Faculty Development in the Health Professions: A Review of the Literature

By Linda S. Behar-Horenstein, Huan Kuang, Lissette A. Tolentino & Huibin Zhang

Faculty development programs are typically designed to broaden college/university faculty teaching expertise, increase research knowledge, or augment mentoring practice. Other than improved knowledge, skills and confidence levels, long-term benefits remain relatively unknown. The authors reviewed 414 empirical health professions faculty development studies published between 2006-2018; 97 that met inclusion criteria are reported. First, the authors read each paper to ensure a consistent understanding of its goals. Second, the authors provide a descriptive analysis of the categories, methods, and results. Third, the authors provide an overview of other characteristics such as the participants studied, researcher recommendations, elements of analysis, and critique the methods used. The findings showed that there were an overwhelming number of positive outcomes, which suggests that faculty development is beneficial. To improve data collection and increase the use of robust methods, the authors recommend strategies related to measurement types, data sources, research design, and researcher positionality. Improving study procedures may shed light on the creditable benefits of faculty development initiatives and lend support for continuous resource commitment.

Faculty development is an approach to providing “the continuous learning that professionals may need … to [develop] maintain, enhance, and broaden their professional competence” (Gottlieb, Rogers & Rainey, 2002, p. 280). Faculty development programs are intended to augment college/university faculty teaching expertise, research knowledge, or mentoring practice (O’Sullivan & Irby, 2011; Steinert et al., 2006). Often, faculty development initiatives aim to empower faculty members to excel in their role as educators and, in so doing, to create organizations that encourage and reward continual learning (Irby, 1993).

Increasingly, faculty development has become an important component in professional education (Behar-Horenstein, Schneider-Mitchell, & Graff, 2008; Behar-Horenstein, Mitchell, & Graff, 2009; Behar-Horenstein, Ginn, & Garvan, 2016; O’Sullivan & Irby 2011; Steinert et al. 2006). Used to manage change and develop professional skills, one of the primary purposes of faculty development has been to improve instructional practice. From novice to veteran professor in academia, the practice of educating students in the health professions requires lifelong learning. Faculty development programs range from single session workshops to multi-year-long commitments.

The need for faculty development is underscored by the observation that professional schools rarely prepare prospective academic faculty for the responsibilities of college and university teaching. Although health professions instructors are typically quite skillful in the practice of their respective disciplines, often, they lack formal teacher training and relevant experience to provide mentoring or to conduct research studies. Some faculty who have little to no experience may figure out how to teach or acquire the mentoring or research skills on their own, while others prefer to learn alongside their colleagues (Starcke, 2005).

The landscape of teaching and learning has changed dramatically in the last two decades. Without pedagogical training, and if left to discover on their own what is expected of them once they enter the academy, faculty have varying degrees of success. At the same time, universities and colleges recognize that faculty retention depends on the successful transition of individuals into the related roles and responsibilities of the professoriate (Irby, Lynch, Boswell, & Kappler Hewitt, 2017). Creating a collegial community of learners that supports risk-taking, reflection and sharing victories relies on the intentional delivery of tailored faculty development initiatives which encourage professional
growth and facilitate faculty vitality and retention (Behar-Horenstein, Catalanotto, Garvan, Su, & Feng, 2015a; Behar-Horenstein et al., 2009; Behar-Horenstein et al., 2010).

Doctoral level programs in the health professions designed to prepare clinicians for professional careers, frequently do not train them for the instructional and research demands of college/university faculty positions. Meanwhile, professional school promotion requirements often require evidence of scholarly activities such as a record of continuous publications, garnering grant funding, and teaching proficiency. Online certificate master’s degree, and continuing education programs, aimed at fostering related knowledge and skills, proffer options for post-graduate study. Programs may cultivate requisite skill development in conducting research, acquiring instructional practices that promote active learning, or promoting mentor competencies. However, enrollment frequently is not mandated.

Previous research highlights how university/college faculty development programs enhance faculty proficiency (Wallace & Infante, 2010). Such initiatives have been linked to faculty’s decision to remain in academics. For example, in a study of 167 baccalaureate dental hygiene faculty, the majority of the participants (96%, n=107) reported that the advancement opportunities were a somewhat or very important factor in their decision to leave their current position and accept another position (Collins, Zinski, Kesluka, & Thompson, 2007). Similarly owing to a lack of professional growth opportunities, 40% (n=35) of participants reported that they were somewhat or very dissatisfied with the amount of time allowed for keeping abreast of new and emergent knowledge in the field (Collins et al.). Feeling pressured to publish was also a somewhat or very important factor in the majority’s (85%, n=96) decision to leave their current position and accept another position. Disadvantages, attributable in part to a lack of faculty development initiatives, were highlighted by participants who reported lacking time to engage in professional growth activities and to dedicate to research/scholarship. Concurrently they acknowledged that productivity in these domains was a deciding factor in tenure decisions. When asked to rate the most important skills for future dental hygiene faculty, more than 90% of 592 participants study identified educational, technological and clinical skills (Coplen, Klausner, & Taichman, 2011). Faculty development programs have been reported to enhance participants’ sense of belongingness, socialization into the workplace environment and productivity. Researchers have shown the benefits of dental faculty development programs for faculty, students, and the community (Behar-Horenstein et al., 2008; Behar-Horenstein et al., 2009; Behar-Horenstein et al., 2010; Behar-Horenstein et al., 2016; Behar-Horenstein et al., 2015a Behar-Horenstein, Yu, & Beck, 2018; Hendricson et al., 2007; O’Neill, & Taylor, 2001).

Over the past three decades, studies pertaining to faculty development programs for dental, medical, and nursing faculty have increased (Hand, 2006; O’Neill, Taylor, 2001). Specifically, their impact on academic careers has been studied among faculty in medicine, pharmacy, nursing, dentistry, and dental hygiene (Behar-Horenstein et al., 2010; Boyce et al., 2009; Enkins, Harmer, Pardue, & Turcato, 2004; Lowenstein, 2011; Srinivasan et al., 2007; Steinert, 2014). Benefits of these initiatives have been realized from varied formats such as workshops, seminars, and courses. Previous research (O’Neill, Taylor, 2001; Hendricson, et al., 2007; Mahler & Benor, 1984; Skeff, Stratos, & Bergen, 1992; Mahler & Neumann, 1987; Stratos, Bergen, Albright, Skeff, & Owens, 1997), systematic reviews (Steinert, 2006), and commentaries (Licari, 2007) accentuate the importance of preparing healthcare professionals for teaching. Researchers have shown that participation in faculty development opportunities leads to enhanced instructional effectiveness, faculty recruitment and retention, academic excellence, and innovation. Engagement in faculty development has fostered positive change in instructional beliefs and practices (Behar-Horenstein et al., 2006; Behar-Horenstein et al., 2008; Behar-Horenstein et al., 2016; Joyce, Weil, & Calhoun, 2004; Steinert et al., 2006). Researchers have highlighted the need for continuous professional opportunities. They stressed the urgency of improving teaching skills, strengthening collegial relationships, and developing the overall academic advancement of healthcare faculty. (Licari, 2007; McAndrew, Motwaly, & Kamens, 2013).

