A National Survey of Faculty Development Evaluation Outcome Measures and Procedures

Katrina A. Meyer
Vicki S. Murrell
The University of Memphis

Abstract

This article presents the results of a national study of 39 higher education institutions that collected information about their evaluation procedures and outcome measures for faculty development for online teaching conducted during 2011-2012. The survey results found that over 90% of institutions used measures of the faculty person's assessment of satisfaction and usefulness of the training itself, rather than student outcomes or changes in teaching methodology. Online evaluations were utilized by 80% of institutions and focus groups were used by 21% of institutions.

Introduction

The Online Learning Consortium's five pillars of quality (Online Learning Consortium, n.d.), look at faculty development vis-à-vis the pillar of Faculty Satisfaction. This pillar stresses the importance of faculty satisfaction within the online teaching experience, and notes the need for faculty to improve their online instruction. The Online Learning Consortium (formerly the Sloan Consortium) characterizes faculty satisfaction as resulting from institutional support, which includes the opportunity for "training in online instructional skills" (Online Learning, n.d., para. 5). Thus, faculty development in online teaching is a critical foundation for quality online education. The Online Learning Consortium created an advisory panel of practitioners and researchers (http://onlinelearningconsortium.org/jaln_advisory_panel_fs) focused on faculty satisfaction, development, and support. A first step in the work of the advisory panel was to identify the current state of knowledge of the evaluation measures and processes for faculty development for online teaching, which is the aim and purpose of this current research study.

The next section answers the question, what can we learn from the research literature on faculty development that can improve our efforts to help faculty learn how to teach online?

Review of Literature

Faculty Development Models for Online Teaching

The literature contains many articles about specific faculty development programs at specific universities. Most of these programs offer several training activities; in other words, they offer multiple training opportunities in different settings using different tactics and cannot be simplified as one approach or one activity. However, the evaluations used for these types of multi-dimensional programs do not allow for assessing each part of the program. As a result, the often-modest evaluations of these faculty

development programs use simplistic outcome measures. Additionally, some of the current research does not provide ample information on the content of the faculty development itself.

For example, a faculty development for online teaching program was named Creating Optimum Learning Environments (CREOLE) and involved collaboration between two unnamed institutions, a community college and an eastern research university (Schrum, Burbank, Engle, Chambers, & Glassett, 2005). Four modules were developed, each by experts in the following fields: "learning theory, motivation research, blending face-to-face classes with online supplements, and creating completely web-based courses" (Schrum et al., p. 280). The four modules were eventually offered as one graduate-level course, and an evaluation was given. However, no further information was provided about the modules or the course. In fact, the article does not provide sufficient detail about CREOLE to allow another institution, developing its own faculty development effort, to know what these institutions are doing.

Another example was found in the New England Center for Inclusive Teaching (NECIT), which provided faculty development seminars at seven colleges and universities (Daly & Dee, 2009). The seminars met weekly for a semester, and included faculty participants reflecting on their professional lives, identifying strengths and competencies, and encountering new ideas. The actual innovative ideas are not described. However, the evaluation was insightful. It involved interviewing faculty to discuss their participation in the seminars as well as "overall growth and development as a faculty member" (Daly & Dee, 2009, p. 14). The interviews were conducted by individuals not involved in the implementation of the seminars. The findings indicated the seminars promoted significant changes to the ways participating faculty pursued teaching.

Pennsylvania State University and its World Campus have been in the forefront of developing faculty development for its online teachers. Ragan et al. have described their faculty development efforts, including an early "Online Learning (OL) Series" that includes 12 courses organized in four levels, with 1000-series courses focused on orientation to online learning for the novice teacher, 2000-series courses focused on pedagogy, 3000-series courses focused on new and emerging technologies, and 4000-series course on authoring online courses (Ragan, Bigatel, Kennan, & Dillon, 2012). Individuals interested in the learning outcomes and teaching behaviors each course produced are encouraged to read the article (Ragan et al., 2012). While it is not impossible to go from the teaching behaviors outlined in the series to the likely content of each course, the precise contents of each course are not spelled out.

These examples are not intended to criticize what these programs and institutions have done; the intent is just to note that the exact contents of the faculty development program are not specified. And in fact, given that these programs involve activities over a semester (Schrum et al., 2005; Daly & Dee, 2009) or possibly several semesters (Ragan et al., 2012), it is difficult to identify which specific activities may be evaluated or tied to an outcome of interest. If a program involved ten activities, which one(s) affected the faculty person's subsequent online teaching?

Many institutions have implemented a variety of faculty development programs aimed at helping faculty design and teach online courses in addition to using technology wisely in a traditional classroom. The University of Central Florida requires all faculty teaching online to participate in a 70-hour faculty development course. Central Michigan University also implemented professional development for online faculty that went beyond the one-time workshop to include weekly tips, online mentoring, and online teaching resources. The University of Houston system created the CampusNet Online Workshop program that includes faculty networking, hands-on practice, and a comfortable environment for asking questions of all kinds (Kidney & Frieden, 2004). The University of Colorado created a Web Camp, offered over the summer and winter months, where faculty participate in a week-long intensive workshop that also includes hands-on training and design (Lowenthal & Thomas, 2010). Michigan State University used master's students enrolled in an instructional design course to help faculty design an online course (Koehler, Mishra, Hershey, & Peruski, 2004). The Open University tackled development of its faculty for mobile learning by providing events, communities, exploratory spaces, and resources (Kukulska-Hulme, 2012). PBS's Teacherline has extensive faculty development opportunities based on problem-based approaches (Southern Regional Education Board, 2009) and developed a comprehensive ("3D")

