select or apply to participate in these programs may not be those who need it the most.

Meaningful reforms require institutional commitment. Some examples of institutional support may include release time for course directors to design, implement, and run new courses; mentoring for postdoctoral teaching fellows with interest in higher education; recognition and rewards demonstrating high cultural valuation for faculty members' contributions to teaching and mentoring, including in the tenure review process (Dasgupta, Symes, and Hyman 2015, Gutlerner and Van Vactor 2013, Henderson, Beach, and Finkelstein 2011). Re-evaluating priorities in academic hiring as well as in tenure review to favor faculty hires with more diverse pre- and postdoctoral experience, with aptitude and interest in pedagogical innovation, and with a commitment to continuing professional development in leadership, communication, and mentoring skills—despite the research time lost—may help programs better to serve future graduate students with increasingly diverse identities, prior preparation, and professional goals.

Admissions

A recent National Academies report concluded that most studies—including both descriptive correlational studies and those designed to support causal inference—provide evidence of positive effects of undergraduate research experiences (UREs) on persistence in STEM fields (National Academies of Sciences 2017). Some evidence also suggests that UREs confer equal benefit in terms of research skill development when made mandatory, as when students self-select (Gilmore et al. 2015). Given this, graduate programs that openly encourage prospective applicants to pursue UREs, then prioritize applicants with such experiences may see greater persistence and retention, as those who apply would likely have greater self-knowledge of their own 'taste' for scientific research.

One clear concern with such an approach, however, is that high-quality URE opportunities are not distributed uniformly or proportionally across undergraduate institutions, and to the extent that there exist concentrations of such opportunities at, e.g., highly selective liberal arts colleges which tend to have lower shares of historically underrepresented minorities, first-generation, or economically disadvantaged students, such an approach might exacerbate existing bias in admissions. Previous research in the University of California system on student-faculty research interactions among undergraduates found that African American and first-generation college students were significantly less likely to have assisted faculty on research (Kim and Sax 2009). Because research skills are subject to cumulative advantage, students who arrive in PhD programs having had no prior research experience (either undergraduate or postbaccalaureate) may see their research skill gap widen further over their first year of study, relative to their peers (Feldon et al. 2016). In essence, small skill disadvantages become larger as faculty preferentially provide research opportunities and other resources to those more-skilled.

The Council of Graduate Schools' PhD Completion Project suggests a resolution to these issues: prospective graduate programs could provide promising undergraduates with early research experiences, for example between students' sophomore and junior years, as a recruitment tool, allowing both the student and the program to assess the potential match before students apply (Council of Graduate Schools 2008). Yet, from an implementation perspective, UREs often require more intensive faculty mentoring, thus have limited scalability at any given site, and relatively little established evidence exists for how to best mentor students (graduate or undergraduate) in learning research skills (Linn et al. 2015, Fuchs et al. 2016).

One promising approach is demonstrated by the NIH Postbaccalaureate Research Education Program (PREP) housed by research-intensive institutions, which offer a 1-year program primarily focused on research training—for example, critical analysis of scientific literature and laboratory skills—but also include professional development and enculturation activities (Hall et al. 2016). Nationally, about two-thirds (65%) of PREP scholars entered PhD programs, and a similar share were retained to complete their degrees, comparable to national averages overall (Hall, Mann, and Bender 2015). Few PREP scholars previously had UREs (Hall, Mann, and Bender 2015). The UNC-Chapel Hill PREP program reported 95% retention rate in PhD programs for participants from the first five years of its program (Hall et al. 2016). One key recruiting mechanism for the PREP program was from the department's own graduate admissions process, contacting their own PREP-eligible but declined applicants, many of whom were considering other career paths after initially being declined from the this department's and others' graduate programs (Hall et al. 2016). Beyond PhD program retention, however, evidence is mixed on the effects of URE interventions on graduate students' subsequent research productivity (Hall, O'Connell, and Cook 2017).

Aside from UREs, other frequently-used measures of applicants' aptitude and prior professional socialization are not significant predictors of graduate student retention (Lovitts 2001, 2004). Specifically, Lovitts' surveys of completers and non-completers found no significant difference between completers and non-completers in their undergraduate preparation and socialization characteristics with regard to GPA, GRE scores, mentoring, previous teamwork experience, involvement with department committees or campus organizations, subscribing to professional journals, attending professional meetings, or presenting papers outside the classroom.

More recent studies concur that the GRE is ‘particularly ineffective’ as a predictor of graduate student productivity, time to degree, or completion, and due to scores’ correlation with gender and race/ethnicity, reliance on these scores disadvantages individuals from underrepresented groups (Hall, O'Connell, and Cook 2017, Moneta-Koehler et al. 2017). Qualitative research likewise indicates successful transitions from coursework stage to independent research have relatively little to do with the analytical intelligence that GRE scores purport to test, though quantitative GRE scores in some fields may be predictive of the finished dissertation’s quality (Lovitts 2008). A meta-analysis of GRE scores’ predictive power across master’s and doctoral programs did find GRE scores predictive of grades in graduate coursework (Kuncel et al. 2010). GRE Verbal scores do seem correlated with faculty evaluations of students’ performance and ability to keep up with coursework at both the master’s and doctoral levels (Kuncel et al. 2010, Moneta-Koehler et al. 2017).

For PhD program retention and completion, Lovitts’ null finding with respect to prior professional socialization opportunities may partly be attributable to the typically greater opportunities for these activities at small liberal arts colleges. Follow-up interviews with non-completers indicated that many who had experienced close, collegial relationships with faculty as undergraduates—an experience apparently more common at liberal arts colleges than at large research-intensive universities—went to graduate school expecting more of the same, but were ultimately disappointed (Lovitts 2004).

Programs that choose to admit a larger initial cohort or wider range of applicants then apply ‘weed-out’ or ‘sink-or-swim’ approaches may inadvertently lose the ‘best and brightest’ students they hoped to retain, as many students—including high performers—respond poorly to the program’s apparent lack of support (Lovitts 2004).

Finally, men—most especially high-ability men, as measured by undergraduate GPA—are more likely than women to enroll in PhD programs in a recession, but for women this is not the case: women’s only countercyclical enrollment in graduate programs is for professional degrees (Bedard and Herman 2008). If not accounted for, gendered differences in responses to economic conditions may confound programs’ efforts to evaluate effects of changes in recruitment and admissions processes.

Lab Rotations

Reading and coursework may decrease students’ motivation, whereas hands-on research experience helps graduate students in science disciplines self-identify as researchers (Mantai 2015). In Lovitts’ (2004) study, PhD non-completers were six times more likely than completers to have been assigned to a faculty advisor, rather than finding a match by other means, such as through their own selection at time of application, through graduate coursework, or via word-of-mouth reputation. Another study found the strongest predictor of PhD student satisfaction with an advisor was satisfaction with the process by which their advisor was selected, but many students are unaware they can change advisors without penalty (Cockrell and Shelley 2011). Together, these and other studies focused on mentoring (see section below) indicate early lab rotations may offer students an opportunity to find better mentoring matches, increasing retention and subsequent research productivity (Weidman, Twale, and Stein 2001, McGaskey 2015). Lab rotations also provide graduate students an earlier opportunity to learn more about the research-intensive academic PI career path, which may allow them sooner to identify any desire to explore alternative careers (Fuhrmann et al. 2011). Students who are more involved with faculty on research activities tend also, unsurprisingly, to have higher research output (McGaskey 2015).

Along with first-year lab rotations, integrated first-year PhD curricula can provide students with greater flexibility to make better-informed decisions about their thesis research or degree-path specializations. For example, at one university, admission to its 15 doctoral degree-granting programs in biomedical sciences was consolidated into a single process, with all students taking the same first-year core courses and given opportunity for 3 or more lab rotations (Dasgupta, Symes, and Hyman 2015). Reported advantages of the program included decreasing redundancy in course content, thereby decreasing faculty time spent lecturing and permitting more time for active learning activities, while also promoting 'collegiality' among the incoming cohort.

Funding Mechanisms

Prospective graduate students in STEM fields are strongly motivated to enroll when greater financial support is offered, especially if offered a fellowship to attend (Bersola et al. 2014, Blume-Kohout and Clack 2013). Receiving scholarship funding may be perceived as providing validation and legitimacy for the PhD student's choice (Mantai 2015). For federal agencies, fellowships and traineeships may seem attractive as a means to control more directly the number of graduate students in training, and thus ultimately the STEM workforce labor supply (Stephan 2012, Blume-Kohout and Clack 2013). High-level guidance from multiple expert committees has encouraged moving towards fellowships (and away from research assistantships) to fund graduate students, but little evidence exists to support such recommendations from the perspective of training outcomes and scientific workforce retention.

The Andrew W. Mellon Foundation's Graduate Education Initiative concluded that providing funding in the form of fellowships directly to students or to universities did not substantially reduce attrition rates, increase completion rates, or reduce time-to-degree in the humanities and related social sciences (Groen et al. 2008, Ehrenberg et al. 2007). Groen et al. also conclude that, to the extent there were positive impacts, these were mediated by increases in financial support and increases in student quality. However, responses did vary by gender and departmental proportions of female students, with the largest improvements seen in departments with relatively higher proportions of female students.

Assistantships can significantly influence graduate students' attitudes and preferences with respect to their career options (Weidman, Twale, and Stein 2001). Teaching assistants may also have more contact with their peers, while research assistants often have more contact with faculty but less contact with peers (Weidman, Twale, and Stein 2001, Austin 2002). However, while both research assistantships and teaching assistantships require graduate students to come to campus and interact with faculty and fellow students, students on full fellowships may not have professional responsibilities that require them to come in to campus and interact, so may have fewer structured and chance interactions, leaving them less socially and academically integrated (Lovitts 2004). On the other hand, students who share lab and office space become more integrated with their research groups (Crede and Borrego 2012). Consistent with the above, evidence is mixed on the impact of fellowships (either institutional or external) versus assistantships on degree completion, with some studies indicating negative impacts (Lovitts 2001, 2004). Moreover, at least among biomedical sciences PhDs, those primarily supported on fellowships are less likely than those primarily supported on research assistantships to remain in scientific research-focused positions when they graduate (Blume-Kohout and Adhikari 2016). Similarly, although STEM PhD students and postdocs both may appreciate portable fellowships that increase their intellectual freedom and autonomy, their satisfaction declines if this is associated with lower levels of interaction with an advisor (Miller and Feldman 2015).

A recent evaluation of the National Science Foundation's Graduate Research Fellowship program (GRFP) also found GRFP awardees reported less practice applying for grants or contracts during graduate school, and fewer opportunities to engage in research activities or receive training or instruction on research, as compared to GRFP applicants who received Honorable Mentions (Bartolone et al. 2014). Because women are both more likely to receive fellowship support and less likely to serve as research assistants, some studies suggest women's underrepresentation in, respectively, STEM PhD programs and STEM careers, could be partly attributable to the fellowship funding mechanism. On the other hand, McGaskey (2015) found that—*controlling for the extent to which students conducted research with faculty*—Black PhD students with fellowship funding were more likely than those with other funding sources to have published. McGaskey (2015) concludes that funding agencies and institutions should "consider requiring all assistantships and fellowships to develop the recipients' research skills and encourage their participation in research-related activities such as presentations, attendance of professional meetings, and submission of papers for publication."