Following immersion in faculty development programs, participants have reported an improved personal value for and an ability to engender a
more positive teaching culture. Wallace and Infante (2010) reported that faculty development focused on the reinforcement of clinical teaching practices and methods led to increased knowledge in competency-based teaching and improved communication skills with students. Others concur. For example, clinical teachers demonstrated improved teaching skills following participation in faculty development programs (Møystad, Lycke, Barkvoll, & Lauvås, 2014). In contrast, initiatives focused on the knowledge and skills necessary to conduct educational research has not received a lot of attention (Behar-Horenstein et al., 2018). Overall evidence that supports the sustainability of newly acquired knowledge, skills and practices is sparse.

Periodic assimilation and analysis of a body of research is recommended to guide practice and prospective research inquiries. To address this need, the authors of this study present a review of the literature and a synthesis of quantitative and qualitative studies by: (a) programmatic aims, (b) measures or assessments used to document program or participant impact, (c) term of measures, (e.g., immediate, pre- and/or post-test, medium-term post-test up to 5 months, and a longer duration of more than 5 months), (d) reported outcomes (e.g., positive, negative, neutral or mixed findings), (e) methodological weaknesses, and (f) recommendations for future studies.

Methods

Search Methods

The researchers conducted an extensive review of published articles between 2006 and 2018 to identify studies. Search results were limited to empirical studies in peer-reviewed journals. Published conference papers and dissertations were excluded. Although the search was not limited by country of practice, the review was limited to articles published in English. The databases searched included: PubMed, ERIC, MEDLINE, CINAHL, and Professional Development Collection in education. We chose these databases because they cover professional development in the health professions. Articles with substantial overlap were identified; duplicates were removed.

Keywords used in the search included: health science center, professional development programs, faculty development, staff development, program implementation, program evaluation, and program effectiveness, faculty development programs with outcome measures/assessments, pharmacy, nursing, medicine, veterinary medicine, dentistry, physical therapy and public health resulting in 414 studies. The first and second authors read the titles and abstracts identified by each of the searches. Studies were reviewed to determine if they met the following inclusion criteria: (1) presence of qualitative or quantitative data and the (2) presence of outcome scores, descriptive, or qualitative findings. Ninety-seven studies that met the inclusion criteria were passed onto the next stage. The full version of the paper was read by all authors independently to determine suitability for inclusion.

Table 1. Study Categories and Conceptual Definitions (n=97)

<table>
<thead>
<tr>
<th>Category</th>
<th>Conceptual Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mentoring studies (n=15)</td>
<td>Outcomes of training opportunities that sought to enhance faculty knowledge and skills, coaching students and colleagues in identifying research goals, and advising their career development.</td>
</tr>
<tr>
<td>Workshop-based teaching studies (n=57)</td>
<td>Outcomes of training opportunities that sought to enhance participant knowledge and skills in instruction.</td>
</tr>
<tr>
<td>Workshop-based other than teaching studies (n=24)</td>
<td>Outcomes of training opportunities that sought to enhance participant knowledge and skills in content areas (e.g., geriatrics, evidence-based findings, and interprofessional education).</td>
</tr>
<tr>
<td>Fellowship studies (n=11)</td>
<td>Outcomes or experiences among participants who received funded training opportunities.</td>
</tr>
<tr>
<td>Literature studies (n=10)</td>
<td>Reviews of published studies related to faculty promotion or research productivity, mentoring; pharmacy, family medicine; and medicine teaching and learning development programs, characteristics of training programs; or faculty developer competence.</td>
</tr>
</tbody>
</table>

We worked in teams of two to categorize articles. The first author initially categorized each article and then conferred with one of the other three authors to reach consensus. The categories included: Mentoring studies (n=15), Workshop-based teaching studies (n=37), Workshop-based other than teaching studies (n=24), Fellowship studies (n=11) and Literature studies (n=10). We created conceptual definitions to guide the organization and summary of related findings within each set (see Table 1). We quantified the number of qualitative and quantitative articles by methods and research
design for all categories except for Category 1, Mentoring Studies and Category 5, Literature Studies. Next, we described the study’s purpose, treatment, and recommendations using researchers’ descriptions, not necessarily quantification. For example, when reviewing quantitative mentoring studies, we asked ourselves what purpose describes this group of studies. For this set of studies, we identified the overarching purposes as assessing participant satisfaction, gains in mentor knowledge, and skills. We categorized the qualitative articles by asking ourselves the following questions:

1. What was the intent of this program or literature review?
2. Did the program aim to enhance instructional skills, individual knowledge of how to teach, or how to coach students?
3. Were funded experiences, aimed at career development, offered?
4. What was the aim of the literature review?

Second, we report findings related to studies in: (a) Mentoring, (b) Workshop-based Teaching, (c) Workshop-based other than teaching, (d) Fellowship, and (e) Literature. The findings for each category are organized by quantitative and where applicable qualitative studies.

For quantitative studies, the synthesis is a summary of program aims, measurements (published protocols or researcher-constructed instruments), research designs, and reported outcomes. For the qualitative studies, the synthesis is a summary of programmatic aims and assessments (interviews, focus groups, and reported outcomes).

**Category 1: Mentoring Studies**

Quantitative studies

Fifteen quantitative studies (nearly 16% of the reviewed publications) focused on mentoring. Of these, the majority focused on participant satisfaction and gains in mentor knowledge and skills (Dutta et al., 2011; Feldman et al., 2012; Fornari et al., 2018; Pfund et al., 2013; Pfund et al., 2006; Pfund et al., 2014). Other studies examined the impact of mentor training on retention and career development (Chen, Sandborg, Hudgens, Sanford, & Bachrach, 2016; Phitayakorn, Petrusa, & Hodin, 2016), scientific dissemination (Ogunyemi, Solnik, Alexander, Fong, & Azziz, 2010), and program effectiveness in directed, individual mentoring (Haines & Popovich 2014; Thorndyke, Gusic, & Milner, 2008). Studies included faculty participants in medicine, nursing, pharmacy, and clinical translational research.

Byington et al (2016) analyzed institutional data; all other studies used surveys. Eight studies used the post-test only design with non-random assignment (Chen et al., 2012; Jackevicius et al., 2014; Ogunyemi et al., 2010; Pfund et al., 2013; Pfund et al., 2006; Pfund et al., 2015; Phitayakorn et al., 2016). Six studies used a pre- and post-test design with non-random assignment (Brody et al., 2016; Dutta et al., 2011; Fornari et al., 2018; Haines, & Popovich, 2014; Pfund et al., 2014; Thorndyke et al., 2008). With the exception of Dutta et al. (2011), which reported that mentees’ baseline expectations were higher than those perceived achievements at 6 months or one-year follow-up, all other findings were positive and showed statistically significant gains or improvement.

---

Table 2. Frequency of the Faculty Development Program Studies by Discipline (n= 97)

<table>
<thead>
<tr>
<th>Discipline</th>
<th>Frequency (n%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dentistry</td>
<td>14 (14.4%)</td>
</tr>
<tr>
<td>Medicine</td>
<td>39 (40.2%)</td>
</tr>
<tr>
<td>Nursing*</td>
<td>9 (9.3%)</td>
</tr>
<tr>
<td>Pharmacy</td>
<td>6 (6.2%)</td>
</tr>
<tr>
<td>Other**</td>
<td>29 (29.9%)</td>
</tr>
<tr>
<td>Total</td>
<td>97 (100%)</td>
</tr>
</tbody>
</table>

* One article included both medicine and nursing faculty but was categorized only in medicine.
** Denotes studies related to practice-based skill geriatrics and mentoring, and content knowledge in humanism, interprofessional education, and research skills.