evaluation in the Appendix (Storandt, Lacher, & Dossin, 2012). Capella University uses a META model (for Mentoring, Engagement, Technology, and Assessment) for its faculty development for online teaching (Dittmar & McCracken, 2012). The University of Cincinnati funded grants that were proposed by faculty or departments (Camblin & Steger, 2000). Colorado State University used active mastery learning, using Bloom's taxonomy and systems theory to create faculty development for online courses (Puzziferro & Shelton, 2008). Fetters and Duby (2011) described a faculty development program at Babson College which tied Rogers' (2003) theory of innovation diffusion to blended learning. Florida Atlantic University developed a detailed plan for a new central eLearning unit (Orozcco, Fowkles, Jerzak, & Musgrove, 2012). A community of practice approach pulled nursing faculty together across multiple campuses for faculty development (Reilly, Vandenhouten, & Gallagher-Lepak, 2012). A three-tiered approach was used for online faculty development, from orientation, to mentoring, and ongoing support (Vaill & Testori, 2012). Finally, a three-week training session at University of Wisconsin-La Crosse was described in Koepke and O'Brien (2012). Given this array of training models, it is challenging to provide comparisons of effectiveness or, indeed, of evaluation metrics.

Disentangling Treatments

In order to attempt to make sense of the available research, it will help to delve into the details of the evaluation tools. Once again, a few in-depth examples will have to stand in for several instances. Daly and Dee (2009) state that the NECIT program asked faculty to reflect on their professional (especially teaching) skills. The interview questions asked of the faculty participants are provided in an Appendix and appear to be comprehensive (Daly and Dee, 2009). However, the questions focus on what was going on in the faculty person's thinking and do not tie these reflections to any particular activities undertaken during the seminars. In other words, the evaluation tool focuses on the ultimate goals of the training (faculty perceptions and understandings) rather than on the training itself. That is not a criticism of the evaluation, but a comment on the inability of the evaluation to help the developers identify activities or interview questions that worked and provided useful information.

The CREOLE project provided several mean responses on statements ("I gained knowledge," "I improved DE [distance education] teaching") (Schrum et al., 2005, p. 285-286). This is another example of a good summative evaluation that does not provide the developers with detailed input. However, it is important to note that interviews with participants did elicit specific comments about the training that could provide important feedback.

Penn State also provided a detailed evaluation of its Online Learning Series, including a factor analysis of 64 statements regarding the types of behaviors important for online instructors. Based on their analysis, Penn State synthesized these statements into seven areas of competency: Active Learning, Administration/Leadership, Active Teaching/Responsiveness, Multimedia Technology, Classroom Decorum, Technological Competence, and Policy Enforcement (Ragan et al., 2012). These competencies make sense and are consistent with the trend within higher education to focus on assessing learning outcomes; however, the way in which participant responses could be used to inform faculty developers as they upgrade their coursework is not clear.

It is likely that each of the programs profiled above has a process of evaluating the basic components of its training, but this information is not included in the published research. The intent of this discussion is not to infer that such evaluations do not occur, but that this data has not been reported.

A larger issue is this: with the numerous deliveries of various training content available, how can evaluations unravel all of these treatments so that faculty developers can know which one is more likely to be effective? This point is not just about knowing what is working (and what is not), but about identifying those parts of the development effort that can be eliminated should institutional and program budgets be constrained. It is likely that faculty developers may increasingly be asked to produce more or better results with the same or lower budget and thus faculty development professionals must ask which activities to keep and which may need to be dropped. To make this critical decision, faculty developers need to analyze how each element of the faculty development effort contributes to the changes it wants.

King (2004) attempted to disentangle the influences on faculty by asking them which activities influenced their perspective transformation: 86.1% of participants mentioned learning activities, which were further broken down into discussion (69.4% of participants), journals (52.8%), reflection (47.2%) and readings (47.2%). A total of 72% of participants also mentioned the influence of other persons, including a professor (33%), classmate (28%) or other student (28%). Disentangling a multi-dimensional (and multi-week) effort is enormously difficult to do and will require a different approach to evaluation, perhaps exploring just-in-time evaluation (an evaluation screen that pops up at the completion of an activity), reflective evaluations (asking participants to identify what activity helped them to learn or understand a concept as was done by King, 2004), or authentic assessments (asking participants – or program completer—to produce an example of the learning intended). Program evaluations may need to go well beyond what is usually conducted under the name of formative and summative evaluations, and might include multiple formative evaluations, designed to capture learning that is the result of a specific activity, as well as longer-term evaluations, which would capture the learning that may take time to be integrated. The goal of such efforts to disentangle the effects of a variety of treatments is to understand—precisely—what is working and for whom. It is also important to determine what is not working and for whom.

Rigorous Evaluations

In an earlier review of the faculty development literature, one of the criticisms made of faculty development evaluations is their lack of rigor (Meyer, 2014). This point is different from disentangling complicated faculty development programs. Rigor has more to do with a lack of clear and definable objectives, measureable outcomes, and a data collection method that does not preclude hearing bad news. Good evaluations allow faculty developers to ask the tough questions and to get the news that something is not working (or working as assumed) and should therefore be revised or eliminated.