Historically, students primarily supported as teaching assistants (TAs) have tended to take longer to complete their degrees, as compared to those supported on fellowships or as research assistants (cf. Ehrenberg and Mavros (1995), Kim and Otts (2010)). For foreign students who are non-native English speakers, however, a TA position may be worth the delay. Unlike native U.S. citizen students, foreign PhD students whose primary financial support came from TA positions remain quite likely to take research-oriented positions in the U.S. scientific workforce after graduation, perhaps due to greater language skills development TA positions afford, or signaling of communication skills through their TA activities (Blume-Kohout and Adhikari 2016). However, more recent studies indicate that some structured teaching development experiences, such as those offered through the National Science Foundation's discontinued GK-12 program, may not significantly lengthen time-to-degree, and furthermore, graduate students who serve as *both* research assistants *and* either (a) teaching assistants in undergraduate classrooms, or (b) GK-12 fellows, demonstrate greater development in their ability to generate testable hypothesis and valid research designs, versus those who serve only as research assistants (Feldon et al. 2011, Trautmann and Krasny 2006). Along these lines, Bettinger, Long, and Taylor (2016) recently found that graduate students who teach more frequently are more likely to graduate on time, and more likely to secure academic jobs after graduation, most especially at less-research-focused colleges and universities. The authors acknowledge this result likely reflects selection rather than a causal impact of teaching opportunities on employment outcomes, but argue that from a policy perspective this distinction is of little importance, as it appears departments are successfully identifying graduate students who will teach their undergraduates well, and that signal serves graduate students well on the academic job market. Nonetheless, as Austin (2002) notes, TA roles "sometimes are structured more to serve institutional or faculty needs than to ensure a high quality learning experience for graduate students."

Pedagogical Training: Teaching Future Faculty How to Teach

For fifty years or more, concerns have been raised about American higher education's reliance on graduate student instructors, and more broadly the lack of pedagogical training academic faculty receive enroute to the PhD (Nowlis, Clark, and Rock 1968). The number of studies examining effectiveness of innovative practices in undergraduates' learning continues to increase, and even more papers provide instructions for replicating innovative approaches for teaching specific topics or entire courses, but just as Carroll (1980) reported decades ago, there are relatively few empirical studies investigating effectiveness of various types of pedagogical training for future faculty.

Love Stowell et al. (2015) argue that graduate students, as pre-service faculty, should receive systematic training and practice in pedagogy, leveraging the growing evidence base for teaching practices. This idea is not new: a 2002 report of the Council of Graduate Schools and Association of American Colleges and Universities argued for more planned and structured professional development in teaching for future faculty (Pruitt-Logan, Gaff, and Jentoft 2002). With the increasing trend toward non-tenure-track teaching careers, structured teaching experience may be particularly useful for STEM PhDs. Also, in many STEM fields—most especially in engineering—women PhDs are disproportionately drawn to academia, to teaching-focused positions, and to teaching-focused institutions like liberal arts colleges (Blume-Kohout

2014, Conti and Visentin 2015, Fox and Stephan 2001, Gibbs et al. 2014, Golde and Dore 2004, Shauman 2017). Thus, lack of formal pedagogical training and experience may disproportionately burden women graduates, leaving them relatively less prepared for their typically heavier postdoctoral teaching responsibilities.

Available evidence and Cognitive Apprenticeship Theory support best practice as having future faculty undergo experiential classroom training scaffolded with explicit pedagogical preparation, reflection, and repeated (multiple, periodic) expert classroom observations with feedback (Addy and Blanchard 2010, Greer, Cathcart, and Neale 2016, Bond-Robinson and Rodriques 2006, Nurrenbern, Mickiewicz, and Francisco 1999, Wyse, Long, and Ebert-May 2014, Miller, Brickman, and Oliver 2014, Cherrstrom et al. 2017) However, even though the overwhelming majority of STEM PhDs do serve as TAs at some point during their training, many serve only as graders or laboratory assistants, and few report any structured progression of responsibilities (Pruitt-Logan and Gaff 2004, Austin 2002). A prospective study of master's and PhD students who initially expressed interest in academic teaching careers found that those who held laboratory assistant positions—rather than progressing towards full responsibility for a course—did not identify their activity as teaching, and seemed not to have developed any greater understanding of the role of the teacher, how students learn, or how teachers can facilitate learning through their experiences (Wulff et al. 2004). By contrast, PhD candidates who serve as instructor of record in a course may gain tacit professional knowledge, as well as improving their skills in oral communication to diverse audiences and time management (Love Stowell et al. 2015).

Across several large-scale surveys, STEM PhD students report their programs provide little to no training in how to teach undergraduate students, and often this training consists of a single workshop or seminar rather than extended instruction (Golde and Dore 2004, Fagen and Suedkamp Wells 2004, Schussler et al. 2015, Palmer 2011). Engineering faculty also indicate integration of pedagogical training and instruction for PhD students lags relative to its importance (Jamieson and Lohman 2012).

Graduate students report little systematic or formative feedback on their teaching from faculty mentors (Wulff et al. 2004). DeChenne and colleagues also cite several studies indicating that many graduate teaching assistants who *do* receive professional development for teaching find it is inadequate, and offer a rubric for systematic evaluation of GTA outcomes (DeChenne, Enochs, and Needham 2012, DeChenne et al. 2012). Because, across many fields, a declining share of STEM PhDs are going to go on to academic faculty careers, and because teaching assistantships are associated with longer time-to-degree, some argue pedagogical training is more appropriately received by junior faculty as on-the-job training, rather than added as part of the PhD curriculum. The difficulty for junior faculty then, of course, is finding time and incentive, especially in departments where they perceive research is primary in promotion and tenure decisions (Brownell and Tanner 2012, Lattuca, Bergom, and Knight 2014, Jamieson and Lohman 2012, Froyd et al. 2013).

One promising alternative to the above scenarios is demonstrated in the proliferation of teaching-oriented postdoctoral fellowships. As one example, Harvard's Curriculum Fellows program hires PhD scientists with career goals in teaching or STEM education research to provide evidence-based instructional support to faculty seeking improvement of their course delivery (Gutlerner and Van Vactor 2013). Others, like the National Institutes of Health (NIH) Fellowships in Research and Science Teaching (FIRST) and Seeding Postdoctoral Innovators in Research and Education (SPIRE) programs, combine training in both teaching and research. The FIRST and SPIRE programs also incorporate mentored teaching of an undergraduate course at a Minority-Serving Institution, providing postdocs with experience teaching students from diverse and underrepresented backgrounds (Hue et al. 2010, Keen-Rhinehart et al. 2009). Though program evaluations for FIRST and SPIRE cannot completely rule out selection on ability, postdocs in both programs published at approximately the same rate as non-participating postdocs based at the same research-intensive institutions, despite spending 15-25% of their time on teaching activities.

The combined program further demonstrates to prospective academic employers the fellows' capacity to balance research and teaching responsibilities, just as faculty must do, and unsurprisingly tenure-track placement rates from both programs have been excellent (Keen-Rhinehart et al. 2009, Holtzclaw et al. 2005, Rybarczyk et al. 2011, Rybarczyk et al. 2016).

An extensive review of literature on change strategies for undergraduate STEM instruction found successful interventions last a full semester or more, incorporate performance evaluation, reflective exercises and feedback, with a focus on changing attitudes and beliefs, not just transmitting knowledge (Henderson, Beach, and Finkelstein 2011). Commonalities of FIRST, SPIRE, and similar institution-based programs include their substantially longer duration as compared to typical TA training in PhD programs, their focus on developing teaching practices rather than tools, providing iterative feedback and mentoring, and culmination in the trainee's developing and teaching at least one full course as instructor. Ebert-May et al. (2015) also infer that such programs have relatively greater benefit in terms of actual behavior change for inexperienced teachers—perhaps indicating those who hadn't taught as graduate students had fewer bad habits to break. Finally, although the proliferation of postdoc positions has often been criticized, postdocs who identify skill development as an important factor in choosing their appointment are less likely to be dissatisfied (Miller and Feldman 2015). Reserving intended academics' intensive pedagogical training for the postdoc would seem to fulfill the skill development criterion. Nonetheless, some exposure to pedagogical training in graduate school is associated with greater tendency toward student-centered instruction among faculty (Lattuca, Bergom, and Knight 2014).

In contrast with the above, studies of short-term pedagogical workshops, seminars, and intensive short courses (e.g., over 2-3 days) often report increases in graduate students', postdocs', or new junior faculty's teaching self-efficacy, with some exceptions in the case of inquiry-based pedagogy for new and inexperienced graduate teaching assistants (Wyse, Long, and Ebert-May 2014, DeChenne et al. 2015). Although self-efficacy is an important and arguably necessary characteristic for performance, it is not sufficient, and the limited evidence on instructors' actual subsequent implementation of learned strategies is mixed (Kurdziel et al. 2003, Gardner and Jones 2011, Nurrenbern, Mickiewicz, and Francisco 1999, Latham and Locke 2007). Multiple studies have also found STEM faculty overestimate (or overstate) their use of evidence-based instructional strategies in the classroom (Froyd et al. 2013). Even more troubling, Wulff et al. (2004) found graduate students' sense of self-efficacy as teachers improved over time, *even in the absence of TA experience,* because their teaching confidence improved with mastery of content-area knowledge, as evidenced by successful coursework completion or qualifying exams. Studies that use simple pre-test/post-test designs to evaluate self-efficacy outcomes from pedagogical training interventions thus may also suffer from maturation bias.

Evidence from small case studies indicate that one-on-one classroom observation and feedback with expert consultant-observers may significantly improve faculty members' teaching as measured by student course evaluations (Sword 2014). Where such one-on-one consulting is an institutional requirement for faculty, the evidence is more plausible, as we can rule out selection bias (i.e., only instructors most motivated to improve may choose to seek consults).

Faculty mentoring of graduate TAs may influence the students' desire for a teaching career and willingness to pattern their own teaching practices on that of the professor. Where there is little knowledge of effective research-based instructional practices among the faculty, this lack of knowledge is transmitted to the TA. On the other hand, to the extent their supervising faculty adopt modern, evidence-based teaching techniques and support TAs experimentation, graduate students who serve as TAs can learn some of these approaches (DeChenne et al. 2015). In at least one case, TAs leading student-centered, inquiry-based labs in a course designed around those principles were able to adopt those methods with minimal training (Ryker and McConnell 2014).

One innovative intermediate approach between TA and instructor-of-record experience for STEM PhD students is team-taught instruction. For example, Emory University's "On Recent Discoveries by Emory Researchers" (ORDER) course for first-year undergraduate students was team-taught by ten graduate students with supervision by two faculty members, with each graduate student 'teacher-scholar' preparing and independently teaching a module spanning five class meetings centered on his or her own area of research. Collaboratively identifying themes across their modules' multiple disciplines also encouraged participating graduate students to identify interdisciplinary connections and commonalities in processes of scientific research across their fields. Focusing on their own research area increased confidence for their first attempt at independent instruction, and teaching first-year undergraduates improved the teacher-scholars' self-assessed ability to communicate research effectively. Finally, undergraduate students in the seminar showed improvement in critical thinking and science-related skills, and a significant share went on to pursue research training themselves (Sales et al. 2007, Sales et al. 2006, Hue et al. 2010).

Evaluations of the National Science Foundation's Graduate Teaching Fellows in K-12 Education (GK-12) program similarly found it improved fellows' communication and teaching skills, and was more useful than a typical TA position to the student's career, allowing them to examine alternative career choices (Trautmann and Krasny 2006). Strengths of the program include having the K-12 teacher in the classroom to advise and evaluate effectiveness of fellows' teaching strategies, and weekly seminars where fellows learn new teaching strategies and discuss successes and challenges (Trautmann and Krasny 2006). Applicants to the fellowship program generally express interest in teaching and preference for teaching-focused careers, and report little change in their career goals but stronger skills in teaching inquiry-based science, mentoring and facilitating student research projects, curriculum development, teamwork, educational outreach (explaining STEM research and concepts to public audiences), facilitating group discussions, and time management (Trautmann and Krasny 2006, Gamse et al. 2010). Of those fellows who completed their degrees by Spring 2009, median time-to-degree was 2.7 years for master's students, and 6.0 years for doctoral students (Gamse et al. 2010).