Results

In this section, first, we report the frequency of the faculty development program studies within disciplines (see Table 2). The majority (39.8%, n=39) occurred in medicine, followed by other 29.6% (n=29). Other refers to practice-based skills geriatrics and mentoring, and content knowledge in humanism, interprofessional education, and research skills. Studies in dentistry (14.3%, n= 14), nursing (9.2%, n= 9) and pharmacy (6.1%, n=6) constituted slightly less than one third of the studies reviewed.
Category 2: Workshop-Based Teaching Studies

Overall, there were thirty-seven studies, accounting for 38% of the reviewed publications, including twenty-five quantitative and twelve qualitative studies.

Quantitative studies

In the quantitative publications, researchers addressed the effectiveness of faculty development programs that aimed to improve teaching and learning in health-related disciplines (Al-Eraky, Donkers, Wajid, & Van Merrienboer, 2015; Branch et al., 2014; Gadbury-Amyot, Smith, Overman, & Bunce, 2015; Gates et al., 2013; Gjerde, Hla, Kokotalo, & Anderson, 2008; Hölzer, Freytag, & Sonntag, 2017; Hull, Arora, Stefanidis, & Sevda, 2015; Johansson, Skeff, & Stratos, 2012; Johnson, 2016; Kogan, Conforti, Yamazaki, Iobst, & Holmboe, 2017; McAndrew et al., 2013; McAndrew, Eidtson, Pierre, & Gillespie, 2011; McAndrew & Pierre, 2013; Roh, Kim, & Tangkawianich, 2016; Mookherjee, Monash, Wentworth, & Sharpe, 2014; Mosyad, Lycke, Barkvoll, & Lauvås, 2015; Paige, Arora, Fernandez, & Seymour, 2015). Researchers reported outcomes from studies of teaching content or discipline-related strategies, such as practicing humanism (Branch et al., 2014; Branch et al., 2017), addressing pedagogic principles of teacher knowledge (Gadbury-Aymot et al., 2014; Ebrahimi & Kojuri, 2012; McLeod, Brawer, Steinert, Chalk, & McLeod, 2008), associating student characteristics and educational expectations (McAndrew et al., 2013), assessing and evaluating student outcomes (Gjerde et al., 2008; Johnson, 2016; McAndrew et al., 2013; McAndrew & Pierre, 2013), giving feedback (McAndrew et al., 2013), and maximizing simulation instructional competency (Roh et al., 2016). They also described factors that impacted training program effectiveness (Gates et al., 2013; Hölzer et al., 2017; Kogan et al., 2017). Johnstone-Dodge, Bowen, Calley, and Peterson (2014) assessed the impact of faculty development designed to enhance online teaching. Researchers reported positive outcomes in all studies. First, participants were satisfied with the programs (Al-Eraky et al., 2015; Ebrahimi & Kojuri, 2012; Gadbury-Aymot et al., 2014; Gadbury-Amyot et al., 2015; Hölzer et al., 2017; Johansson et al., 2009; Johnson, 2016; McAndrew et al., 2013; McAndrew & Pierre, 2013, Mookherjee et al., 2014; Moystad et al., 2015; Roh et al., 2016). Second, participants scored significantly higher on the teaching practice effectiveness than control groups following their pre-test measures (Branch et al., 2014; Branch et al., 2017; Gjerde et al., 2008; Johansson et al., 2012; Johnson, 2016; McAndrew et al., 2013; McAndrew & Pierre, 2013; Roh et al., 2016). Participants reported changes in teaching habits (Gates et al., 2012; Hull, Arora, Stefanidis, & Sevda, 2015; Johansson, Skeff, & Stratos, 2009; Johansson, Skeff, & Stratos, 2012; McAndrew et al., 2011; Paige et al., 2015). Faculty development programs were shown to have exerted positive impact on participant knowledge and behaviors (Ebrahimi & Kojuri, 2012; Gadbury-Aymot et al., 2014; McLeod et al., 2008.). A statistically significant relationship between motivation for change and commitment locus was reported (Kogan et al., 2017). Johnstone-Dodge, Bowen, Calley and Peterson (2014) reported significant pre- and post-test increases among dental hygiene faculty participants in their ratings of the importance of four practices: activities promoting relevant, lifelong learning, faculty communication fostering a sense of community, encouraging students’ self-introduction and encouraging productive dialogue and respecting diverse opinions. Third, increases in responsibility, teaching hours, and scholarly and academic outcomes were observed (Gjerde et al., 2008; McAndrew et al., 2013). Fourth, program participation was associated with higher faculty retention (McAndrew et al., 2013). Likewise, participation in workshop training increased individual development and application of substantial learning objectives (Freytag et al., 2017). Participants reported greater confidence in teaching, demonstrated an increased willingness to discuss teaching issues with peers, and reported taking on active roles as educational leaders (Gjerde et al., 2008; McAndrew et al., 2013).

Qualitative studies

In the twelve qualitative studies, researchers assessed programs which were designed to enhance participants’ skills as educators. Some interventions sought to expand faculty implementation of instructional methodologies such as case-based learning...
(CBL) (Behar-Horenstein, Catalanotto, & Nascimento, 2015b). In other studies, researchers assessed the impact of efforts aimed at augmenting student engagement (Hartford, Nimmon, & Stenfors, 2017; Wallace, & Infante, 2008; Weber, Cable, & Wehbee-Janek, 2016). Researchers explored the benefit of faculty development programs focused on simulation and identified the main characteristics of effective training (Balmer & Richards, 2012; Greenwood, & Ewell, 2018; Krautscheid, Kaakin, & Warner, 2008). Others emphasized the development of critical thinking via implementation of specific instructional strategies (Behar-Horenstein, et al., 2009; Behar-Horenstein, et al., 2010; Behar-Horenstein, et al., 2008; Christofilos, DeMatteo, Penciner, 2014; Liben, Chin, Boudreau, Boillat, Steinert, 2012).

Category 3: Workshop-Based Other Than Teaching Studies

Twenty-four, or nearly 25% of the total articles, twenty-one quantitative and three qualitative studies, comprised the workshop-based other than teaching set of articles.