Kucsera and Svinicki (2010) conducted a literature review of nine journals that published faculty development evaluations between 1992 and 2007. Unfortunately, they concluded that only a few studies "met best practice standards" (Kucsera, & Svinicki, 2010, p. 5) for program evaluations or for precision in those evaluations. To put this insight into perspective, only 47% of the articles included in this review could be construed as "research" (defined broadly as using either quantitative or qualitative methods and including any outcome measures at all). The reason for the lack of good program evaluations may be because faculty development programs are complex (comprising many parts and activities, as noted above), take place over an extended period of time, and enroll small samples of faculty who are evaluated immediately at the end of the training rather than being followed over time. While randomization and other qualities of good evaluation may never be possible, given the constrained budgets of faculty development programs, the authors conclude that perhaps qualitative research methods—such as ethnographies, anthropological methods, and case studies—would be more likely to lead to useful insights into the training provided to faculty.

Evaluations of faculty development programs or trainings have increasingly depended on qualitative research methods (as recommended by Kucsera & Svinicki, 2010, above) and eschewed the identification of outcome measures *a priori*. For example, Lackey (2011) interviewed six participants in faculty development programs for teaching online and found that one-on-one assistance as well as both technical and pedagogical training were most beneficial for preparing them to teach online. While an example of good qualitative research, more studies like this one and McQuiggan's (2012) are needed as well as larger studies that focus on more and more diverse faculty and teaching changes over time. So far, the literature lacks rigorous research comparing the effects of different faculty development models, programs or activities, or comparing these effects across different institutions. It is understandable why this has not been done, since it would be costly to conduct and gather such cross-institutional data. However, perhaps individual faculty developers can pool resources across institutions and across institutional types to undertake such an endeavor in the future.

What is disconcerting is the lack of stringent evaluations of some (although not all) of the faculty

development programs found in the literature. Evaluations can help developers at other institutions decide which interventions work best based on particular outcome measures that support the conclusions. In all fairness to these institutions and their hard-working faculty development staff, collecting detailed evaluations may not have been of immediate concern since they were likely experiencing pressures to get something underway and respond quickly to a felt need. However, to build expertise and understanding of what specific activities work and why, the field of faculty development may need to contemplate a number of changes to the program evaluations they are currently doing.

Outcome Measures

To design a quality program evaluation, the faculty developer needs to know what the training is intended to achieve. For example, are the outcomes of interest a change in faculty perceptions of their teaching roles, as in Daly and Dee (2009), or a set of teaching competencies as in Ragan et al. (2012)? The evaluation should measure how well the learning objectives were achieved, but this tenet is not reflected in the research. An earlier review of the literature on faculty development for online teaching (Meyer, 2014) found that not many evaluation measures are named, and the few that are mentioned are not particularly clear or robust. Here is a partial list: number of new educational programs added (Gruppen, Frohna, Anderson, & Lowe, 2003), opinions about effectiveness of the training (Maxwell & Kazlauskas,1992), adoption of case studies in instruction (Atheny & Hoffman, 2007), improved teaching (DiLorenzo & Heppner, 1994), professional growth of faculty (Lindman & Tahamont, 2006), usefulness (Steinert et al., 2006), satisfaction or relevance to participant (Lavoie & Rosman, 2007), use of portfolios (Haviland, Shin, & Turley, 2010), more cooperation across disciplines (Camblin & Steger, 2000), and confidence with and attitudes about assessment (Edwards et al., 2001). Steinert et al. (2006) also reports outcomes of training by intervention type (e.g., workshops, short courses) which begins the process of tying outcomes to treatments, although in less detail than may be helpful. Lavoie and Rosman (2007) compiled a variety of outcome measures used in faculty development efforts for medical educators, from a positive change in attitudes to increased knowledge of and change in teaching behavior or student learning. Storandt et al. (2012) included such outcome measures of the faculty development provided to PBS online teachers as a score on the course rubric, learner course grades, and turnover of faculty. Edwards et al. (2001) also identified possible moderating variables: faculty who think of themselves as facilitators of learning (rather than disseminators of information) or have a higher sense of personal efficacy were more successful in completing all of the faculty development modules. Knowledge of pedagogy and innovative course design were also important for successful change. Koepke and O'Brien (2012) found that training changed faculty's conceptions or "myths" of online learning, away from more critical or negative points of view as well as changed several teaching behaviors: from adding video and audio files to providing more, and more prompt, feedback. Orozco et al. (2012) found that faculty development yielded such outcomes as increased comfort with using technology but also 27 detailed evaluations of the training provided (from "objective clarity" to "ease of interaction" to "discussion effectiveness"). While this list is deliberately brief, it is perhaps clear that the program outcomes are described in less detail than what may be helpful to other faculty developers. Some do a better job of delineating outcomes of interest and either developing an evaluation instrument to measure those or delving more deeply into the factors that lead to faculty change (Edwards et al., 2001; Skeff et al., 1997). Others are too amorphous or poorly defined to be identified with any confidence.

Commonly used outcome measures for faculty development for online teaching include usefulness (as assessed by the participant); willingness to recommend the training to another faculty; self-reported knowledge or ratings of self-efficacy; and changes in behavior, beliefs, and attitudes (Grant, 2004). Schrum et al. (2005) included many of these measures, but also utilized self-reports of participants altering their pedagogy, redesigning their courses, or experiencing community online. In an attempt to understand the reasons for satisfaction with faculty development, Grant (2004) investigated an individual's intrinsic factors (convenience, comfort, interests) and extrinsic factors (external pressure to teach online). In a more detailed evaluation of Purdue University at Calumet's online courses for faculty, 47 participants underwent a year-long development program and then were followed over four years; the

program evaluation included 72 evaluation items, from "I am satisfied" to "My on-campus teaching has improved" (Hixon, Barczyk, Buckenmeyer, & Feldman, 2011). Potter and Meisels (2005) included such authentic measurements as giving an example of how the faculty development impacted the individual's "ability to think critically and use information to solve problems and answer questions," "understanding of science in the news," and applications of "problem-solving approaches learned" in the training (p. 194). These examples are then enhanced with further reflective questions that focus faculty persons' attention on their teaching beliefs and application of concepts to other courses they teach.