Finally, a survey of graduate student teaching certificate programs found one-third had no explicit teaching experience requirement. About 40% relied solely on credit-bearing courses, while about half required workshops but no credit-bearing courses (von Hoene 2011). The author noted that reflective assessments and end-products—for example, statements of teaching philosophy, syllabi, teaching portfolios—were a 'hallmark' of the certificate programs. However, only half of programs included working with a faculty mentor, and a quarter did not include classroom observation. In one specific case, a two-year intensive Teaching and Learning certificate program improved self-efficacy and confidence using evidence-based methods in the classroom, provided participants with a supportive cohort for socialization, and created positive spillover effects in teaching practices in participants' home departments. Participants in the program were also more likely to receive teaching awards and publish in prestigious higher education journals; however, it is unclear how much of this reflects self-selection (Sword 2014).

Graduate-Level Classroom Instruction

Brint and Clotfelter (2016) relate, “At the classroom level, the sole aim broadly endorsed by the faculty and the public alike is for faculty to provide instruction that contributes to students’ learning.” And yet, evidence indicates that graduate program reputational prestige measures, largely based on faculty research productivity, may substantially diverge from alumni ratings of educational effectiveness (Morrison et al. 2011). In Lovitts’ (2004) study, non-completers were more likely than completers to report dissatisfaction with the contribution of their graduate coursework to their intellectual development, citing disorganized and haphazard instruction, and competitive or combative class discussions.

Qualitative studies with PhD students have indicated a lack of individualized training, and many students felt discouraged from seeking nonacademic professional skills through complementary extradepartmental coursework, internships, or community outreach programs (Fagen and Suedkamp Wells 2004, Callier, Greenbaum, and Vanderford 2015, Szelényi, Bresonis, and Mars 2016).

Retrospective program evaluations by STEM PhDs 10 years after their doctorates called for curricula to be continuously updated to reflect the current state-of-the-art, to include greater breadth and interdisciplinary coursework, and to provide career planning information and relevant training opportunities for those seeking industry, government, or teaching-focused academic careers (Nerad, Aanerud, and Cerny 2004). For example, scientific research and discovery both inside and outside the academy continues to increase use of advances in data science, but standalone data science programs are still in their infancy, and integration of these transdisciplinary methods into other STEM graduate curricula is slow (Aikat et al. 2017).

At the same time, an “overcrowded curriculum” with a growing amount of required discipline-specific content poses a threat to inclusion of additional skills- and process-based training (Jamieson and Lohman 2012). Faculty who “have devoted time, effort, and emotion” into delivering existing curriculum may also resist implementing change, and institutional policies—for example, regarding course approvals, scheduling, and registration—may constrain curriculum experimentation and reform (Dasgupta, Symes, and Hyman 2015).

Since the Accreditation Board for Engineering and Technology (ABET) shift to outcomes-based accreditation for U.S. undergraduate engineering programs in the late 1990s, many engineering faculty and program chairs have spent more time learning about evidence-based pedagogical methods, and have increased use of collaborative, experiential, and inquiry-based active learning methods in their classrooms, including at the graduate level (Jamieson and Lohman 2012).

At the undergraduate level, meta-analysis of over fifty studies indicated that hybrid or blended instruction significantly improves student performance versus traditional face-to-face lectures (Means et al. 2009). In these courses, basic materials like readings or short video lectures are provided to students online, to be completed before coming to class, and accountability is enforced through online or in-class quizzes. Class time is then used for interactive collaborative learning exercises, and for student-centered discussions providing substantial time for students to ask questions.

Rigorous studies of flipped-hybrid instruction in graduate biomedical sciences education have found similar strong positive effects for student performance, with students praising the greater focus on understanding and applying the material, self-reflection, and accountability (Tune, Sturek, and Basile 2013, Khanova et al. 2015, Belfi et al. 2015, O'Connor et al. 2016). Qualitatively, graduate engineering students also exhibited higher engagement in a flipped classroom (Zalewski and Schneider 2016). However, Khanova et al. (2015) find effectiveness of (and student satisfaction with) the flipped classroom model is eroded when course design and implementation are misaligned, for example when instructors’ in-class activities remained largely lecture-based, and when assessments focused more on memorized knowledge than critical thinking and applications. For example, Zalewski and Schneider (2016) acknowledge not having used online or in-class quizzes to ensure accountability for class preparation activities, which may have reduced effectiveness of their first flipped classroom implementation. More generally, inappropriate or superficial adoption of pedagogical innovations coupled with inadequate measures of learning effectiveness may lead faculty to conclude these methods are ineffective (Besterfield-Sacre et al. 2014). Accordingly, a recent systematic review found only mixed evidence that flipped classrooms improve knowledge acquisition beyond traditional methods in graduate medical education, and concluded that more research is needed on effects on long-term knowledge retention and application (Chen, Lui, and Martinelli 2017). Consistent with this critique, Jensen, Kummer, and Godoy (2015) found similar outcomes for both flipped and non-flipped undergraduate biology classes when both used active learning approaches.

Team-based learning (TBL) is a specific form of active learning conducive to hybrid instruction environments, with specific characteristics including instructor pre-formed groups designed to maximize within-group heterogeneity and cross-group homogeneity, pre-class preparation enforced by some accountability measure, adaptive instruction to fill in any observed gaps in student knowledge, and finally group application exercises. A recent meta-analysis of 89 studies of TBL found significant gains in subsequent exam performance and course grades, as compared to traditional didactic instruction. Notably, over one-third of these studies were medical education, and over half were graduate health sciences courses. Limited evidence, again from graduate medical education, also indicates TBL improves both academic performance and attitudes toward teamwork in laboratory courses (Huitt, Killins, and Brooks 2015). More generally, student-centered pedagogy that is adaptive and differentiated to students' learning needs improves student outcomes, and TBL along with other active learning instructional strategies have been found, in some studies, to increase academic achievement among students with diverse prior preparation (Sisk 2011, Krupat et al. 2016). For graduate students in engineering, active learning pedagogies may also intentionally include industry internships and service learning (Jamieson and Lohman 2012).

Short-format courses—encompassing boot camps, nanocourses, and workshops—often report positive impacts on participants' self-efficacy, but evidence for long-term effects is both limited and mixed. Feldon and colleagues (2017) found no difference in research self-efficacy, scholarly writing, or publication productivity between participants versus non-participants in summer bridge or pre-matriculation boot camp programs for PhD students in life sciences. Unfortunately, the authors' stated conclusions are only consistent with the presented statistical evidence if one believes there are no systematic, baseline differences between students who choose to participate in such programs, and those who do not. Identity-based controls (e.g., gender, race and ethnicity) are inadequate to this task. If students self-select into bridge and boot camp programs due to their real or perceived needs for remediation, or to enhance community-building and professional socialization for individuals from underrepresented groups, then the authors' finding that outcomes for these students do not significantly differ from others' could equally be interpreted as a success. Bridge programs for students from underrepresented groups may also help them build supportive peer networks and social capital to mitigate the unequal access to resources these students often find—e.g., for faculty interaction and research experiences in faculty members' labs—as they progress through their programs (Winkle-Wagner and McCoy 2016, Amelink et al. 2016, Gardner 2009b).

A distinction must be made between short-format workshops offered at times convenient for the program, versus the more effective student-centered 'just-in-time' teaching approaches, versus intensive classroom-based seminars and workshops followed by continuing practicum and periodic reflection exercises. For example, nanocourses can permit students to learn technical skills "just-in-time" when the need for those skills arises, and can also provide flexibility for continuous updating and supplementation of curriculum content to reflect new scientific advances (Bentley, Artavanis-Tsakonas, and Stanford 2008, Gutlerner and Van Vactor 2013). When made open to the broader academic community—including advanced graduate students, postdocs, practitioners, and faculty—these courses serve as continuing professional development, but can also promote professional socialization and engagement for students otherwise focused on their theses. One program found graduate students who took nanocourses for credit were most likely to have done so for topics and techniques relevant to their thesis, and for the opportunity to interact with faculty mentors; however, over two-thirds of students said they attended as an antidote to specialization inherent in graduate training, to learn something new that was unrelated to their thesis or prospective postdoc research (Bentley, Artavanis-Tsakonas, and Stanford 2008).

Finally, a recent review of literature on interdisciplinary graduate programs found that while some studies have considered the problem of interdisciplinary graduate student supervision from a theoretical perspective, only very few have examined empirical evidence, and these primarily have focused on evaluation criteria for interdisciplinary work, rather than processes for and outcomes from interdisciplinary graduate training (Vanstone et al. 2013).

Learning Communities and Cohorts

Over the past ten years, some graduate departments have explored use of the Learning Community (LC) model, originally designed for undergraduate education. In undergraduate education, residential learning communities focusing on STEM have not been found to increase persistence in STEM majors (Soldner et al. 2012). For example, one such program in environmental and earth sciences emphasizes integration of core coursework and multidisciplinary learning, a sense of community as the new graduate cohort progresses through core courses and socializes together, use of active learning practices in the classroom, and guided self-reflection (Romsdahl and Hill 2012). The LC works collaboratively in one or more teams on a project that integrates their core coursework and has regional and policy importance, and presents their findings in written papers and conference presentations. This process encouraged students to reflect on the challenges of collaborative writing, as a research skill. However, both terms—'cohort' and 'learning community'—are applied loosely, with 'cohort' often referring simply to an incoming class of graduate students who progress through first-year core classes together, and 'learning community' describing a range of social, professional development, integrative and reflective activities, organized at the department or institutional level. Given the lack of consistency in terminology, it appears there is little published evidence to evaluate regarding LC models and practices at the graduate level. One limited exception to this is for distance-education programs offered primarily online, some of which include short periods with the cohort in-residence. Such programs are of potential interest for further research as they may provide greater access for students who have family, geographic, or financial constraints, or students with disabilities precluding relocation and/or full-time study.

Pedagogy and Curriculum: Responsible Conduct of Research

Though some pre-test/post-test studies have indicated improvements in capacity for ethics-based decisions with Responsible Conduct of Research (RCR) training, one large-scale study found the most common method used for online RCR instruction had little or no positive impact on decision ethicality and socio-behavioral responses to ethical dilemmas (Antes et al. 2010). Case-based learning (CBL or, alternatively, case-based reasoning, CBR) is commonly used in professional graduate programs, for example business, medicine, engineering, and teacher training programs. Fewer examples are found for graduate programs in the natural sciences, though one consistent application of CBL/CBR across STEM programs is in ethics education, including to fulfill federal agency requirements for RCR training (Thiel et al. 2013, Johnson et al. 2012). However, a recent study focusing on environmental science graduate programs concluded student competence is generally not assessed (Hall et al. 2017), and preliminary review of the literature indicates little is known about the qualities of CBL/CBR case studies used in ethics training that would increase retention of principles and their real-world transfer. One study employing TBL for RCR training, using the same ethics assessment as in Antes et al. (2010), demonstrated positive improvements in ethicality of professional practices, but mixed results for metacognitive reasoning, and no benefits for study conduct, data management, or business practices (McCormack and Garvan 2014).

Pedagogy and Curriculum: Writing

Writing—that is, effectively communicating research results—is fundamental to value creation in scientific research. In some sectors, evidence indicates that graduate students have greater self-confidence in their writing skills than their prospective employers think warranted. However, particularly in the experimental sciences, the phrase ‘write it up’ seems to marginalize the importance of writing in research (Kamler and Thomson 2006). Students also identify writing as isolating and a grueling process (Mantai 2015, Aitchison et al. 2012).

Pedagogically, embedding learning to write within the subject context is better than providing it as an add-on, for example through workshops. Participants in one structured writing group—primarily PhD students, from a wide variety of disciplines—reported that the regular social interaction around text and sense of mutual obligation in preparing for group meetings improved their writing skills and productivity, including via critique of peers’ texts (Aitchison 2009). In contrast with mixed evidence found at the undergraduate level for effectiveness of peer critique and feedback in improving writing, for doctoral students learning to give and receive useful critiques have been found valuable for their learning (Aitchison 2009, Caffarella and Barnett 2000). Given that giving and receiving peer review is an essential part of most researchers’ practice, it would seem that providing doctoral students with this experience serves at least three desirable objectives: reducing writers’ feelings of isolation, improving their own writing skills and productivity, and learning to critique and accept critique from others.