Quantitative studies

The aims of these studies included: assessing program outcomes (Abu-Rish Blakeney, et al., 2016; Baker et al., 2010; Beck, Wingard, Zuniga, Heffetz, & Gilbreath, 2008; Krichbaum, Kaas, Wyman, & Van Son, 2015; Love et al., 2016; Sheehan, & Brannan, 2013; Taylor & Berry, 2008; Williams, Schigelone, Fitzgerald, & Halter, 2008; Yanamadala, Criscione-Schreiber, Hawley, Hellin, & Shah, 2016); increasing administrative knowledge and skills (Sheehan, & Brannan, 2013); conducting medical interviews (Hatam et al., 2007); expanding health literacy and geriatric care (Google, Hackett, Owens, Ansello, & Mathews, 2016; Evans et al., 2014; Williams et al., 2008); fostering leadership development (Carney et al., 2015); facilitating online program participation (Chan et al., 2018); exploring the research productivity among faculty who were promoted (Littman, Sonne, & Smith, 2017); identifying factors associated with faculty retention (Ries et al., 2009); examining the status of an interprofessional education (IPE) program (Alexandraki et al., 2017); promoting teaching evidence-based medicine concepts (Laird, George, Sanford, & Coon, 2010); writing multiple choice questions (Abdulghani et al., 2017); identifying components that ought to be included in new faculty orientation program (Holyfield & Berry, 2008); and assessing the effectiveness of writing programs (Isenburg, Lee, & Oermann, 2017). Pharmacy, medical, and dental faculty as well as medical residents participated in these studies.

In two studies, researchers reported findings from descriptive analyses (Isenburg, et al., 2017; Littman et al., 2017). Findings from seven studies were based on survey post-test analyses (Abu-Rish Blakeney, et al., 2016; Abdulghani, et al., 2017; Alexandraki et al., 2015; Carney et al., 2015; Chan et al., 2018; Holyfield, & Berry, 2008; Ries et al., 2009); 12 others reported findings from pre-/post-test surveys (Baker, 2010; Beck et al., 2008; Google et al., 2016; Evans et al., 2014; Hatam et al., 2007; Krichbaum et al., 2015; Laird et al., 2010; Love et al., 2016; Sheehan, & Brannan, 2013; Taylor, & Berry, 2008; Williams et al., 2008; Yanamadala et al., 2016). All study outcomes including statistically significant change scores, or improvement in participant confidence, knowledge, or skills, were positive.

Qualitative studies

Of the three qualitative studies, one (Heinrich, Hurst, Leigh, Oberleitner, & Poirrier, 2009) focused on impact of workshops on faculty productivity, which showed positive findings. In another study (Wilson, McCormack, & Ives, 2008), researchers reported how action learning improved participant confidence in working with others, a willingness to engage in introspection, and better patient care. During the course of the study, participants became more attuned to the impact of the environment on the patient and family, attentive to the need for pain management during invasive procedures and recognized the importance of multidisciplinary communication. Noticeable increases were reported in their attention to procedures related to nursing handover, breastfeeding, parent rooming in and orientation to care and getting ready for discharge.

In the remaining study (Baker et al., 2018), while using grounded theory, researchers explored the actions, experiences, and perceptions of faculty developers in their facilitator role. Findings highlighted three novel processes: negotiating, constructing, and attuning, and suggested that fac-
ulty developer competence is best understood as a situated construct shaped by the interaction between the properties of the environment and identity of that individual.

**Category 4: Fellowship Studies**

There were eleven studies, including seven quantitative (two of which included qualitative analyses), and four qualitative fellowship studies, or 11% of total articles reviewed publications.

**Quantitative studies**

The purpose of quantitative studies was to evaluate funded training program opportunities that aimed to influence early career outcomes (Brown et al., 2008); assess educational outcomes for senior family medicine residents (Reamy, Williams, Wilson, Goodie, & Stephens, 2012); determine program effectiveness on administrative training program (Sheehan & Brannan, 2012); measure program effectiveness (Campbell et al., 2013); evaluate effectiveness of measurable Accreditation Council of Medical Graduate Education (ACGME) competencies; evaluate literature on two newly defined competencies (Lurie, Mooney, & Lyness (2009); and assess the outcomes of a program that sought to better prepare doctoral students (Koblinsky, Hrapczynski, and Clark 2015). In another study, the goal was to evaluate nursing faculty scholar program cohorts, where the aim was to improve the career outcomes of junior nursing faculty (Hickey et al., 2014). Participants in these studies included faculty of senior family medicine residents, osteopathic residency directors, professors, students in the sciences, and faculty in nursing.

Data were collected from interviews, questionnaires, self-assessment, self-reported, public online data, and meta-analysis. In one study, researchers used institutional data to compare outcomes (Brown et al., 2008). Others reported use of a questionnaire, but they did not indicate whether or not it was validated (Reamy et al., 2012) or if it was researcher-constructed questionnaire (Sheehan & Brannan, 2012). Campbell et al. (2013) used self-reported and publicly available online data, while Lurie et al. (2009) completed a meta-analysis.

The majority of program outcomes were positive. Researchers from one study reported that as a result of engagement in faculty development initiatives, participants received funding earlier in their careers and had an increase in published work (Brown et al., 2008). Another set of researchers noted that participants were hired for faculty positions either immediately or shortly after the program, that they developed interest in becoming faculty, and increased faculty-related skills (Reamy, et al., 2012). Participants in one program reported having higher knowledge and skills related to residency director administration (Sheehan & Brannan, 2012). Another set of researchers reported an increase in publications, citing that training received was correlated with publications, especially in higher impact journals. They also noted that participants developed new skills and sharpened current ones after participating in the training programs, and that half of its participants became eligible for tenure (Campbell et al., 2013). Meta-analysis results from Lurie et al. (2019) showed that no effective measures have been used to evaluate the competencies set out by the American Council for Graduate Medical Education (ACGME).

**Qualitative studies**

Using findings from a qualitative analysis of programs, researchers identified four strengths of a training initiative: development of academic community, cultivation of mentoring relationships, broadened knowledge regarding diverse employment opportunities, and preparation for future employment (Koblinsky, et al., 2015). Researchers reported that students rated professional development activities highest. Program alumni reported that they felt well trained for research. Nine of 10 respondents reported feeling satisfied with their career tract (Koblinsky et al., 2015). Hickey et al.’s (2014) results demonstrated that during the program, several participants published articles and presented their work at conferences and meetings. Overall, participants reported that the program helped them attain professional and personal goals (Hickey et al., 2014). Study themes resulting from the open-ended questionnaire suggested that mentoring was important and inclusive to teaching (Hickey et al., 2014). Mentoring was cited as a crucial to learning how to review articles and grant proposals; it was deemed instrumental in participant success (Hickey et al., 2014).

Researchers reported outcomes among faculty
fellowship recipients engaged in a longitudinal program (Robins, Murphy, & Zierler, 2016); perspectives of participants enrolled in a fellowship program (Lown, Newman, & Hatem, 2009); the impact of a fellowship program for international faculty development (Kim et al., 2017); and outcomes from a research supported program (Archibald et al., 2017).

In one study, researchers included a team building session (though they did not describe its time or duration) (Robins et al., 2016). Other researchers evaluated fellowship opportunities where the goal was to develop and enhance faculty skills among participants who met weekly for two-hour length seminars throughout the academic year (Lown et al., 2009). Researchers evaluated an eight consecutive week fellowship program where the goal was to understand the factors surrounding the program’s success (Kim et al., 2017). A group of researchers evaluated the impact of a faculty development program designed to enhance research productivity and improve education among funded participants (Archibald et al., 2017). Researchers used surveys (Robins et al., 2016), semi-structured interviews (Lown et al., 2009; Kim et al., 2017), and program evaluation (Archibald et al., 2017).