What is clear from this information is that outcome measures are often poorly defined or poorly measured and they depend on the honesty and self-understanding of those undergoing the training. While many faculty persons possess these qualities, evaluations of faculty development should not depend solely on simple or easy measures (such as, "Are you satisfied with your training?"). For example, what is actually being measured when a faculty person indicates satisfaction with training? Does satisfaction capture how much they enjoyed themselves, that they liked the facilitators, or that they thought the training would be helpful to them? Here is another complication: Liu (2012) surveyed 11,351 students taking distance education courses at 29 colleges and found that student reasons for taking the course influenced their ratings of instructors' teaching skills. Therefore, and by extension, the faculty person's reasons for undergoing training may also affect the evaluation of the training as were the students' evaluations of teaching in Liu (2012)

Faculty developers need to improve ways of identifying authentic outcomes of training. Although reliable and rigorous assessments are often more cumbersome and more costly than simple Likert-scale items, perhaps they are a way to give flesh to the bones of our current set of outcome measures.

Who Should Evaluate?

What may be clear from the above discussion is that faculty persons who undergo training are the predominant source of evaluative comments or judgments. This begs the question of whether others might provide more objective assessments, or at least an alternative view, of the training's results. For example, should students have a role in evaluating instruction provided by faculty persons who have been through training to teach online? Stehle et al. (2012) extended the issue of using students to evaluate a faculty person's online teaching by also assessing the correlation between student evaluations of teaching effectiveness and student learning; unfortunately, the results were equivocal (Stehle, Spinath, & Kadmon, 2012). These studies encourage us to explore the usability or practicality of having students evaluate their teachers' online teaching as a way to judge the effectiveness of faculty development or correlating faculty development experiences with student learning. Another possibility is administrators' (or other outside entity such as instructional designers from another institution) evaluations of faculty development for online teaching. Allen and Seaman (2012) asked a national sample of faculty and administrators to respond to a number of questions, two of which dealt with faculty development. For both questions ("My institution offers excellent training and support for using digital tools in the classroom" and "My institution offers excellent training and support for the use of lecture capture"), administrators chose "strongly agree" by over 10 percentage points more frequently than faculty. This may be a case of "what you see depends on where you stand" but could be worth further exploration.

Differences by Carnegie Classification

Higher education institutions in the United States have been classified into like groupings based on work originally done in 1973 by the Carnegie Foundation for the Advancement of Teaching. The classifications have been regularly updated as criteria were revised to address changes in higher education institutions and to better capture the diversity of institutions. The Carnegie classification also captures how institutions at "lower" levels work to "raise" their classification, also known as "Mission Creep" (Longanecker, 2008) or "striving" (O'Meara, 2007). The Carnegie classification has been a useful tool in research covering a wide range of topics, from funding to faculty to information technology. For example, Carnegie classifications have been instrumental in understanding differences in uses of institutional websites. Meyer (2008a and 2008b), Wilson & Meyer (2009), and Jones & Meyer (2012); found that

types of Carnegie institutions may have more, or less, funding or staff to provide support to students or the general public through the use of various technologies (e.g., web-based services, Facebook sites). Based on this kind of information, it is reasonable to expect that having the expertise and staffing to design, implement, and use detailed evaluations may also vary by the Carnegie type of the institution.

Research Questions

The four research questions for this study are:

What outcome measures are higher education institutions using to evaluate their faculty development for online teaching?

Are there differences in use of outcome measures by the institutions' Carnegie type?

When and how are faculty asked to evaluate the training they receive?

Are there differences in evaluation options by institutional Carnegie type?

Methodology

Research Design and Instrument

This study is based on survey research that collected information from participating higher education institutions. As this is one of the first attempts to assess faculty development for online teaching practices in a national sample, a survey research approach is appropriate.

The instrument used was developed by the first author and was based on a thorough review of the published literature on faculty development for online teaching (Meyer, 2014).

A draft of the instrument was reviewed by the Online Learning Consortium Advisory Panel for Faculty Satisfaction as well as representatives of the Online Learning Consortium and WCET organizations, including organizational leaders and researchers, faculty developers, and faculty who conduct research on this topic. Because this would be a national study of faculty development for online teaching, the findings would be of interest to members in both organizations. This process resulted in many additions and revisions that resulted in a cleaner and more comprehensive instrument. Given the face validity of the items, the data resulting from the instrument are valid; a test for reliability was not conducted due to the small number of responses which should be considered a limitation of the study.

The present study focuses on three items from the instrument (which included a total of 26 items). Two items dealt with the two issues in the research questions. First, the institution would indicate which outcome measures it used to evaluate training including faculty satisfaction with training, faculty assessment of the usefulness of training and relevance of training, faculty developed skill or competency, faculty willingness to recommend training to other faculty, faculty assessment of elements of training (explanation of instructional design principles, research behind training, mentoring relationship, clear purposeful communications, improvement in teaching, changes to their face-to-face teaching, changes in attitude toward online learning, changes in perception of teacher's role), student evaluations of faculty teaching, students' course grades, students' cumulative GPA, students' grades for specific course assignments. Answers provided allowed for two responses (numerical coding in parentheses): Yes (2), No (1).