Co-authorship with supervisors is ‘standard and important’ in many fields of science, but evidence suggests early research collaboration and ‘scaffolding’ co-writing experiences are important in building students’ subsequent research productivity and learning to ‘speak’ in discipline-specific ways. For example, as Kamler (2008) discusses, during a review, revise and resubmission process “the supervisor’s confident persistence accomplished a kind of identity work for the student, allowing her to perform the resilient scholar, even before she felt like one.” For some supervisors, students’ writing for publication is explicitly incorporated in their research mentoring practice (Aitchison et al. 2012). This notion is supported by Paglis, Green, and Bauer (2006) empirical results showing statistical correlation between psychosocial mentoring (with its attendant effects on self-efficacy) and early research collaboration (with its attendant effects on research productivity). Qualitative evidence suggests that institutions that clearly value and have explicit expectations for student co-authorship—for example, embedding it in their tenure and promotion criteria—incentivize faculty to engage students in the writing process (Maher et al. 2013). However, many supervisors do not consider themselves expert in writing or teaching of writing, especially for non-native English speakers (Aitchison et al. 2012). Thus supervisors are in tension, serving as ‘gatekeeper’ accountable for the quality of students’ written work submitted for publication, while often feeling inadequate to teach or develop students’ writing skills, such that for many faculty their preferred solution is simply to edit or rewrite the paper directly (Aitchison et al. 2012).

Individual Development Plans

Substantial evidence exists demonstrating positive effects of goal-setting on task performance, and high expectations combined with supervisor support likewise yield positive effects on employee performance (Latham and Locke 2007). Learning and performance goals are most effective when proximal goals are specific and attainable, and when the goal-setter has necessary information and feedback to advance toward their goals. More broadly, self-determination theory proposes that workers experience greater satisfaction and demonstrate greater persistence when they feel ownership over their goals, and derive a sense of accomplishment and competence from their activities (Lewis et al. 2016). Leveraging this literature, individual (or personal) development plans (IDPs) with specified learning and/or performance goals are commonly used in industry and government for personnel evaluation and career management (Grohnert, Beausaert, and Segers 2014). The Federation of American Societies for Experimental Biology (FASEB) recommends use of IDPs for postdocs, and the American Association for the Advancement of Science (AAAS) now offers a web-based myIDP tool[1] that graduate students and postdocs can use to explore 'fit' of alternative careers based on their values and interests, assess personal strengths and weaknesses, and create a list of skill- and project-oriented near-term goals for their career advancement.

Very little rigorous evidence appears to be available regarding the uptake, use, and effectiveness of IDPs or other formal, written mentoring and professional development plans for graduate education. However, concerns have been raised for years about lack of systematic professional development and feedback in doctoral education (Austin 2002). Even among postdocs, one recent study found less than half of postdocs and only 20% of FASEB mentors were aware of IDPs (Hobin et al. 2014). The primary reason mentors said they had not developed a written career plan with their postdocs was because their postdocs hadn't asked for it, and many postdocs who knew about IDPs but didn't make one with their mentors said they were uncomfortable bringing up the topic.

Another study suggests that early introduction to IDPs—in this case, among post-baccalaureate students in a research bridge program—may increase graduate students' confidence and efficacy in initiating conversations on professional development and eliciting feedback from their mentors. In a study of UNC's PREP program, the IDP process seemed to help advisors to be more aware of their mentees' strengths and weaknesses, and allowed students to clarify and better meet expectations so they earned stronger letters of recommendation (Hall et al. 2016). Alumni of the program also reported that the IDP process helped prepare them for frank conversations with their subsequent PhD advisors.

Because myIDP is a subjective self-assessment, its usefulness relies in part on the trainee's relationship to mentors who can provide honest, informed, and constructive feedback on the trainee's self-assessment, and who can help the trainee access resources necessary to achieve their career development goals. On the other hand, effectiveness of IDPs for graduate student trainees may be limited, if faculty mentors are uninformed (due to their own lack of experience outside academia) with regard to skills necessary for non-academic career paths, or are otherwise uninterested in the content or process (Callier, Greenbaum, and Vanderford 2015, Grohnert, Beausaert, and Segers 2014). In addition, in broader workplace contexts where IDP use is required, many employees indicate the tool is ineffective or unnecessary, in that it does not stimulate learning or feedback that would not otherwise take place (Grohnert, Beausaert, and Segers 2014). Still, by reflecting on and systematically cataloging their skill development in conversation with their mentors—especially transferable or generalizable skills—trainees may improve their ability to identify and articulate those transferable skills to non-academic employers (Manathunga and Lant 2006).

[1] AAAS ScienceCareers myIDP website: https://myidp.sciencecareers.org, last accessed 24 October 2017.

Mentoring

A substantial literature has shown the primary importance of the student-advisor relationship in graduate students' retention, academic success, research productivity, timely completion, professional socialization, self-efficacy, career aspirations and commitment, and overall satisfaction (O'Meara, Knudsen, and Jones 2013, Mollica and Nemeth 2014). Abedin et al. (2012) identify 19 specific competencies needed across six thematic areas, including communication and relationship management, psychosocial support, career and professional development, professional enculturation and scientific integrity, and development as both a researcher and a principal investigator, where the latter includes learning to manage a research team effectively. Regular, directive interactions between faculty mentors and graduate students are instrumental to students' research skill development (Feldon et al. 2016).

Paglis, Green, and Bauer (2006) note that supervisory mentors are "key individuals in the newcomer socialization process," providing both psychosocial support and career advancement, the latter by helping the newcomer build connections within the organization. However, in graduate programs it appears that faculty may focus their mentoring efforts on the most-qualified candidates as measured by GRE scores and prior research experience at time of entry, and such positional favoritism may be more stark in top-ranked programs (Conley and Önder 2014, Noy and Ray 2012).

After controlling for entering student attitudes, GRE scores, and prior research experience, one study found adviser mentoring did not have any positive impact on PhD students' commitment to a research career. However, PhD students who reported collaborative mentoring by the PhD advisor on research in year two had significantly higher research productivity four years later, while those who reported psychosocial mentoring in year two had greater sense of research self-efficacy, which was found to be a positive predictor of research submissions (Paglis, Green, and Bauer 2006).

For engineering research groups, Crede and Borrego (2012) found no impact of research group size (i.e., number of graduate students and postdocs per faculty member) on graduate students' learning outcomes, though in larger groups more learning came from interactions with more experienced peers. By contrast, smaller research groups (e.g., N=2 students per faculty member) resulted in more direct faculty contact, quicker time to establish their research projects, and generally more secure funding, but less peer-to-peer interaction between students, and less enculturation into a community of practice. For lower-performing students, being part of a larger research team including more advanced student-peers allows them to supplement their knowledge and reliance on their advisor, increasing their collaborative research activity (Feldon et al. 2016).

Correspondence between PhD student self-efficacy, mentor appraisal, and performance-based competence in research activities is mixed (Feldon et al. 2015). Discordance between student and advisor perceptions may reflect lack of adequate feedback, but Feldon et al. (2015) also indicate that advisors' perceptions of their mentees' skills and knowledge may also be skewed. If faculty advisors are not truly engaged as mentors and indeed lack knowledge of their students' capabilities, intentionality and guided professional development of students as 'cognitive apprentices' may be compromised.

PhD students most satisfied with their advisors indicated these mentors challenged and stimulated their thinking, were helpful and encouraging, were enthusiastic about the student's research, contributed to the student's professional development, were approachable and generous with their time, gave appropriate amounts of freedom and direction, and provided regular and constructive feedback on both research and academic progress (Lovitts 2004, Zhao, Golde, and McCormick 2007). Failure to provide any of these gave cause for dissatisfaction, but some dissatisfied students also sought a more personal, warm, open and informal relationship with their advisor (Lovitts 2004). This last point may reflect unrealistic or inappropriate expectations on the part of some graduate students regarding the professional advising relationship; however, many have noted the lack of training faculty receive for their advising role, in particular with regard to personal-social or emotional competencies (O'Meara, Knudsen, and Jones

2013). In the absence of psychological safety, honest and open conversations about professional development may not take place (Grohnert, Beausaert, and Segers 2014).

On average, women PhD students are less satisfied with the advising relationship, and in STEM fields women students are significantly less likely to say their primary advisor is respectful of their ideas (Zhao, Golde, and McCormick 2007, Noy and Ray 2012). However, women and underrepresented minority PhDs in STEM and social science fields appear also more often to seek out and receive both affective emotional support and instrumental support from a secondary advisor (Noy and Ray 2012).

One alternative supervisory structure or set of techniques borrowed from professional development and management literatures is doctoral coaching. Coaching is consistent with use of the IDP, as the approach encourages a focus on goals, and seeks to improve autonomy, internal locus of control, and cognitive-behavioral responses (Godskesen and Kobayashi 2016). However, unlike the student's faculty advisor, the coach is not meant to judge the student's work, is not invested in the student's decision to continue or quit their graduate program, and has no formal power or authority. In one Danish study of doctoral coaching, PhD students reported mainly discussing topics like self-confidence and self-esteem, workplace relationships including with their supervisors, planning and time management, self-motivation, stress, own and others' expectations, and the writing process (Godskesen and Kobayashi 2016).

Another longitudinal randomized trial of a coaching intervention demonstrated significant increases in perceived achievability of academic research careers, and significant defense against reduction in desirability of academic research careers, for women and URM PhD students in the biomedical sciences (Williams, Thakore, and McGee 2016). Coaches met with groups of students, not one-on-one, and key aspects of the career coaching intervention included systematic and social-science-based training in diversity and providing 'safe space' for discussions, dedicated time and space for career mentoring conversations, and independent advice (coaches from a different institution, no work-based relationship with the students coached).

Mentor Training: Can Emotional Competencies Be Taught?

A randomized controlled trial across 16 academic health centers investigated impact of an eight-hour, case-based curriculum adapted from *Entering Mentoring*, a program developed for mentors in biological sciences. The curriculum focused on six core competencies: maintaining effective communication, establishing and aligning expectations, assessing mentees' understanding of scientific research, addressing diversity in mentoring relationships, fostering independence, and promoting career development. Faculty mentors receiving the intervention self-reported significantly greater competency ex-post in maintaining effective communication (e.g., active listening and providing constructive feedback), establishing expectations, and helping mentees set goals for their professional development. Interestingly, retrospective 'pre-test' indicates that faculty participating in the intervention re-evaluated their initial competencies in several aspects lower than in actual pre-test, most especially in various aspects of effective communication, aligning expectations, and diversity (accounting for differences in background and biases). Accordingly, "intervention group mentors reported a significantly greater degree of change in their awareness of mentoring competencies and need to implement behavioral changes." (Pfund et al. 2014) The Mentoring Competency Assessment instrument developed by the research team also appears to exhibit strong validity (Fleming et al. 2013).

Unfortunately, most studies of mentor training interventions—even those that purport to strong experimental design, for example randomized controlled trials—are plagued by selection bias, as comparisons are made only among those who applied to participate in the program (see, e.g., Gandhi and Johnson (2016)). Outcome measures typically include faculty mentors' self-efficacy and self-reported skills, most often directly following the intervention, though a couple of studies have examined longer-term outcomes. For example, the University of California, San Francisco, Clinical and Translational Science Institute's Mentor Development Program (MDP) found immediate positive impact on faculty mentors' self-efficacy and self-assessed mentoring skills, and in a follow-up survey 1-3 years later MDP graduates reported that MDP had helped them work more effectively with mentees to determine long-term career plans and provide emotional support (Feldman et al. 2012).