Results suggest that programs offered participants opportunities to network, disseminate ideas and acquire best practices (Robins et al., 2016). Researchers reported an increase in participant knowledge surrounding their specialty (Lown et al., 2009) and successful outcomes for international faculty (Kim et al., 2017). Findings from Archibald et al. (2017) showed that program participation improved the overall departmental cultures.

**Category 5: Literature Studies**

Overall, there were 10 study reviews of publications related to faculty promotion, research productivity, or mentoring; pharmacy, family medicine, and medicine teaching and learning development programs; characteristics of training programs; or faculty developer competence that represented 10% of the total articles.

Buddeberg-Fischer and Herta (2006) examined the types of structured mentoring programs for doctors and medical students, what short- and long-term goals these projects pursued, and searched for evidence about program effectiveness and efficiency. Overall, programs lacked concrete structure as well as short- and long-term evaluation. Buddeberg-Fischer and Herta (2006) described findings in terms of great interest and satisfaction but not program effectiveness or the efficiency. Guglielmo et al. (2011) searched for studies from major healthcare professions academic organizations regarding programs intended to enhance faculty development. Relatively few publications were specific to pharmacy and none adequately evaluated the impact of various faculty-development programs on associated outcomes. Leslie, Baker, Egan-Lee, Esdaile, and Reeves (2013) assessed the scope and quality of faculty development programs in medical education. Most programs consisted of a short or longitudinal series of seminars, intended to improve teaching effectiveness. A few programs focused on improving scholarship and leadership. Surveys were typically used to assess self-reported behavioral change, acquisition of knowledge and skills, and changes in attitudes and perceptions. Meagher, Taylor, Probsfield, and Fleming (2011) evaluated 10 studies related to research mentors in clinical translational research. Although most studies evaluated mentor satisfaction and general opinions on mentoring during a single time period, the authors found considerable heterogeneity among the methods used. Sambunjak, Straus, and Marušić, 2006 reviewed 42 articles describing 39 studies about mentorship and its relationship to career development. Of these, 34 relied on the use of self-report survey results with small samples. Although mentoring was perceived to influence personal development, career guidance, career choice, and research productivity, including publication and grant success, women reported having more difficulty finding mentors. Sorinola and Thistlethwaite (2013) reviewed 46 studies of faculty development in family medicine. Most studies focused on teaching skills and relied on self-report measures. Participants reported that they valued engaging in faculty development. The findings, which demonstrated learning and behavioral change, support its benefit. However, transformation in organizational practice and student learning were infrequent. Strang and Baia (2016) assessed publications representing 20 programs that studied resident teaching programs and one program that described faculty development. Most programs were delivered across one
year and offered instruction on teaching methodologies and assessment measures. All programs except one included an experiential component. Thirteen publications presented outcomes data on satisfaction and self-perceived improvement.

Steinert et al. (2006) analyzed 53 studies to determine the effects of faculty development on participant knowledge, skills, and attitudes. All studies focused on teaching improvement. Interventions were delivered using workshops, seminar series, short courses, and longitudinal programs. Participants reported high satisfaction with programs, positive changes in attitudes toward faculty development and teaching, increased knowledge of educational principles, and gains in teaching skills. Steinert et al. (2016) reviewed 111 studies published between 2002 and 2012. Overall participant satisfaction with faculty development programs was high. Participants reported increased confidence, enthusiasm, and awareness of effective educational practices, gains in knowledge and skills, and self-reported changes in teaching behaviors.

Following a review of 17 articles that met criteria, one study aimed to identify characteristics of a training program or faculty developer competence (Ratka, Zorek, & Meyer, 2017). They identified successful program characteristics (i.e., ensuring institutional support; identifying objectives and outcomes based on IPE competencies; focusing on consensus-building and group facilitation skills; ensuring flexibility and remaining responsive to institution- and participant-specific characteristics; incorporating assessment strategies; assessing faculty developer competence such as trainer knowledge, skills and identity) but pointed out that faculty developer competence is better understood as a situated construct. All of the literature review studies included faculty participants in medicine and pharmacy.

The measures in this set of studies included self-reported data and researcher-constructed instruments (Buddeberg-Fischer & Herta, 2006; Guglielmo et al., 2011; Meagher et al., 2011; Leslie et al., 2013; Ratka et al., 2017; Sambunjak et al., 2006; Sorinola & Thistlethwaite, 2013; Steinert et al., 2006; Steinert et al., 2016; Strang & Baia, 2016).

Methodological weaknesses included small sample sizes, a lack of control group or use of pre-testing without repeated measurements, and failure to consider the potential impact of personal traits on training program effectiveness. Other limitations were omitted response rates, lack of description about the statistical methods used, or basic background information critical to understanding the context of the intervention such as its duration or the discipline.

Table 3. Summary of Research Designs Across All Studies (N = 97)

<table>
<thead>
<tr>
<th>Research Design Categories</th>
<th>#/ Sampling*</th>
</tr>
</thead>
<tbody>
<tr>
<td>One-shot case study survey</td>
<td>25/NR</td>
</tr>
<tr>
<td>One group, Pre-/post-test survey</td>
<td>36/NR</td>
</tr>
<tr>
<td>Two groups, Post-test survey</td>
<td>3/ R (1), NR (2)</td>
</tr>
<tr>
<td>Descriptive</td>
<td>6/ NA</td>
</tr>
<tr>
<td>Mixed methods</td>
<td>1, NR</td>
</tr>
<tr>
<td>Meta-Analysis</td>
<td>1, NA</td>
</tr>
<tr>
<td>Qualitative</td>
<td>25, NA</td>
</tr>
</tbody>
</table>

* Sampling refers to: random (R) non-random NR, not applicable, NA.

Research Design

The majority of researchers conducted survey studies (25 one-shot case studies, 36 one group, Pre-/post-test, and three two group, post-test.) Of these 62 (63.9%) studies, all but one (a two group, post-test study) used non-random sampling (see Table 3). More than a quarter of the studies were qualitative (n= 25, 25.7%). The remaining 10 studies (10.4%) were descriptive, mixed method or meta-analysis. With the exception of the one survey study that used random sampling the other survey studies are considered pre-experimental (Campbell & Stanley, 1963).

The pre-experimental design is known for both its lack of control group and random assignment. As the least powerful among research designs, it is also the most vulnerable to threats to internal validity. For example, studies that lacked control or comparison groups might have suffered from poor measurement reliability. The use of convenience samples, poor sampling procedures, and low response rates as well as a lack of reported psychometric questionnaire information limit the generalizability of studies (Hickey, et al., 2014). Survey results based on self-reporting that were not correlated with observable behaviors or outcomes (Lown, et al., 2009) is another limitation of the pre-experimental design.
For the pre-experimental studies that used only one group, it is difficult to assert a causal relationship between treatment and participant change since researchers did not describe if and how threats to validity were properly addressed. Campbell and Stanley (1963) refer to the pre-experimental study as “worth doing where nothing better can be done” (p. 7). We encourage authors of pre-experimental studies to acknowledge the limitations of such studies. Confounding and extraneous variables such as history, maturation, and testing could have jeopardized the internal validity of these studies. Not surprisingly, the veracity of pre-experimental study-related conclusions and recommendations cannot be taken at face value. The random selection and assignment of participants in empirical research has implications for making claims that statistical differences are attributable to treatment.