Second, the institution indicated how it conducted faculty evaluations to evaluate their training. Institutions were allowed to check all options that were used. Options included: after all training is completed, after each elements of training is completed, after all training of a particular type is completed (all presentations, all discussions, etc.), paper evaluation tool, online evaluations, delaying evaluation (after a passage of time), one-on-one interviews, and focus groups. Two responses were provided (numerical coding in parentheses): Yes (2), No (1).

Third, the individual completing the survey was asked to indicate the Carnegie classification of the institution, if this was known. If it was not known, this item was to be left blank and the first author used the following link: http://classifications.carnegiefoundation.org/lookup_listings/institution.php. This

look-up function is offered by the Carnegie Foundation for the Advancement of Teaching.

Population and Sample.

A request to complete the survey was sent from an Online Learning Consortium officer to the official representatives of higher education institutions that are members of the Online Learning Consortium. This totals 407 institutions. The first author also sent a request to complete the survey to the online WCET (WICHE Cooperative for Educational Telecommunications) Discussion Board, which is open to any individual who is an employee of the 295 WCET member organizations. The request asked that the survey information be forwarded to the individual responsible for faculty development at the institution. It also asked institutions to beware of duplicative emails, since institutions can and are members of both organizations. Two items were included in the instrument (institution name and individual's name) so that duplicative responses could be identified and one eliminated.

This sampling procedure was chosen since it would most efficiently get to institutions that offer online learning and would be most likely to offer faculty development for online learning. Since it was these practices of faculty development for online learning that were the focus of the research, sending the instrument to a wider range of institutions would not likely be effective. Therefore, while the results reflect the practices of institutions that offer faculty development for online learning, it is a limitation of this study that the results cannot be generalized to all higher education institutions.

Responses were received from 39 institutions including 13 doctoral/research institutions, 12 master's institutions (institutions offering degrees up to and including master's degrees), three baccalaureate institutions, and 11 associate's institutions. Given that institutions do belong to both the Sloan and WCET organizations, it is impossible to calculate a response rate. Given the low number of baccalaureate institutions that responded, results are reported but are not interpreted or discussed in comparisons to other Carnegie types. Responses from one special focus institution and one international institution were deleted for the analysis that used Carnegie classifications in order to protect the anonymity of responses. However, not all 39 institutions answered every question on the survey, so the number of institutions upon which results are based is noted throughout the results section.

The survey was completed by individuals responsible for faculty development with titles including coordinator, director, dean, and Vice President, and located in Academic Affairs (57.9%), Chief Information (Technology) Office (23.7%), an academic department (18.4%), an academic college (13.2%), or Central/System Office (5.3%). This diversity of job titles and locations seems to imply faculty development for online teaching is occurring in many different locations within institutions and under different personnel. In other words, institutions have approached faculty development with a variety of organizational locations and titles.

Data Collection

The instrument was created within SurveyMonkey.com, which provides a flexible set of question types for the researcher with options for long-term data storage. The initial request to institutions for responses to the instrument was sent January 4, 2013 and the deadline for receipt of responses was February 1, 2013. On this date, the survey was closed to further responses and analysis began.

Data Analysis

For research question 1 ("What outcome measures are higher education institutions using to evaluate their faculty development for online teaching?"), answers are reported by frequency of the answer options (yes, no, don't know, will use in future).

For research question 2 ("When and how are faculty asked to evaluate the training they receive?"), answers are reported by frequency of evaluation options. For research questions 3 and 4, the answers are reported by Carnegie classification. No statistical analyses of differences were attempted due to the low sample size.

Findings and Discussion

Outcome Measures

      Table 1 reports the types of outcome measures used in evaluations of faculty development for online teaching by frequency (number of institutions reporting use of the measure, percent of institutions using this measure, and the rank order of the measure based on frequency of use). Due to the number of ties among the items, the rank of an item should be interpreted as suggestive; references to rank in the discussion which follows are intended to help clarify findings.

Table 1 Outcome Measures Used in Evaluations (n= 39 institutions)

Outcome measure	Frequency	Percent of institutions	Rank
Faculty satisfaction with training	37	95	1
Faculty assessment of usefulness of training	35	90	2
Faculty assessment of relevance of training	34	87	3
Faculty developed skill or competency	28	72	4
Faculty willingness to recommend training to other faculty	24	62	5 (tie)
Faculty assessment of clear, purposeful communication	24	62	5 (tie)
Faculty assessment of improvement in teaching	22	56	6
Faculty assessment of changes to their face-to-face teaching	19	49	7 (tie)
Student evaluations of faculty teaching	19	49	7 (tie)
Faculty assessment of changes in attitude toward online learning	18	46	8
Faculty assessment of changes in perception of teacher's role	17	44	9
Students' course grades	11	29	10
Faculty assessment of explanation of instructional design principles	9	23	11
Cost of training	8	21	12
Faculty assessment of mentoring relationship	7	18	13 (tie)
Students' grades for specific course assignments	7	18	13 (tie)
Students' cumulative GPA	5	13	14
Faculty assessment of research behind training	2	5	15