Where both mentor and protégé are included in the intervention, nominally to evaluate actual behavior changes from the protégé's perspective, their participation requires joint willingness, which may imply existence *a priori* of a healthy, high-functioning mentoring relationship (Lewis et al. 2016). However, not all faculty are interested in serving as mentors, except to the extent that it helps promote their productivity and reputation (Brunsma, Embrick, and Shin 2017). These realities limit generalizability of the curriculum and intervention, for example if such training were broadly mandated. In addition, many studies of mentoring interventions rely on self-reported and immediate or short-term measures of self-efficacy, so effectiveness for sustained behavioral change is unknown. One careful study of a diversity mentor-training intervention found initial improvement reported about a month after training, but no significant impact on mentoring behavior after a year (Lewis et al. 2016). Such findings reinforce the notion discussed above, that short-format training is typically insufficient to effect behavioral change.

Some mentoring interventions seek also (or instead) to improve the graduate student's intrapersonal communication and leadership skills. However, here again, shorter interventions may be inadequate to produce behavioral change, particularly given the power dynamic and implied onus to "manage up" toward relatively untrained faculty advisors. For example, Brockman, Nunez, and Basu (2010) report on a conflict resolution workshop with pre- and post-test evaluations of graduate student participants' conflict management preferences, and found significant change in students' collaboration versus accommodation and avoidance preferences and self-efficacy in navigating the student-advisor power differential, but only modest change in their application of the techniques. They conclude that while workshop-based exposure to communication techniques may be useful for setting expectations and giving students more options/tools, a single three-day workshop does not create lasting change in the dominant avoidant/accommodating behavior of students.

None of the above should be read to imply that mentoring, leadership, and communication skills cannot be taught. To the contrary, as one example, a study of a MBA program change requiring a new leadership course found improvement in both self-rated and direct behavioral measures among its graduates (Boyatzis, Stubbs, and Taylor 2002). The more fundamental question is whether faculty supervisors most in need of such training will choose to seek it, and what incentives and support their institutions will provide to ensure faculty supervising graduate students are motivated to become better mentors.

Career Choices and Program Alternatives

A survey of 8,373 University of California system doctoral students across nine campuses found 51% of women and 45% of men were married or partnered, and more than three-quarters of them expressed concern about 'family friendliness' of their career choices, with tenure-track faculty positions in research-intensive universities rated least family-friendly of all options, including policy or managerial careers, non-academic research careers, faculty careers at teaching-intensive colleges, and non-tenure-track faculty positions (Mason, Goulden, and Frasch 2009). The most common reasons cited as "very important" by both men and women in their shifting career preference was "negative experience as PhD student," but women also disproportionately cited professional activity as too time-consuming, and a combination of family and geographic location concerns. This concern is echoed in national data as well: women STEM PhDs are more likely than their male counterparts to have a partner whose job requires at least technical expertise of a bachelor's in STEM, and are more likely to cite location and family as reasons for changing jobs (Blume-Kohout 2014).

DeFraine, Williams, and Ceci (2014) argue that Millennials differ from earlier generations in their preference for "work/life-interaction"—that is, integration of work and non-work or family lives. This contrasts with a "traditional" model for (academic) STEM careers: "long hours in the lab in addition to teaching, and in general, dedication to one's work above all else." This depiction is consistent with other studies demonstrating that—in some fields, particularly those like computer sciences and engineering that remain overwhelmingly male-dominated—faculty advisors prize students who are "motivated" and driven to "work very hard" (Gardner 2009a). Millennials were significantly more attracted by the prospect of joining a lab that emphasized collaboration, teamwork, and cohesiveness, but faculty who were themselves trained in a traditional work-focused lab were far more likely to use a work-focused approach in training their own graduate students (DeFraine, Williams, and Ceci 2014). This may be one reason late-stage STEM PhDs are more likely to say an academic research career is unattractive, compared to early-stage PhDs (Gibbs et al. 2014, Sauermann and Roach 2012, Fuhrmann et al. 2011).

In biology, chemistry, and physics, as well as in the social sciences, recent evidence shows PhD students feel their advisors encourage or strongly encourage academic research careers, and PhD students feel relatively uninformed about other, non-academic career options (Roach and Sauermann 2010, Sauermann and Roach 2012, Morrison et al. 2011). More research is needed to understand the causes of significantly greater switch towards preference for non-research careers observed between PhD entry and completion among underrepresented minority women in STEM, and the implications of these disparate career preferences for retention and training paths (Gibbs et al. 2014).

Internships and certificate programs may help both to inform students of their options and help them individualize their graduate education. For example, the University of Illinois' Certificate in Entrepreneurship and Management, targets graduate and professional students in life sciences, to train them in the business side of biotechnology venturing (Agarwal and Sonka 2010). Industry sponsorship reduces the tuition cost to students, and students receive assistance in obtaining internships and other experiential activities. A similar Certificate in Business program is offered to graduate students seeking nonbusiness degrees. However, while many published papers describe campus initiatives, their popularity and enrollments, very few studies systematically examine student outcomes.

A recent survey of U.S. engineering programs found they were comfortable collaborating on curricular innovations with industry, future employers, and other academic STEM fields (Jamieson and Lohman 2012). Over two-thirds of faculty indicated departments 'routinely' collaborated with industry and employers for educational innovations including lab work and other research experiences, design competitions, and co-operative education and internships (Jamieson and Lohman 2012). Internationally, previous studies of science and engineering PhD students have found little if any effect on student productivity or career ambitions from industry collaborations, but students reported greater satisfaction and optimism about their career prospects, more positive attitudes toward non-academic careers, and were also more likely to rate their supervisor's interpersonal skills as satisfactory or very satisfactory (Harman 2002, Thune 2009, Louis et al. 2007). However, while U.S. engineering programs did support continuing professional development for faculty, the report found less enthusiasm for faculty members having actual industry experience (Jamieson and Lohman 2012).

Qualitative and quantitative studies indicate that exposure to industry-funded R&D, technology transfer, and entrepreneurial start-ups in graduate school—for example, working as a research assistant on several industry-funded projects—encourage STEM PhDs to consider alternative (non-academic) career paths, and moreover, this exposure appears to be more salient in subsequent career choices for women STEM PhDs than for men (Szelényi, Bresonis, and Mars 2016, Blume-Kohout 2017, Roach 2017, Mangematin 2000). Despite early concerns that industry sponsorship might have a chilling effect on graduate students' research, this has not been borne out (Behrens and Gray 2001).

Finally, in addition to support from science funding agencies, government-provided education subsidies include various tax incentives for education expenditures, including a tax exemption for employer-provided tuition assistance. When a tax exemption for employer-provided aid is available, students ages 24 to 30 who are employed full- or part-time are more likely to enroll and persist in full-time graduate programs, especially at public universities (Bednar and Gicheva 2013). More research is needed to understand the rationales and incentives for employer-supported graduate training, and whether there exists room to grow such private sector support of human capital investments.

Professional Science Master's Programs

As a final note, this review failed to uncover useful evidence to inform design or policies for academic, research-oriented master's degree programs. While many professional master's programs in fields like social work, nursing, counseling, and K-12 education yield significant educational research efforts, it appears less interest and effort is focused toward terminal master's degrees, which in many fields seem to represent some failure to transition from graduate coursework to doctoral research. De facto, for some students graduating from less rigorous bachelor's degree programs, a master's degree may serve simply to make the student competitive in the job market vis-à-vis students graduating from more prestigious schools.

One exception deserves mention, given the trends noted above with regard to curriculum relevance and career choices. Professional Science Master's (PSM) programs have addressed students' desire for 'real world' practical experience, and alumni surveys indicate strong post-graduation placement results with roughly 90 percent attaining jobs "somewhat related" or "closely related" to their degree (Komura 2014). In principle, PSM programs and the professional development experiences they incorporate should be directly informed and supported by industry. These programs typically leverage expertise across multiple departments, to offer multidisciplinary programs like North Carolina State University's PSM program in Microbial Biotechnology (Hamilton, Luginbuhl, and Hyman 2012). Illinois also offers three coursework-based PSM programs that combine science or mathematics courses with business/professional courses, an industry seminar series, and an internship. Programs include agricultural production, bioenergy, food science and human nutrition. There, too, courses for the programs are taught by multiple different departments.

PSM programs' internships, capstone projects, and semester-long team practicum projects hone skills in collaboration, teamwork, customer orientation, writing, and oral presentation (Cramer and Hamilton 2017). Recognizing the relevance, placement success, and long-run sustainability of these programs, the National Academies recommended several actions to accelerate development of PSM programs across the country (NRC 2008).

References

Abedin, Zainab, Ewelina Biskup, Karin Silet, Jane M. Garbutt, Kurt Kroenke, Mitchell D. Feldman, Richard McGee, Jr., Michael Fleming, and Harold Alan Pincus. 2012. "Deriving competencies for mentors of clinical and translational scholars." *Clinical and Translational Science* 5 (3):273-280.

Addy, Tracie M., and Margaret R. Blanchard. 2010. "The problem with reform from the bottom up: instructional practises and teacher beliefs of graduate teaching assistants following a reform-minded university teacher certificate program." *International Journal of Science Education* 32 (8):1045-1071.

Agarwal, Rajshree, and Steven Sonka. 2010. "Different strokes for different folks: university programs that enable diverse career choices of young scientists." In *Spanning Boundaries and Disciplines: University Technology Commercialization in the Idea Age*, edited by Gary D. Libecap, Marie Thursby and Sherry Hoskinson, 139-164. Bingley, UK: Emerald Group Publishing Limited.

Aikat, Jay, Thomas M. Carsey, Karamarie Fecho, Kevin Jeffay, Ashok Krishnamurthy, Peter J. Mucha, Arcot Rajasekar, and Stanley C. Ahalt. 2017. "Scientific training in the era of big data: a new pedagogy for graduate education." *Big Data* 5 (1):12-18. doi: 10.1089/big.2016.0014.

Aitchison, Claire. 2009. "Writing groups for doctoral education." *Studies in Higher Education* 34 (8):905-916.

Aitchison, Claire, Janice Catterall, Pauline Ross, and Shelley Burgin. 2012. "'Tough love and tears': learning doctoral writing in the sciences." *Higher Education Research & Development* 31 (4):435-447. doi: 10.1080/07294360.2011.559195.

Amelink, Catherine T., Benjamin David Lutz, Marvin K. Karugarama, and John J. Lesko. 2016. "Graduate summer bridge program: building community and preparedness for success among engineering graduate students." American Society for Engineering Education, New Orleans, LA.

Antes, Alison L., Xiaoqian Wang, Michael D. Mumford, Ryan P. Brown, Shane Connelly, and Lynn D. Devenport. 2010. "Evaluating the effects that existing instruction on Responsible Conduct of Research has on ethical decision making." *Academic Medicine* 85 (3):519-526.

Austin, Ann E. 2002. "Preparing the next generation of faculty: graduate school as socialization to the academic career." *Journal of Higher Education* 73 (1):94-122. doi: 10.1353/jhe.2002.0001.

Bartolone, Jake, Marie L. Halverson, Thomas B. Hoffer, Gregory C. Wolniak, Lisa Setlak, E. C. Hedberg, Evan Nielsen, Vi Nhuan-Le, and Melissa Yisak. 2014. Evaluation of the National Science Foundation's Graduate Research Fellowship Program. Chicago, IL: National Opinion Research Center (NORC) at the University of Chicago.

Bedard, Kelly, and Douglas A. Herman. 2008. "Who goes to graduate/professional school? The importance of economic fluctuations, undergraduate field, and ability." *Economics of Education Review* 27 (2):197-210.

Bednar, Steven, and Dora Gicheva. 2013. "Tax benefits for graduate education: Incentives for whom?" *Economics of Education Review* 36:181-197. doi: 10.1016/j.econedurev.2013.07.001.

Behrens, Teresa R., and Dennis O. Gray. 2001. "Unintended consequences of cooperative research: impact of industry sponsorship on climate for academic freedom and other graduate student outcome." *Research Policy* 30 (2):179-199. doi: 10.1016/s0048-7333(99)00112-2.