**Methods**

The methods used in these studies ranged from descriptive findings to correlation measures such Wilcoxon rank-sum test, t-tests, Pearson’s r or with Kendall’s tau-b, and ANOVA. Infrequently, researchers used repeated measures analysis (regression or logistic regression analysis (Kogan et al., 2017), mixed-effects regression modelling or meta-analysis techniques. Effect sizes were not reported in any of the studies. Several researchers also assessed program outcomes across successive cohorts (Byington et al., 2016; Chen et al., 2016; Feldman et al., 2012; Fornari et al., 2018; Ogunyemini et al., 2010; Pfund et al., 2014; Pfund et al., 2013; Pfund et al. 2006; Pfund et al., 2015; Thordyke, Gusic, & Milner, 2008). Infrequently, researchers compared outcomes between those who did or did not participate (Brown et al., 2008).

Surveys were used to capture responses in the majority of the studies. Most study authors used researcher-constructed surveys. Infrequently, those surveys were not validated for purposes very closely related to the ones being studied (Branch et al., 2014; Roh et al., 2016). The predominant use of non-validated researcher-constructed surveys raises questions regarding whether potential outcomes were appropriately or effectively assessed. Notably, a survey validated for a use in one setting may not be valid in another. Survey items were used to report demographics, current teaching practice, changes in participants’ attitudes, changes which that participants planned to make in the future, and participant opinions about the specific faculty development program (Branch et al., 2014; Gjerde et al., 2008; Johnson, 2016; McAndrew et al., 2013; McAndrew & Pierre, 2013; Roh et al., 2016). In addition to the use of close-ended items, researchers also employed open-ended questions (Hartford et al., 2017).

Qualitative data were collected primarily through semi-structured interviews (Behar-Horenstein, et al., 2010; Christofilos et al., 2014; Liben et al., 2012; Weber et al., 2016), focus groups (Behar-Horenstein, et al., 2015b; Wallace & Infante, 2008), reflective writing (Behar-Horenstein, et al., 2018) and discipline-based teaching presentations (Behar-Horenstein, et al., 2009; Behar-Horenstein, et al., 2010). Qualitative study results were likely limited by participant subjectivity due to convenience sampling and failure to mitigate researcher bias (Archibald et al., 2017; Baker et al., 2018; Greenwood, & Ewell, 2018; Hartford et al., 2017; Kim et al., 2017; Power et al., 2018).

Specifically, all mentoring studies relied on convenience samples. With the exception of Byington et al. (2016) other studies suffered from threats to maturation, a lack of randomization, limitations inherent to self-reporting and the potential for social desirability bias. All but three studies (Fornari, et al., 2018; Jackevicius, et al., 2014; Pfund, et al., 2006) lacked corroboration for their findings. These researchers included a qualitative component such as an analysis of reflective writing or open-ended responses.

Workshop-based teaching studies quantitative articles suffered from potential selection bias. Participants were not randomly selected, or randomly assigned to treatment and control groups. A few studies lacked use of a control group. Second, two studies with less than 20 participants for different conditions may have suffered from insufficient power to support the statistical inference (VanVoorhis & Morgan, 2007). Studies with low response rates, or a limited pool of invited participants typically reduce generalizability (Gadbury-Aymot et al., 2014). Third, studies that relied on self-reported data lack objectivity and might suffer social desirability bias. Specifically, in the pre- and post-test design studies, there may have been a response shift bias resulting
from self-reported data. Also, several researchers did not report which statistical test they applied when trying to detect pre- and post-test differences or when they tested their hypotheses (Branch et al., 2017; Johnson, 2016; McLeod et al., 2008).

The most frequent methodological problem among all workshop-based teaching quantitative studies was a lack of baseline measurement. The absence of a pre-test measurement weakens conclusions, especially in those studies where researchers reported gains in knowledge or skills. This observation suggests that objective measures of the effects of faculty development on learners are needed (Wallace & Infante, 2008). In some studies, the number of individuals who completed the training and participated in interviews or focus groups was low. One additional issue relates to those who did and did not participate since the reported findings might not be representative of overall perspectives.

Overall, all studies relied on small sample sizes (Behar-Horenstein et al., 2015b) although it is important to point out that this might be unavoidable in faculty development programs. Small sample size limitation is well known for its often inability to achieve desired statistical power. Similarly, Liben et al. (2012) reported limited generalizability due to small sample size and an imbalance of male and female participants. Researchers from another study mentioned a small sample size, self-reported data, and the length of time between the program and the interviews as possible study biases (Christofilos et al., 2014). Similarly, Greenwood and Ewell (2018) identified small sample size and response bias as methodological weaknesses.

Nearly all of the workshop-based other than teaching studies used convenience samples. The lack of randomization, comparison groups, threats due to maturation and the inherent limitations of self-report data as well as the potential for social desirability bias were methodological disadvantages. In one study the researchers (Abu-Rish Blakeney, et al., 2016) attempted to corroborate its numeric findings via the findings from structured interviews.

**Treatment**

The structure of treatment varied widely. Dutta et al. (2011) implemented an uncontrolled pilot study. Several researchers reported that their studies often had little to no structured training opportunity interventions (Al-Eraky et al., 2015; Alexandraki et al., 2017; Archibald et al., 2017; Baker et al., 2018; Balmer, & Richards, 2012; Brody et al., 2016; Gadbury-Amyot et al., 2015; Greenwood, & Ewell, 2018; Haines, & Popovich, 2014; Hartford et al., 2017; Holyfield & Berry 2008; Hull et al., 2015; Jackevicius et al., 2014; Johansson et al., 2009; Kogan et al., 2017; Krautscheid et al., 2008; Littman et al., 2017; McAndrew et al., 2011; Mookherjee et al., 2014; Møystad et al., 2015; Phitayakorn et al., 2016; Ratka, et al., 2017; Rics et al., 2009).

Treatments ranged from brief, single sessions to medium and lengthy interventions. Single interventions included 4-hour workshops. Medium interventions included 10 case-based seminars held during monthly half-day meetings over 5 months (i.e., Campbell et al., 2013; Google et al., 2016; Kim et al., 2017; Koblinsky et al., 2015; Reamy et al., 2012; Yanamadala et al., 2016). In another study, researchers measured effects of a program that consisted of an 8 to 10-week training internship, where the goal was to support faculty development and career trajectories (Campbell et al., 2013). Lengthy interventions were evidenced by year-long (Abu-Rish Blakeney, et al., 2016; Chan et al., 2018; Sheehan, & Brannan, 2013; Taylor & Berry, 2008) or longer, namely an 18-month small-group curriculum that combined experiential learning of teaching skills with critical reflection, and a three-year program that evaluated a three-year program that offered mentor and financial support to its students (Hickey et al., 2014). In Hickey et al. (2014) participants were expected to attend meetings, trainings, and networking events, as well as disseminate a research project. Others reported lengthy treatment interventions (Isenburg et al., 2017; Williams et al., 2008). Sheehan & Brannan (2012) reported the implementation of a one-year fellowship program that utilized a blended curriculum, online modules, and in-person meetings in order to improve skills and knowledge. In a two-year long program, some faculty received funding that permitted participation in mentoring opportunities to disseminate research, and conference attendance where the goal was to promote career development (Brown et al., 2008). However, other researchers did not specify the nature of training opportunities or the length of each session.
Recommendations for Future Studies Among Articles Reviewed