It appears that this sample of institutions is consistent in its reliance on three faculty assessments of the training received: their satisfaction, assessment of usefulness, and assessment of relevance (ranks 1, 2 and 3). This is consistent with the findings of Grant (2004) and Steinert, et al. (2006), who reported usefulness (as assessed by the participant) to be a commonly used metric of faculty development, and with Lavoie and Rosman (2007), who used satisfaction or relevance to the faculty person in evaluation measures. These assessment measures may be seen as the reliable, tried-and-true measures by trainers who work in a variety of settings and teach a variety of topics and are reflected in the research. What is perhaps more interesting is the reliance by at least 72% of the institutions on some sort of assessment of the faculty's skill or competency with online teaching (rank 4), as reported by DiLorenzo and Heppner, (1994) and Hixon et al. (2011). This may be a worthwhile parallel to the emphasis on competency assessments for student learning required by accreditors. Three outcome measures of particular interest (ranks 5, 6, and 7)—because they emphasize the importance of encouraging critical reflection on the part of the faculty person (the faculty participants' assessment of improvement in their teaching, changes they made to their face-to-face teaching, and changes in perception of their teaching role)—are used by only half of the institutions. This lack of reflective evaluation is consistent with what was found in the review of literature. If the greater use of transformational learning among faculty developers (as noted in Meyer, 2014) were responsible for this finding, more faculty might be encouraged to undertake some serious reflection on their beliefs about teaching as they undergo faculty development for online teaching.

At least half of the institutions use student evaluations of faculty teaching (rank 7), but only one-third use student course grades (rank 13). It is both good and bad news that one-third of the institutions are attempting to tie the success of faculty training to student learning. However, the rank of this measure, as well as the low rates for other measures that attempt to capture student learning (rank 13 and 14)—student grades for specific assignments and cumulative GPAs—may be for two reasons. First, an institution may face technical difficulties of gathering such data, given the complications of student databases. Second, it may be too time-consuming to connect student data to faculty participating in development activities. Similarly, given the multiple influences on student grades, there may be too many confounding variables to tell the developers anything of importance. As Stehle et al. (2012) found, resulting data may be ambiguous. It might be useful for those institutions that use such measures to share what they are learning about the validity of using these measures so that they can help other developers improve the evaluations of training at their own institutions.

Two other infrequently used evaluation measures are of particular concern. First, few (21%) of institutions attempt to capture the cost of training (rank 12), and fewer yet (5%) evaluate how and whether faculty participants understand the research that provides the basis for the training (rank 15). This latter finding has several explanations, including research findings in the training are not valued by faculty developers, it is not considered appropriate for the faculty, or an outcome measure has not been devised to assess this. There may be other explanations which need to be explored further in future research studies. Given the findings of Meyer and Murrell (2014a), where only 48.9% of the institutions indicated they included research on online learning in their faculty development materials, the simplest explanations may be "all of the above."

To ease comparisons between Tables 1 and 2, the outcome measures in Table 1 are presented in the same rank order in Table 2. Table 2 presents the outcome measures used by institutions and are grouped by Carnegie classification.

Table 2 Outcome Measures by Carnegie Classification (n= 39 institutions)

Outcome measure	Research/Doctoral		Master's		Baccalaureate		Associate's	
	%	Rank	%	Rank	%	Rank	%	Rank
Faculty satisfaction with training	100	1	92	1	100	1	91	1
Faculty assessment of usefulness of training	92	2	92	1	100	1	82	2
Faculty assessment of relevance of training	85	3	92	1	100	1	82	2
Faculty developed skill or competency	77	4	58	3	66	2	82	2
Faculty willingness to recommend training to other faculty	69	5	58	3	66	2	55	4
Faculty assessment of clear, purposeful communication	62	6	75	2	66	2	55	4
Faculty assessment of improvement in teaching	54	7	50	4	66	2	58	3
Faculty assessment of changes to their face-to-face teaching	62	6	33	5	66	2	45	5
Student evaluations of faculty teaching	69	5	33	5	66	2	36	6
Faculty assessment of changes in attitude toward online learning	46	8	58	3	33	3	36	6
Faculty assessment of changes in perception of teacher's role	46	8	50	4	66	2	27	7
Students' course grades	23	11	25	6	33	3	36	6
Faculty assessment of explanation of instructional design principles	38	9	17	7	33	3	9	9
Cost of training	15	12	17	7	66	2	18	8
Faculty assessment of mentoring relationship	31	10	17	7	0	4	9	9
Students' grades for specific course assignments	23	11	33	5	33	3	18	8
Students' cumulative GPA	0	14	8	8	33	3	27	7
Faculty assessment of research behind training	8	13	0	9	33	3	0	10

The results in Table 2 seem to portray several differences in the usage of various outcome measures by Carnegie classification. Research/doctoral institutions seem less interested in assessing whether faculty developers assess participants' attitudes toward online learning (46%) and perception of the teacher's role (46%), both receiving a rank of 8. These measures seem to be more highly ranked by master's institutions (rank 3 and 4, respectively) and associate's institutions (rank 6 and 7, respectively). Use of student measures (course grades, assignment grades, cumulative GPA) is ranked lower for research/doctoral institutions (ranks of 11, 11, and 14) than other institutions (master's institutions rank these at 6, 5, and 8; associate's institutions rank these at 6, 8, 7), and surprisingly, research/doctoral institutions rank the use of an outcome measure capturing the research behind the training as the lowest item as do all other institution. This is consistent with the findings of Meyer and Murrell (2014a), in which the use of research about online learning during the development of training was reported by fewer than 25% of faculty developers.

Details on Faculty Evaluation of Trainings

Table 3 presents the results from the survey on the type of evaluations used by institutions, focusing both on the form of delivery and its timing.

Table 3 When and How Faculty Evaluate Training (n=38 institutions)

Evaluation type	Frequency	Percent of institutions	Rank
Online evaluations	30	79	1
After all training is completed	27	71	2
After each element of training is completed	15	40	3
After all training of a particular type is completed	14	37	4
Paper evaluation tool	13	34	5
Delayed evaluations (after passage of time)	11	29	6
Focus groups	8	21	7
One-on-one interviews	7	18	8

Several insights can be gleaned from these results. First, no one type of evaluation seems to be universal; even online evaluations, perhaps the least labor-intensive method, were used only by 80% of institutions. But this number is intriguing in comparison to the lesser-used paper evaluation tool (34%). Many institutions seem to have moved online to conduct their evaluations, which pairs logically with teaching faculty how to teach online.