Belfi, Lily M., Roger J. Bartolotta, Ashley E. Giambrone, Caryn Davi, and Robert J. Min. 2015. "'Flipping' the introductory clerkship in radiology: impact on medical student performance and perceptions." *Academic Radiology* 22 (6):794-801.

Bentley, Anna M., Spyros Artavanis-Tsakonas, and Jennifer S. Stanford. 2008. "Nanocourses: a short course format as an educational tool in a biological sciences graduate curriculum." *CBE-Life Sciences Education* 7 (2):175-183. doi: 10.1187/cbe.07-07-0049.

Bersola, Samuel H., Ellen Bara Stolzenberg, Janice Love, and Kevin Fosnacht. 2014. "Understanding admitted doctoral students' institutional choices: student experiences versus faculty and staff perceptions." *American Journal of Education* 120 (4):515-543. doi: 10.1086/676923.

Besterfield-Sacre, Mary, Monica F. Cox, Maura Borrego, Kacey Beddoes, and Jiabin Zhu. 2014. "Changing engineering education: views of U.S. faculty, chairs, and deans " *Journal of Engineering Education* 103 (2):193-219.

Bettinger, Eric P., Bridget Terry Long, and Eric S. Taylor. 2016. "When inputs are outputs: The case of graduate student instructors." *Economics of Education Review* 52:63-76. doi: 10.1016/j.econedurev.2016.01.005.

Blume-Kohout, Margaret E. 2014. Understanding the Gender Gap in STEM Fields Entrepreneurship. Washington D.C.: U.S. Small Business Administration Office of Advocacy Report.

Blume-Kohout, Margaret E. 2017. "Gender and postgraduate training environments in STEM fields entrepreneurship." In *Gender and Entrepreneurial Activity*, edited by Albert N. Link, 289-307. Cheltenham, UK: Edward Elgar.

Blume-Kohout, Margaret E., and Dadhi Adhikari. 2016. "Training the scientific workforce: Does funding mechanism matter?" *Research Policy* 45 (6):1291-1303. doi: 10.1016/j.respol.2016.03.011.

Blume-Kohout, Margaret E., and John W. Clack. 2013. "Are graduate students rational? Evidence from the market for biomedical scientists." *PLoS ONE* 8 (12). doi: 10.1371/journal.pone.0082759.

Bond-Robinson, Janet, and Romola A. Bernard Rodriques. 2006. "Catalyzing graduate teaching assistants' laboratory teaching through design research." *Journal of Chemical Education* 83 (2):313. doi: 10.1021/ed083p313.

Boyatzis, Richard E., Elizabeth C. Stubbs, and Scott N. Taylor. 2002. "Learning cognitive and emotional intelligence competencies through graduate management education." *Academy of Management Learning and Education* 1 (2):150-162.

Brint, Steven, and Charles T. Clotfelter. 2016. "U.S. higher education effectiveness." *RSF: The Russell Sage Foundation Journal of the Social Sciences* 2 (1):2-37.

Brockman, Julie L., Antonio A. Nunez, and Archana Basu. 2010. "Effectiveness of a conflict resolution training program in changing graduate students style of managing conflict with their faculty advisors." *Innovative Higher Education* 34 (4):277-293.

Brownell, Sara E., and Kimberly D. Tanner. 2012. "Barriers to faculty pedagogical change: lack of training, time, incentives, and… tensions with professional identity?" *CBE-Life Sciences Education* 11 (3):339-346.

Brunsma, David L., David G. Embrick, and Jean H. Shin. 2017. "Graduate students of color: race, racism, and mentoring in the white waters of academia." *Sociology of Race and Ethnicity* 3 (1):1-13.

Caffarella, Rosemary S., and Bruce G. Barnett. 2000. "Teaching doctoral students to become scholarly writers: the importance of giving and receiving critiques." *Studies in Higher Education* 25 (1):39-52.

Callier, Viviane, Steve Greenbaum, and Nathan L. Vanderford. 2015. "The traditional training of PhDs threatens the technology transfer and entrepreneurship pipeline while innovative programs provide unique recovery opportunities." *2015* 2:51-58.

Carroll, J. Gregory. 1980. "Effects of training programs for university teaching assistants: a review of empirical research." *Journal of Higher Education* 51 (2):167-183.

Chen, Fei, Angela M. Lui, and Susan M. Martinelli. 2017. "A systematic review of the effectiveness of flipped classrooms in medical education." *Medical Education* 51 (6):585-597.

Cherrstrom, Catherine A., Ra'sheedah Richardson, Debra Fowler, Robin Autenrieth, and Mark J. Zoran. 2017. "Creating teaching opportunities for STEM future faculty development." *Journal of STEM Teacher Education* 52 (1):art. 7.

Cockrell, Casey N., and Kyna Shelley. 2011. "The relationship between academic support systems and intended persistence in doctoral education." *Journal of College Student Retention* 12 (4):469-484.

Conley, John P., and Ali Sina Önder. 2014. "The research productivity of new PhDs in economics: the surprisingly high non-success of the successful." *Journal of Economic Perspectives* 28 (3):205-215.

Conti, Annamaria, and Fabiana Visentin. 2015. Science and engineering Ph.D. students' career outcomes, by gender. *PLOS One* 10 (8). Accessed 10 August 2017. doi:https://doi.org/10.1371/journal.pone.0133177.

Council of Graduate Schools. 2008. "Promising Practices: Student Selection and Admission." accessed 6 September. http://www.phdcompletion.org/promising/selection.asp.

Cramer, Jason M., and Paul T. Hamilton. 2017. "An internship may not be enough: enhancing bioscience industry job readiness through practicum experiences." *Journal of Microbiology & Biology Education* 18 (1):1-5.

Crede, Erin, and Maura Borrego. 2012. "Learning in graduate engineering research groups of various sizes." *Journal of Engineering Education* 101 (3):565-589.

Dasgupta, Shoumita, Karen Symes, and Linda Hyman. 2015. "Leading change: curriculum reform in graduate education in the biomedical sciences." *Biochemistry and Molecular Biology Education* 43 (2):126-132. doi: 10.1002/bmb.20862.

DeChenne, Sue Ellen, Larry G. Enochs, and Mark Needham. 2012. "Science, technology, engineering, and mathematics graduate teaching assistants teaching self-efficacy." *Journal of the Scholarship of Teaching and Learning* 12 (4):102-123.

DeChenne, Sue Ellen, Natalie Koziol, Mark Needham, and Larry Enochs. 2015. "Modeling sources of teaching self-efficacy for science, technology, engineering, and mathematics graduate teaching assistants." *CBE-Life Sciences Education* 14 (3):1-14. doi: 10.1187/cbe.14-09-0153.

DeChenne, Sue Ellen, Kristin Lesseig, Shawn M. Anderson, Sissi L. Li, Nancy L. Staus, and Celeste Barthel. 2012. "Toward a measure of professional development for graduate student teaching assistants." *Journal of Effective Teaching* 12 (1):4-19.

DeFraine, William C., Wendy M. Williams, and Stephen J. Ceci. 2014. Attracting STEM talent: Do STEM students prefer traditional or work/life-interaction labs? *PLOS One* 9 (2). Accessed 10 August 2017. doi:doi:10.1371/journal.pone.0089801.

Ebert-May, Diane, Terry L. Derting, Timothy P. Henkel, Jessica Middlemis Maher, Jennifer L. Momsen, Bryan Arnold, and Heather A. Passmore. 2015. "Breaking the cycle: future faculty begin teaching with learner-centered strategies after professional development." *CBE-Life Sciences Education* 14 (2):ar22, 1-12. doi: 10.1187/cbe.14-12-0222.

Ehrenberg, Ronald G., George H. Jakubson, Jeffrey A. Groen, Eric So, and Joseph Price. 2007. "Inside the black box of doctoral education: what program chacteristics influence doctoral students' attrition and graduation probabilities." *Educational Evaluation and Policy Analysis* 29 (2):134-150.

Ehrenberg, Ronald G., and Panagiotis G. Mavros. 1995. "Do doctoral students' financial support patterns affect their times-to-degree and completion probabilities." *Journal of Human Resources* 30 (3):581-609.

Fagen, Adam P., and Kimberly M. Suedkamp Wells. 2004. "The 2000 National Doctoral Program Survey: an on-line study of students' voices." In *Paths to the Professoriate: Strategies for Enriching the Preparation of Future Faculty*, edited by Donald H. Wulff and Ann E. Austin, 74-91. San Francisco, CA: Jossey-Bass.

Feldman, Mitchell D., Jody E. Steinauer, Mandana Khalili, Laurence Huang, James S. Kahn, Kathryn A. Lee, Jennifer Creasman, and Jeanette S. Brown. 2012. "A mentor development program for clinical translational science faculty leads to sustained, improved confidence in mentoring skills." *Clinical and Translational Science* 5 (4):362-367.

Feldon, David F., Soojeong Jeong, James Peugh, Josipa Roksa, Cathy Maahs-Fladung, Alok Shenoy, and Michael Oliva. 2017. "Null effects of boot camps and short-format training for PhD students in life sciences." *Proceedings of the National Academy of Sciences* 114 (37):9854-9858.

Feldon, David F., Michelle A. Maher, Melissa Hurst, and Briana Timmerman. 2015. "Faculty mentors', graduate students', and performance-based assessments of students' research skill development." *American Educational Research Journal* 52 (2):334-370. doi: 10.3102/0002831214549449.

Feldon, David F., Michelle A. Maher, Josipa Roksa, and James Peugh. 2016. "Cumulative advantage in the skill development of STEM graduate students: a mixed-methods study." *American Educational Research Journal* 53 (1):132-161. doi: 10.3102/0002831215619942.

Feldon, David F., James Peugh, Briana E. Timmerman, Michelle A. Maher, Melissa Hurst, Denise Strickland, Joanna A. Gilmore, and Cindy Stiegelmeyer. 2011. "Graduate students' teaching experiences improve their methodological research skills." *Science* 333 (6045):1037-1039. doi: 10.1126/science.1204109.

Fleming, Michael, Stephanie C. House, Vansa Shewakramani, Lan Yu, Jane Garbutt, Richard McGee, Kurt Kroenke, Zainab Abedin, and Doris M. Rubio. 2013. "The Mentoring Competency Assessment: validation of a new instrument to evaluate skills of research mentors." *Academic Medicine* 88 (7):1002-1008.

Fox, Mary Frank, and Paula E. Stephan. 2001. "Careers of young scientists: preferences, prospects and realities by gender." *Social Studies of Science* 31 (1):109-122.

Froyd, Jeffrey E., Maura Borrego, Stephanie Cutler, Charles Henderson, and Michael J. Prince. 2013. "Estimates of use of research-based instructional strategies in core electrical or computer engineering courses." *IEEE Transactions on Education* 56 (4):393-399.

Fuchs, Jonathan, Aminta Kouyate, Liz Kroboth, and Willi McFarland. 2016. "Growing the pipeline of diverse HIV investigators: the impact of mentored research experiences to engage underrepresented minority students." *AIDS and Behavior* 20:S249-S257.

Fuhrmann, C. N., D. G. Halme, P. S. O'Sullivan, and B. Lindstaedt. 2011. "Improving graduate education to support a branching career pipeline: recommendations based on a survey of doctoral students in the basic biomedical sciences." *CBE-Life Sciences Education* 10 (3):239-249. doi: 10.1187/cbe.11-02-0013.

Gamse, Beth, W. Carter Smith, Amanda Parsad, John Dreier, Kristen Neishi, Jennifer Carney, Linda Caswell, Eddie Breaux, Tom McCall, and Jon Spader. 2010. Evaluation of the National Science Foundation's GK-12 Program. Cambridge, MA: Abt Associates.