Researchers of the studies reviewed offered a plethora of suggestions for future studies. For example, among the mentor studies, Byington et al. (2016) and Feldman et al. (2012) recommended adding a comparison and control group to determine if impacts were due to the program. Researchers recommended determining the long-term impact of training on actual mentoring performance (Feldman et al., 2012; Pfund et al., 2013) and whether participants would translate skills to clinical environments (Fornari et al., 2018). Pfund et al. (2015) suggested assessing differences by contexts and demographic variables, while Thorndyke et al. (2008) called for multilevel outcome assessment beyond 12 and 18 months to demonstrate impact and return on investment. As an extension of the applicability of findings, Phitayakorn, et al. (2016) recommended that more research is needed to understand the best method to pair mentors and mentees and more objective measurements of academic surgery success. Brody et al. (2016) urged studying both vertical and lateral mentoring. The latter was suggested to better understand how mentoring crosses institutional boundaries and helps individuals early in their academic career expand their networks.

Workshop-based teaching study researchers argued that how the program affects participants’ long-term development should be examined in future studies (Branch et al., 2014; McAndrew & Pierre, 2013). They recommended that future studies seek to determine the influence of the programs on the learning climates at institutions as well as on students’ learning outcomes (Johnson, 2016; Roh et al., 2016; McAndrew & Pierre, 2013). In terms of the design, some suggested employing experimental research designs with comparison groups. We also recommend the use of mixed methods and using objective structure measures instead of self-report as an alternative to improving the study design. Addressing the effect of the faculty development training on student outcome measures is also needed. Similarly, assessing students’ perceptions of the extent that participants’ pedagogical skills changed should be investigated. In the future, the research may focus on exploring the sustainability of perceived practice changes in the teaching environments.

Researchers recommended accumulating small randomized trials and focusing on well-defined components of the subject area (Branch et al., 2017). Other study researchers recommended ongoing program evaluation to determine quality of faculty development (Gadbury-Aymot et al., 2014); increasing study participants (Freytag et al., 2017); and assessing potential long-term outcomes (Narayan et al., 2012). Researchers of four studies did not include future recommendations (McLeod et al., 2008; Johansson et al., 2012; Power et al., 2018; Ebrahimi & Kojuri, 2011).

Researchers also suggested including a pretest and posttest of the students’ knowledge to determine the impact of the training, and a comparison between students who are taught using CBL with those who are taught using a lecture approach (Behar-Horenstein, et al., 2015b). Liben et al. (2012) recommended: (a) refining the assessment tool used, (b) assessing whether repeated exposure of sessions in narrative medicine leads to increase narrative learning, (c) investigating why narrative medicine is used in teaching and (d) assessing its impact on the student and the patient. Although unstated, limitations in Christofilos et al. (2014) suggested including larger sample sizes and stricter control over additional factors that may influence participant’s likelihood of following through with their statement.

Several workshop-based other than teaching study researchers recommended assessing the long-term impact of their programs. For example, Love et al. (2016) suggested longitudinal tracking to see if participants published in the subsequent two to five years after completing the program. Evans et al. (2014) advised measuring the impact of participants’ teaching skills on trainees and patients. Chan et al. (2018) proposed that future research determine if participation in online learning resulted in increased scholarly output. Beck et al. (2008) called for assessing the program’s national impact on numbers of students and residents involved, number of patients served, and long-term impact on student career choices to work with the underserved, and their philosophies of health care. To determine how well a formalized orientation program prepared and socialized nurse educators for academic roles and if there was an increase in job satisfaction and retention, Baker (2010) urged conducting large scale, nationalized studies.
Sheehan & Brannan (2013) and Love et al. (2016) stated that future studies should include a non-participant or other type of comparison group to lessen the limitations due to selection bias. Similarly, Ries et al. (2009) recommended studying the impact of faculty development on participants versus non-participants to support the usefulness of and resource expenditure on faculty development initiatives. Goole et al. (2016) advocated validating self-developed instrument through wider use of the scale and implementing ways to corroborate self-efficacy gains. Ries et al. (2009) called for developing a uniform, centralized faculty personnel database with information about hire and separation dates to ensure the veracity of studies using survival analysis. Hatem et al. (2007) recommended that faculty development program instructors consider tailoring their methodology to ensure that individual learners could establish their own learning goals and objectives and promote individualized achievement.

Heinrich et al. (2009) advised exploring the scholar impostor syndrome as it relates to career development, professional identity and familiarity with scholarship. Wilson et al. (2008) stressed the role of evaluation in understanding the process and outcome of action learning and how practitioners make sense of the new knowledge and skills they develop. Baker et al. (2018) recommended exploring faculty developer competence and its impact on learning outcomes and organizational change skills.

To enhance generalizability of quantitative fellowship studies, researchers suggested including a control or comparison group, longer follow-up period, and random assignment (Brown et al., 2008), evaluating program directors as opposed to having a self-assessment and validating these evaluations (Reamly et al., 2012). They also recommended including additional sources of data (Sheehan & Brannan, 2012; Campbell et al., 2013; Lurie et al., 2009) increasing response rates and using different sampling techniques (Hickey et al., 2014). Researchers suggested that future studies examine how prepared students are for a variety of careers, as well as ways to increase preparation for these careers (Koblinsky et al., 2015). They also asserted that future studies address issues associated with the shortage of nursing faculty, while implementing mentorship and peer to peer support (Hickey et al., 2014).

Researchers recommended that qualitative fellowship studies include other sources of data (Robins et al., 2016); correlate results from other observed outcomes (Lown et al., 2009); and use appropriate selection strategies (Kim et al., 2017). One study did not list recommendations (Archibald et al., 2017).

Literature-based review authors recommended that future studies of structured mentoring programs focus on assessing program effectiveness or efficiency (Buddeberg-Fischer & Herta, 2006) assess individual mentors on an ongoing basis and seek the perceptions of both mentors and mentees (Meagher, Taylor, Probsfield, & Fleming, 2011); conduct outcomes oriented studies in pharmacy (Guglielmo et al., 2011; Strang & Baia, 2016; focus on the educational process or the interactions between program implementation and contextual factors that affect the success of faculty development programs (Leslie, Baker, Egan-Lee, Esdaile, & Reeves, 2013), use more rigorous methods, address contextual issues and rely on the use of both quantitative and qualitative approaches simultaneously (Sambunjak, Strauss, & Marušić, 2006; Steinert et al., 2006; Meagher, Taylor, Probsfield, & Fleming, 2011); assess change and transfer to practice over time, use multiple methods and data sources, and embed studies in a theoretical or practical framework (Steinert et al., 2016); and use reliable and valid outcome measures (Steinert et al., 2006). Sorinola and Thistlethwaite (2013) emphasized that the success of future faculty development is highly dependent on the learner. They emphasized that participants must engage in metacognition so that they: (a) focus on how to learn in the future as well as in the present, (b) recognize what they know and do not know, (c) identify how they best learn, (d) develop and implement a plan to acquire the skills and knowledge needed, and (e) capably monitor their success.