Second, the majority (71%) of institutions continue to pursue summative evaluation (after all of the training is completed), although about one-third try to conduct evaluations in a more formative fashion by doing so after an element of training is completed (perhaps after a PowerPoint presentation was made or a lab exercise was done) or after all training of a particular type is completed (e.g., after all the PowerPoint presentations or all of the exercises were completed). Third, evaluations that take more time and resources to conduct seem to also be used by fewer institutions with 29% of institutions trying to evaluate training after a passage of time, 18% of institutions using one-on-one interviews, and 21% of institutions using focus groups. Given pressures on both budgets and staff, these lesser frequencies are understandable, but their loss means that in-depth, rich, and reflective evaluations may be foregone. Perhaps such evaluation tools or approaches can be done on occasion.

To ease comparisons between Table 3 and Table 4, the evaluation types in Table 3 are presented in the same rank order as in Table 4. Both present differences according to Carnegie classification.

Table 4 When/How Training Occurs by Carnegie Classification (n=38 institutions)

Training	Research/Doctoral		Master's		Baccalaureate		Associate's	
	%	Rank	%	Rank	%	Rank	%	Rank
Online evaluations	75	1	67	1	66	1	73	1
After all training is completed	75	1	67	1	66	1	73	1
After each element of training is completed	42	3	17	4	33	2	58	2
After all training of a particular type is completed	50	2	33	3	0	3	36	3
Paper evaluation tool	50	2	17	4	33	2	36	3
Delayed evaluations (after passage of time)	17	5	58	2	33	2	9	4
Focus groups	25	4	33	3	0	3	9	4
One-on-one interviews	8	6	33	3	33	2	0	5

Two things may be seen from Table 4. First, there is consistency among the preferred evaluation types and timings across the Carnegie classifications. Online evaluations and evaluations done after all training is completed were ranked number one by all institutional types. Second, paper evaluations are ranked higher (number two) among research/doctoral institutions than the other Carnegie types (master's rank of number four and associate's rank of number 3). Third, the time- and staff-intensive evaluation types consistently receive lower rankings, although master's institutions use delayed evaluations more frequently than the other Carnegie institutional types and over three times more frequently than research/doctoral institutions.

In fact, the master's institutions are more likely to pursue the last three evaluation types (delayed evaluations, focus groups, and one-on-one interviews). This finding bears further analysis. Are these institutions more committed to evaluation? Do they make more resources for evaluation available than at research/doctoral institutions? Answering these questions might help elucidate whether these are real and substantive differences among Carnegie types or an anomaly of this particular study.

RECOMMENDATIONS AND FUTURE RESEARCH

Based on the findings from this research, seven recommendations seem reasonable. First, it may be valuable for institutions to share their evaluation measures with other institutions. If the measures are too idiosyncratic (or applicable only to a single institution), then perhaps like institutions can coordinate on the development of outcome measures and tools. More informal associations could be formed than those noted in Daly and Dee (2009), but national memberships, like the Online Learning Consortium, might provide a data warehouse of proven evaluation measures. This would encourage institutions to use the best evaluation tools that are currently available and that are suitable for a particular type of institution. For time-pressed faculty developers, this need not be an onerous process; instruments that institutions use to collect evaluation information, especially including unique outcome measures, may be shared online or through an organization of like-minded individuals.

Second, it is understandable that the more time- and staff-intensive approaches to evaluation (one-on-one interviews or focus groups, as described in Daly and Dee, 2009; Lackey, 2011; and McQuiggan, 2012) are not a regular staple of faculty developers' evaluation plans. However, these evaluation approaches often produce insightful results that can help developers adjust what they do in the future. Can there be some attempt to either use these approaches less frequently—perhaps once a year— or to collect such data across institutions? It is difficult to recommend that institutions spend money when budgets are tight, but perhaps some modest investment in the kinds of evaluations that can inform developers on what is working and what is not is money well spent.

Third, while the evaluation types addressed in this study are a good attempt to disentangle treatments, more can be done to provide developers with specific information on what is working to help faculty learn to teach online. Certainly this will be difficult to do, but without such input, how will decisions be made when budgets are cut or the number of faculty to be trained increases? Again, this sharing could be accomplished using an organization of like-minded individuals.

Fourth, more attention needs to be paid to student outcomes and tying these to faculty development. This may require analysis of data from online courses, student assignments, and faculty training. Of course, this presumes that the data are available. However, as institutional budgets continue to be constrained and attention on student learning continues to grow, this is an area that will require greater attention on the part of faculty developers and their institutions.

Fifth, if faculty developers are unsure of their evaluation skills, there are good models to borrow from. Boulemetis and Dutwin (2005) is one such model that guides the tentative evaluator through a number of steps, including identifying suitable outcomes, measures that capture these outcomes, and designing approaches to collect those measures. This may be one way that faculty developers can improve the rigor of their evaluations so that decisions based on evaluation data can be made more confidently.

Sixth, the field may benefit from focusing on what does not work. For example, a group of faculty development professionals that come together to share stories of their failures might go a long way towards finding out the flaws in current assumptions or the differences among faculty that are not included in evaluations at the present. A centralized clearinghouse for this kind of feedback (perhaps this can be done so responses are anonymous) might help ensure the involvement of faculty and faculty developers alike.