Gandhi, Monica, and Mallory Johnson. 2016. "Creating more effective mentors: mentoring the mentor." *AIDS and Behavior* 20:S294-S303. doi: 10.1007/s10461-016-1364-3.

Gardner, Grant E., and M. Gail Jones. 2011. "Pedagogical preparation of the science graduate teaching assistant: challenges and implications." *Science Educator* 20 (2):31-41.

Gardner, Susan K. 2009a. "Conceptualizing success in doctoral education: perspectives of faculty in seven disciplines." *Review of Higher Education* 32 (3):383-406.

Gardner, Susan K. 2009b. *The Development of Doctoral Students: Phases of Challenge and Support*, *ASHE Higher Education Report*. San Francisco, CA: Jossey-Bass.

Gibbs, Kenneth D., Jr., John McGready, Jessica C. Bennett, and Kimberly Griffin. 2014. "Biomedical science Ph.D. career interest patterns by race/ethnicity and gender." *PLoS ONE* 9 (12). doi: 10.1371/journal.pone.0114736.

Gilmore, Joanna, Michelle Vieyra, Briana Timmerman, David F. Feldon, and Michelle Maher. 2015. "The relationship between undergraduate research participation and subsequent research performance of early career STEM graduate students." *Journal of Higher Education* 86 (6):834-863.

Godskesen, Mirjam, and Sofie Kobayashi. 2016. "Coaching doctoral students - a means to enhance progress and support self-organisation in doctoral education." *Studies in Continuing Education* 38 (2):145-161. doi: 10.1080/0158037x.2015.1055464.

Golde, Chris M., and Timothy M. Dore. 2004. "The survey of doctoral education and career preparation: the importance of disciplinary contexts." In *Paths to the Professoriate: Strategies for Enriching the Preparation of Future Faculty*, edited by Donald H. Wulff and Ann E. Austin. San Francisco, CA: Jossey-Bass.

Greer, Dominique A., Abby Cathcart, and Larry Neale. 2016. "Helping doctoral students teach: transitioning to early career academia through cognitive apprenticeship." *Higher Education Research & Development* 35 (4):712-726. doi: 10.1080/07294360.2015.1137873.

Groen, Jeffrey A., George H. Jakubson, Ronald G. Ehrenberg, Scott Condie, and Albert Y. Liu. 2008. "Program design and student outcomes in graduate education." *Economics of Education Review* 27 (2):111-124.

Grohnert, Therese, Simon Beausaert, and Mien Segers. 2014. "Pirfalls of personal development plans -- the user perspective." *Journal of Vocational Education and Training* 66 (1):74-88.

Gutlerner, Johanna L., and David Van Vactor. 2013. "Catalyzing curriculum evolution in graduate science education." *Cell* 153:731-736.

Hall, Alison, Jessica Mann, and Michael Bender. 2015. "Analysis of Scholar Outcomes for the NIGMS Postbaccalaureate Research Education Program." In. Bethesda, MD: National Institutes of Health. https://www.nigms.nih.gov/News/reports/Documents/PREP-outcomes-report.pdf.

Hall, Joshua D., Jessica R. Harrell, Kimberley W. Cohen, Virginia L. Miller, Patricia V. Phelps, and Jeanette G. Cook. 2016. "Preparing postbaccalaureates for entry and success in biomedical PhD programs." *CBE-Life Sciences Education* 15 (3):ar27, 1-13. doi: 10.1187/cbe.16-01-0054.

Hall, Joshua D., Anna B. O'Connell, and Jeanette G. Cook. 2017. "Predictors of student productivity in biomedical graduate school applications." *PLoS ONE* 12 (1). doi: 10.1371/journal.pone.0169121.

Hall, Troy E., Jesse Engebretson, Michael O'Rourke, Zach Piso, Kyle Whyte, and Sean Valles. 2017. "The need for social ethics in interdisciplinary environmental science graduate programs: results from a nation-wide survey in the United States." *Science and Engineering Ethics* 23 (2):565-588.

Hamilton, Paul T., Sarah C. Luginbuhl, and Michael Hyman. 2012. "Preparing science-trained professionals for the biotechnology industry: a ten-year perpsective on a Professional Science Master's program." *Journal of Microbiology & Biology Education* 13 (1):39-44.

Harman, Kay. 2002. "The research training experiences of doctoral students linked to Australian Cooperative Research Centres." *Higher Education* 44 (3/4):469-492.

Henderson, Charles, Andrea Beach, and Noah Finkelstein. 2011. "Facilitating change in undergraduate STEM instructional practices: an analytic review of the literature." *Journal of Research in Science Teaching* 48 (8):952-984.

Hobin, Jennifer A., Philip S. Clifford, Ben M. Dunn, Susan Rich, and Louis B. Justement. 2014. "Putting PhDs to Work: Career Planning for Today's Scientist." *CBE-Life Sciences Education* 13 (1):49-53. doi: 10.1187/cbe-13-04-0085.

Holtzclaw, J. D., Lee G. Morris, Robert Pyatt, Cynthia S. Giver, Joseph Hoey, J. K. Haynes, Robert B. Gunn, Douglas Eaton, and Arri Eisen. 2005. "FIRST: A model for developing new science faculty." *Journal of College Science Teaching* 34 (6):24-29.

Hue, Gillian, Jessica Sales, Dawn Comeau, David G. Lynn, and Arri Eisen. 2010. "The American science pipeline: sustaining innovation in a time of economic crisis." *CBE-Life Sciences Education* 9 (4):431-434.

Huitt, Tiffany W., Anita Killins, and William S. Brooks. 2015. "Team-based learning in the gross anatomy laboratory improves academic performance and students' attitudes toward teamwork." *Anatomical Sciences Education* 8 (2):95-103.

Jamieson, Leah H., and Jack R. Lohman. 2012. "Innovation with Impact: Creating a Culture for Scholarly and Systematic Innovation in Engineering Education." In. Washington, DC: American Society for Engineering Education. https://www.asee.org/member-resources/reports/Innovation-with-Impact (accessed 21 October 2017).

Jensen, Jamie L., Tyler A. Kummer, and Patricia Godoy. 2015. "Improvements from a flipped classroom may simply be the fruits of active learning." *CBE-Life Sciences Education* 14 (1):ar5.

Johnson, James F., Zhanna Bagdasarov, Shane Connelly, Lauren Harkrider, Lynn D. Devenport, Michael D. Mumford, and Chase E. Thiel. 2012. "Case-based ethics education: the impact of cause complexity and outcome favorability on ethicality." *Journal of Empirical Research on Human Research Ethics* 7 (3):63-77.

Kamler, Barbara. 2008. "Rethinking doctoral publication practices: writing from and beyond the thesis." *Studies in Higher Education* 33 (3):283-294.

Kamler, Barbara, and Pat Thomson. 2006. *Helping Doctoral Students Write: Pedagogies for Supervision*. London: Routledge.

Keen-Rhinehart, E., A. Eisen, D. Eaton, and K. McCormack. 2009. "Interactive methods for teaching action potentials, an example of teaching innovation from neuroscience postdoctoral fellows in the Fellowships in Research and Science Teaching (FIRST) program." *Journal of Undergraduate Neuroscience Education* 7 (2):A74-A79.

Khanova, Julia, Mary T. Roth, Jo Ellen Rodgers, and Jacqueline E. McLaughlin. 2015. "Student experiences across multiple flipped courses in a single curriculum." *Medical Education* 49 (10):1038-1048.

Kim, Dongbin, and Cindy Otts. 2010. "The effect of loans on time to doctorate degree: differences by race/ethnicity, field of study, and institutional characteristics." *Journal of Higher Education* 81 (1):1-32.

Kim, Young K., and Linda J. Sax. 2009. "Student-faculty interaction in research universities: differences by student gender, race, social class, and first-generation status." *Research in Higher Education* 50 (5):437-459.

Komura, Kiriko. 2014. Outcomes for PSM Alumni: 2013/14. Claremont, CA: PSM National Office, Keck Graduate Institute.

Krupat, Edward, Jeremy Richards, Amy Sullivan, Thomas J. Fleenor, and Richard M. Schwartzstein. 2016. "Assessing the effectiveness of case-based collaborative learning via randomized controlled trial." *Academic Medicine* 91 (5):723-729.

Kuncel, Nathan R., Serena Wee, Lauren Serafin, and Sarah A. Hezlett. 2010. "The validity of the Graduate Record Examination for master's and doctoral programs: a meta-analytic investigation." *Educational and Psychological Measurement* 70 (2):340-352.

Kurdziel, Josepha P., Jessica A. Turner, Julie A. Luft, and Gillian H. Roehrig. 2003. "Graduate teaching assistants and inquiry-based instruction: implications for graduate teaching assistant training." *Journal of Chemical Education* 80 (10):1206. doi: 10.1021/ed080p1206.

Latham, Gary P., and Edwin A. Locke. 2007. "New developments in and directions for goal-setting research." *European Psychologist* 12 (4):290-300.

Lattuca, Lisa R., Inger Bergom, and David B. Knight. 2014. "Professional development, departmental contexts, and use of instructional strategies." *Journal of Engineering Education* 103 (4):549-572.

Lewis, Vivian, Camille A. Martina, Michael P. McDermott, Paula M. Trief, Steven R. Goodman, Gene D. Morse, Jennifer G. LaGuardia, Daryl Sharp, and Richard M. Ryan. 2016. "A randomized controlled trial of mentoring interventions for underrepresented minorities." *Academic Medicine* 91 (7):994-1001. doi: 10.1097/acm.0000000000001056.

Linn, Marcia C., Erin Palmer, Anne Baranger, Elizabeth Gerard, and Elisa Stone. 2015. "Undergraduate research experiences: impacts and opportunities." *Science* 347 (6222):e1261757.

Louis, Karen Seashore, Janet M. Holdsworth, Melissa S. Anderson, and Eric G. Campbell. 2007. "Becoming a scientist: the effects of work-group size and organizational climate." *Journal of Higher Education* 78 (3):311-336.

Love Stowell, Sierra M., Amber C. Churchill, Amanda K. Hund, Katharine C. Kelsey, Miranda D. Redmond, Sarah A. Seiter, and Nichole N. Barger. 2015. "Transforming graduate training in STEM education." *Bulletin of the Ecological Society of America* 96 (2):317-323.

Lovitts, Barbara E. 2001. *Leaving the Ivory Tower: The Causes and Consequences of Departure from Doctoral Study*. Lanham, MD: Rowman & Littlefield.

Lovitts, Barbara E. 2004. "Research on the structure and process of graduate education: retaining students." In *Paths to the Professoriate: Strategies for Enriching the Prepartion of Future Faculty*, edited by Donald H. Wulff and Ann E. Austin, 115-136. San Francisco, CA: Jossey-Bass.

Lovitts, Barbara E. 2008. "The transition to independent research: who makes it, who doesn't, and why." *Journal of Higher Education* 79 (3):296-325.

Maher, Michelle A., Briana Crotwell Timmerman, David F. Feldon, and Denise Strickland. 2013. "Factors affecting the occurrence of faculty-doctoral student coauthorship." *Journal of Higher Education* 84 (1):121-143.

Manathunga, Catherine, and Paul Lant. 2006. "How do we ensure good PhD student outcomes?" *Education for Chemical Engineers* 1 (1):72-81.

Mangematin, V. 2000. "PhD job market: professional trajectories and incentives during the PhD." *Research Policy* 20:741-756.

Mantai, Lilia. 2015. Feeling like a researcher: experiences of early doctoral students in Australia. *Studies in Higher Education*. doi:10.1080/03075079.2015.1067603.

Mason, Mary Ann, Marc Goulden, and Karie Frasch. 2009. "Why graduate students reject the fast track." *Academe* 95 (1):11-16.

McCormack, Wayne T., and Cynthia W. Garvan. 2014. "Team-based learning instruction for responsible conduct of research positively impacts ethical decision-making." *Accountability in Research* 21:34-49.