**Discussion**

The findings showed overwhelmingly that the studies of participants drawn from the spectrum of those working in the health professions demonstrated positive outcomes, which suggests that faculty development makes a difference. However, the absence of negative findings seems to suggest an appearance of publication bias. Publication bias refers to the potential prejudice in estimated mean

Vol. 33, No. 2, May 2019 / 89
effect size caused by studies’ unequal probability of getting published. First described by Rosenthal (1979) as the “file drawer problem,” he pointed out that empirical studies showing statistically significant results are more likely to be accepted by journals for publication. Publication bias can impose an impediment to being reviewed or assessed in other studies such as meta-analyses (Scargle, 2000) because studies showing non-significant treatment effects tend to be disproportionately “left in the file drawer” and hence remain unpublished. It is important to point out that some if not many of the studies reviewed were designed and/or funded as evaluation projects, not as research studies. While efforts to share results with the larger community through publication are laudatory, the authors urge researchers to candidly acknowledge related limitations associated with study research design, methodology, and treatment implementation.

Other than improved knowledge, skills, and confidence levels, what are the long-term benefits of faculty development? Does participation versus non-participation increase the likelihood of retention, promotion, vitality, or job satisfaction? Given the important potential influence of faculty development, exploring its impact should exceed temporal gains.

Faculty in the health professions play an expanding number of roles: teacher, scholar/researcher, mentor, discipline leader, curriculum developer, information manager, committee member, and clinical practitioner. Other than the delivery of clinical health care, health professions training programs (Hand, 2006) provide little opportunity for the development of these skills. Since faculty are the most important core resource, they must be nurtured, mentored, and provided opportunities to advance their teaching to its fullest potential (O’Neill & Taylor, 2001; Friedman et al., 2004).

Expanding expectations for teaching and curricula effectiveness requires that faculty have the requisite skills to prepare professional students. Given that few faculty are prepared for the professoriate during their terminal studies, administrators are urged to consider investing resources directed towards ensuring that their faculty are equipped. Making a commitment to offering systematic faculty development initiatives may nurture untapped faculty talent and ensure faculty vitality and retention.

Studies regarding faculty development needs have been reported in dentistry, veterinary medicine, and pharmacy. Veterinary medicine and dental faculty reported a need to develop skills relative to grant proposal writing, documenting educational outcomes, and writing an educational research manuscript (Behar-Horenstein et al., 2016). Dental faculty also reported a need to augment their skills in developing research designs (Behar-Horenstein et al., 2014). Junior and senior pharmacy faculty reported a low level of confidence using non-traditional research methods such as qualitative methods, identifying funding resources for educational research, and use of statistical software (Behar-Horenstein et al., 2018).

Learning to teach is a dynamic and evolving process that is influenced by multiple life experiences, peers, feedback, and observation (Weber, Cable, & Wehbe-Janek, 2016). Thus, formal training on teaching is both necessary and fundamental. Teaching cannot be left to happenstance in disciplines where classroom instruction is subsequently translated into patient care (Behar-Horenstein et al., 2008). Professional schools must recognize the essential relationship between teaching and social progress more clearly because how future healthcare providers are taught is likely to impact understanding, communication, and patient care. If resources are invested to: (a) ensure quality teaching; (b) promote mentoring; and (c) facilitate positive student outcomes, then there is a greater likelihood that carefully designed faculty development will aid reaching these outcomes. Faculty development is recursive; it not only influences faculty and those they teach, but also those they care for as well.

Faculty’s lack of formal training in teaching, mentoring, and conducting educational research has significant consequences for students who look to their instructors as respected leaders as they try to emulate their behaviors in thinking, speaking, and writing (Brookfield, 1997). Without a place to learn and model these skills, educators unwittingly contribute to a pre-existing problem, the lack of development of student critical thinking skills. However, when faculty development is grounded in a theory (andragogy in this case) that involves making meaning through experience and reflection (Schon, 1987; Moon, 1999), and occurs during participation
in communities of practice (Lave & Wenger, 1991; Boud & Middleton, 2003), it can result in instructional practice changes beyond the seminar.

In summary, the authors present a review of existing studies on faculty development in the health professions. They provide a descriptive analysis of the categories, methods, and results. The authors include other descriptives such as the participants studied, recommendations, and elements of analysis and a critique of methods used. They also present recommendations for improvements in methodology for future studies on faculty development in the health professions.

**Limitations of This Study**

A potential limitation of this study was the exclusion of program reports and dissertations from this review. Including articles other than journal publications may reduce the publication bias. Furthermore, the databases that were used in this review, namely, PubMed, ERIC, MEDLINE, CINAHL, and Professional Development Collection (education) are likely published in North America, which likely limited the scope of articles which were included. In this review, we did not classify whether studies were conducted in or outside the U.S. because so few studies outside the U.S. were reported.

**Author Recommendations for Future Studies by Article Reviewed by Authors**

Based on the syntheses described in this paper, we advocate that future faculty development studies implement a more robust approach to data collection and reporting findings. We recommend that:

1. Studies include pre-test and post-test measures with control and treatment groups or implement time series studies to assess the impact of the intervention beyond the conclusion of the faculty development initiative.

2. Researchers avoid publishing outcomes based on single shot, pre-experimental designs. Given their associated myriad limitations, researchers are advised to increase the use of more robust research designs, methods that can be replicated, and outcomes that are likely to advance an understanding of the efficacy of faculty development studies.

3. Researchers use validated measures, or at a minimum report the reliability estimates of the instruments that are used.

4. If a single survey is used, researchers simultaneously assess and report the potential of social desirability bias using the Marlowe Crowne Social Desirability Scale. (Loo, & Thorpe, 2000; Strahan, & Gerbasi, 1972). This will help readers appreciate the degree to which outcomes derive from the treatment versus artifacts.

5. Qualitative studies include more than one form of data. Rather than solely relying on interviews or focus groups, include analyses of reflection, writing, observations, or other data sources.

6. Researchers report their positionality, including their knowledge and assumptions regarding the phenomena that they study and how they mitigated personal bias. This type of reporting will heighten the transparency of studies and explain to readers why and how outcomes may be informed by researcher perspective and expertise.

**References**


* Indicates studies cited.


Reamy, B. V., Williams, P. M., Wilson, C., Goodie, J. L., & Stephens, M. B. (2012). Who will be the faculty of the future? Results of a 5-year study growing educators using an immersive third postgraduate year (PGY-3) faculty development mini-fellowship. Medical Teacher 34(6), e459-e461.


Linda S. Behar-Horenstein, Ph.D. is Professor Emeritus, University of Florida. Her research interests include faculty development, cultural competency, and the assessment of behavioral, cognitive, and attitudinal change among healthcare faculty, staff, and trainees.

**Huan Kuang** is a Ph.D. student in the Research and Evaluation Methodology program at the University of Florida. Her research interests are statistical and measurement methods as applied to educational and social science studies.

**Lissette Tolentino** is a Ph.D. student in Research and Evaluation Methodology program at the University of Florida, specializing in psychometrics. Currently, she is an instructor and teaches measurement and evaluation in education for undergraduate students.

**Huibin Zhang** is a Ph.D. student in School Psychology at the University of Florida. His research focuses on the effects of mental health intervention with statistical methods.”