McGaskey, Ferlin G. 2015. "Facilitating the creation of knowledge: an investigation of the factors that influence the research productivity of Black doctoral students at predominantly White institutions." *Journal of Negro Education* 84 (2):187-201.

Means, Barbara, Yukie Toyama, Robert Murphy, Marianne Bakia, and Karla Jones. 2009. "Evaluation of evidence-based practices in online learning: a meta-analysis and review of online learning studies." In. Washington, DC: U.S. Department of Education. http://files.eric.ed.gov/fulltext/ED505824.pdf (accessed 22 October 2017).

Miller, Jennifer M., and Maryann P. Feldman. 2015. "Isolated in the lab: examining dissatisfaction with postdoctoral appointments." *Journal of Higher Education* 86 (5):697-724.

Miller, Kristen, Peggy Brickman, and J. Steve Oliver. 2014. "Enhancing teaching assistants' (TAs') inquiry teaching by means of teaching observations and reflective discourse." *School Science and Mathematics* 114 (4):178-190.

Mollica, Michelle, and Lynne Nemeth. 2014. "Outcomes and characteristics of faculty/student mentorship in PhD programs." *American Journal of Educational Research* 2 (9):703-708.

Moneta-Koehler, Liane, Abigail M. Brown, Kimberly A. Petrie, Brent J. Evans, and Roger Chalkley. 2017. "The limitations of the GRE in predicting success in biomedical graduate school." *PLoS ONE* 12 (1):e0166742. doi: 10.1371/journal.pone.0166742.

Morrison, Emory, Elizabeth Rudd, William Zumeta, and Maresi Nerad. 2011. "What matters for excellence in PhD programs? Latent constructs of doctoral program quality used by early career social scientists." *Journal of Higher Education* 82 (5):535-+.

National Academies of Sciences, Engineering, and Medicine. 2017. Undergraduate Research Experiences for STEM Students: Successes, Challenges, and Opportunities. Washington, DC: The National Academies Press.

Nerad, Maresi, Rebecca Aanerud, and Joseph Cerny. 2004. "'So you want to become a professor!': lessons from the PhDs--Ten Years Later study." In *Paths to the Professoriate: Strategies for Enriching the Preparation of Future Faculty*, edited by Donald H. Wulff and Ann E. Austin, 137-158. San Francisco, CA: Jossey-Bass.

Nowlis, Vincent, Kenneth E. Clark, and Miriam Rock. 1968. *The Graduate Student as Teacher*. Washington, DC: American Council on Education.

Noy, Shiri, and Rashawn Ray. 2012. "Graduate students' perceptions of their advisors: Is there systematic disadvantage in mentorship?" *Journal of Higher Education* 83 (6):876-914.

NRC. 2008. Science Professionals: Master's Education for a Competitive World. Washington, DC: Committee on Enhancing the Master's Degree in the Natural Sciences, Natural Research Council.

Nurrenbern, Susan C., Joseph A. Mickiewicz, and Joseph S. Francisco. 1999. "The impact of continuous instructional development on graduate and undergraduate students." *Journal of Chemical Education* 76 (1):114-119.

O'Connor, Erin E., Jessica Fried, Nancy McNulty, Pallav Shah, Jeffery P. Hogg, Petra Lewis, Thomas Zeffiro, Vikas Agarwal, and Sravanthi Reddy. 2016. "Flipping radiology education right side up." *Academic Radiology* 23 (7):810-822.

O'Meara, KerryAnn, Katrina Knudsen, and Jill Jones. 2013. "The role of emotional competencies in faculty-doctoral student relationships." *Review of Higher Education* 36 (3):315-347.

Paglis, Laura L., Stephen G. Green, and Talya N. Bauer. 2006. "Does adviser mentoring add value? A longitudinal study of mentoring and doctoral student outcomes." *Research in Higher Education* 47 (4):451-476.

Palmer, Michael S. 2011. "Taxonomy of TA training programs." In *Mapping the Range of Graduate Student Professional Development*, edited by Laura L. B. Border, 1-17. Stillwater, OK: New Forums Press.

Pfund, Christine, Stephanie C. House, Pamela Asquith, Michael F. Fleming, Kevin A. Buhr, Ellen L. Burnham, Julie M. Eichenberger Gilmore, W. Charles Huskins, Richard McGee, Kathryn Schurr, Eugene D. Shapiro, Kimberly C. Spencer, and Christine A. Sorkness. 2014. "Training mentors of clinical and translational research scholars: a randomized controlled trial." *Academic Medicine* 89 (5):774-782.

Pruitt-Logan, Anne S., and Jerry G. Gaff. 2004. "Preparing future faculty: changing the culture of doctoral education." In *Paths to the Professoriate: Strategies for Enriching the Preparation of Future Faculty*, edited by Donald H. Wulff and Ann E. Austin, 177-193. San Francisco, CA: Jossey-Bass.

Pruitt-Logan, Anne S., Jerry G. Gaff, and Joyce E. Jentoft. 2002. Preparing Future Faculty in the Sciences and Mathematics: A Guide for Change. Washington DC.

Roach, Michael. 2017. "Encouraging entrepreneurship in university labs: research activities, research outputs, and early doctorate careers." *PLoS ONE* 12 (2). doi: 10.1371/journal.pone.0170444.

Roach, Michael, and Henry Sauermann. 2010. "A taste for science? PhD scientists' academic orientation and self-selection into research careers in industry." *Research Policy* 39 (3):422-434. doi: 10.1016/j.respol.2010.01.004.

Romsdahl, Rebecca J., and Michael J. Hill. 2012. "Applying the learning community model to graduate education: linking research and teaching between core courses." *Teaching in Higher Education* 17 (6):722-734. doi: 10.1080/13562517.2012.678325.

Rybarczyk, Brian, Leslie Lerea, P. Kay Lund, Dawayne Whittington, and Linda Dykstra. 2011. "Postdoctoral training aligned with the academic professoriate." *BioScience* 61 (9):699-705.

Rybarczyk, Brian, Leslie Lerea, Dawayne Whittington, and Linda Dykstra. 2016. "Analysis of postdoctoral training outcomes that broaden participation in science careers." *CBE-Life Sciences Education* 15 (3):ar33, 1-11. doi: 10.1187/cbe.16-01-0032.

Ryker, Katherine, and David McConnell. 2014. "Can graduate teaching assistants teach inquiry-based geology labs effectively?" *Journal of College Science Teaching* 44 (1):56-63.

Sales, Jessica, Dawn Comeau, Kathleen Liddle, Nikki Khanna, Lisa Perrone, Katrina Palmer, and David Lynn. 2006. "Bridging the gap: a research-based approach for teaching interdisciplinary science to undergraduate freshman students." *Journal of College Science Teaching* 35 (6):36-41.

Sales, Jessica, Dawn Comeau, Kathleen Liddle, Lisa Perrone, Katrina Palmer, and David Lynn. 2007. "Preparing future faculty: an interdisciplinary, undergraduate science course taught by graduate and postdoctoral teacher-scholars." *Journal of College Science Teaching* 36 (4):24-30.

Sauermann, Henry, and Michael Roach. 2012. "Science PhD career preferences: levels, changes, and advisor encouragement." *PLoS ONE* 7 (5). doi: 10.1371/journal.pone.0036307.

Schussler, Elizabeth E., Quentin Read, Gili Marbach-Ad, Kristen Miller, and Miriam Ferzli. 2015. "Preparing biology graduate teaching assistants for their roles as instructors: an assessment of institutional approaches." *CBE-Life Sciences Education* 14 (3). doi: 10.1187/cbe-14-11-0196.

Shauman, Kimberlee A. 2017. Gender differences in the early employment outcomes of STEM doctorates. *Social Sciences* 6 (24). doi:doi:10.3390/socsci6010024.

Sisk, Rebecca. 2011. "Team-based learning: systematic research review." *Journal of Nursing Education* 50 (12):665-669.

Soldner, Matthew, Heather Rowan-Kenyon, Karen Kurotsuchi Inkelas, Jason Garvey, and Claire Robbins. 2012. "Supporting students' intentions to persist in STEM disciplines: the role of living-learning programs among other social-cognitive factors." *Journal of Higher Education* 83 (3):311-336.

Stephan, Paula. 2012. "Perverse incentives." *Nature* 484 (7392):29-31.

Sword, Helen. 2014. "Dancing on the bottom line: an unruly cost-benefit analysis of three academic development initiatives." *Higher Education Research & Development* 33 (4):783-793.

Szelényi, Katalin, Kate Bresonis, and Matthew M. Mars. 2016. "Who am I versus who can I become?: exploring women's science identities in STEM Ph.D. programs." *Review of Higher Education* 40 (1):1-31.

Thiel, Chase E., Shane Connelly, Lauren Harkrider, Lynn D. Devenport, Zhanna Bagdasarov, James F. Johnson, and Michael D. Mumford. 2013. "Case-based knowledge and ethics education: improving learning and transfer through emotionally rich cases." *Science and Engineering Ethics* 19 (1):265-286.

Thune, T. 2009. "Doctoral students on the university-industry interface: a review of the literature." *Higher Education* 58 (5):637-651. doi: 10.1007/s10734-009-9214-0.

Trautmann, Nancy M., and Marianne E. Krasny. 2006. "Integrating teaching and research: a new model for graduate education?" *BioScience* 56 (2):159-165. doi: 10.1641/0006-3568(2006)056[0159:itaran]2.0.co;2.

Tune, Johnathan D., Michael Sturek, and David P. Basile. 2013. "Flipped classroom model improves graduate student performance in cardiovascular, respiratory, and renal physiology." *Advances in Physiology Education* 37 (4):316-320. doi: 10.1152/advan.00091.2013.

Vanstone, Meredith, Kathy Hibbert, Elizabeth Anne Kinsella, Pam McKenzie, and Allan Pitman. 2013. "Interdisciplinary doctoral research supervision: a scoping review." *Canadian Journal of Higher Education* 43 (2):42-67.

von Hoene, Linda. 2011. "Graduate student teaching certificates: survey of current programs." In *Mapping the Range of Graduate Student Professional Development*, edited by Laura L. B. Border, 101-123. Stillwater, OK: New Forums Press.

Weidman, John C., Darla J. Twale, and Elizabeth Leahy Stein. 2001. Socialization of graduate and professional students in higher education: a perilous passage? In *ASHE-ERIC Higher Education Report*. San Francisco, CA: Jossey-Bass.

Williams, Simon N., Bhoomi K. Thakore, and Richard McGee. 2016. "Coaching to augment mentoring to achieve faculty diversity: a randomized controlled trial." *Academic Medicine* 91 (8):1128-1135. doi: 10.1097/acm.0000000000001026.

Winkle-Wagner, Rachelle, and Dorian L. McCoy. 2016. "Entering the (postgraduate) field: underrepresented students' acquisition of cultural and social capital in graduate school preparation programs." *Journal of Higher Education* 87 (2):178-205.

Wulff, Donald H., Ann E. Austin, Jody D. Nyquist, and Jo Sprague. 2004. "The development of graduate students as teaching scholars." In *Paths to the Professoriate: Strategies for Enriching the Preparation of Future Faculty*, edited by Donald H. Wulff and Ann E. Austin, 46-73. San Francisco, CA: Jossey-Bass.

Wyse, Sara A., Tammy M. Long, and Diane Ebert-May. 2014. "Teaching assistant professional development in biology: design for and driven by multidimensional data." *CBE-Life Sciences Education* 13 (2):212-223.

Zalewski, Dan, and Kellie Schneider. 2016. "Assessing learning outcomes and evaluating graduate student perceptions of a flipped classroom." American Society for Engineering Education 123rd Annual Meeting, New Orleans, LA.

Zhao, Chun-Mei, Chris M. Golde, and Alexander C. McCormick. 2007. "More than a signature: how advisor choice and advisor behaviour affect doctoral student satisfaction." *Journal of Further and Higher Education* 31 (3):263-281.