The last issue about faculty differences leads us into a concern that is poorly addressed in the literature as well as in our own study. The near-universal lack of attention to faculty differences in the evaluations of faculty development programs seems to imply that developers and those who design and carry out evaluations may believe that faculty members learn in a homogenous fashion. Much of faculty development is packaged as a "one-size-fits-all" endeavor that reveals similar assumptions about how faculty persons think, feel, and learn. Perhaps the ways faculty persons learn are more diverse than our current faculty development models account for, and research-based learning theories (Meyer & Murrell, 2014b) could be better utilized in the development of faculty training.

The opportunities for future research are vast. We need research that explores the faculty person's different perceptions of their roles, student learning, and appropriate pedagogy for their disciplines. We also need to better understand the faculty person's learning preferences, as well as differences by other kinds of personal and professional variables such as rank, age, and gender. A better understanding of faculty may produce faculty development tailored to the individual's learning preferences and skill needs, resulting in optimal outcomes for the faculty person, the students, and the institution.

About the Authors

Dr. Katrina Meyer was professor of higher and adult education at The University of Memphis specializing in online learning and higher education. She is the author of Lessons Learned from Virtual Universities, a 2009 publication in the *New Directions in Higher Education* series, and *Cost-Efficiencies of Online Learning*, a 2006 publication of the ASHE Higher Education Report Series. For over three years, she was Director of Distance Learning and Technology for the University and Community College System of Nevada. Prior to this, she served over eight years as Associate Director of Academic Affairs for the Higher Education Coordinating Board in the state of Washington and was responsible for technology planning and policy related to online learning.

Dr. Vicki Murrell is on the faculty in the department of counseling, educational psychology and research at the University of Memphis. She previously served as the Assistant Dean for Online Education at the University of Memphis, providing administrative oversight, faculty development, and instructional guidance in the development of the UMOnline program, which offers more than 40 online degrees and certificates. Dr. Murrell spent over 20 years in educational administration focused on the integration of technology into curricula, at all levels of elementary, secondary, undergraduate, graduate, and professional education.

References

Allen, I.E., & Seaman, J. (2012). Digital faculty: Professors, teaching and technology. Babson Survey Research Group.

Atheny, S., & Hoffman, K.D. (2007). The Master Teacher Initiative: A framework for faculty development. *Marketing Education Review, 17*(3): 1-9.

Boulemetis, J. & Dutwin, P. *The ABCs of Evaluation*. San Francisco: Jossey-Bass, 2005.

Camblin, L.D., & Steger, J.A. (2000). Rethinking faculty development. *Higher Education*, *39*(1):1-18.

Daly, C.J., & Dee, J.R. (2009). Innovative models for organizing faculty development programs: Pedagogical reflexivity, student learning empathy, and faculty agency. Human Architecture: *Journal of the Sociology of Self-Knowledge, 7*(1): 1-22.

DiLorenzo, T.M., & Heppner, P.P. (1994). The role of an academic department in promoting faculty development: Recognizing diversity and leading to excellence. *Journal of Counseling & Development, 72*(5): 485-491.

Dittmar, E., & McCracken, H. (2012). Promoting continuous quality improvement in online education: The META model. *Journal of Asynchronous Learning Networks, 12*(1): 163-176.

Edwards, L., Freeman, L.M., Hodges, J., Johnston, R.A., Odell, M., & Sunal, C.S. (2001). Teaching science in higher education. *School Science and Mathematics, 101*(5): 246-257.

Fetters, M.L., & Duby, T.G. (2011). Faculty development: A stage model matched to blended learning maturation. *Journal of Asynchronous Learning Networks,15*(1): 77-85.

Grant, M.M. (2004). Learning to teach with the web: Factors influencing teacher education faculty. *Internet and Higher Education, 7*:329-341.

Gruppen, L.D., Frohna, A.Z., Anderson, R.M., & Lowe, K.D. (2003). Faculty development for educational leadership and scholarship. *Academic Medicine, 78*(2): 137-141.

Haviland, D., Shin, S., & Turley, S. (2010). Now I'm ready: The impact of a professional development initiative on faculty concerns with program assessment. *Innovative Higher Education, 35*(4): 261-275.

Hixon, E., Barczyk, C., Buckenmeyer, J., & Feldman, L. (2011). Mentoring university faculty to become high quality online educators: A program evaluation. *Online Journal of Distance Learning Administration, 14*(5), 2011. Retrieved from http://www.westga.edu/~distance/ojdla/winter144/hixon_Barczyk_Buckenmeyer_feldman144.html

Jones, S.J., & Meyer, K.A. (2012). The "virtual face" of distance learning at public colleges and universities: What do websites reveal about student support services? *Online Journal of Distance Learning Administration*, *15*(4). Retrieved http://www.westga.edu/~distance/ojdla/winter154/jones_meyer154.html

Kidney, G.W., & Frieden, S. (2004). When the cows come home: A proven path of professional development for faculty pursuing e-learning. *The Journal.* Retrieved from http://thejournal.com/Articles/2004/06/01/When-the-Cows-Come-Home-A-Proven-Path-of-Professional-Development-for-Faculty-Pursuing-ELearning.aspx

King, K.P. (2004). Both sides now: Examining transformative learning and professional development of educators. *Innovative Higher Education, 29*(2): 155-174.

Koehler, M.J., Mishra, P., Hershey, K., & Peruski, L. (2004). With a little help from your students: A new model for faculty development and online course design. *Journal of Technology and Teacher Education, 12*(1): 25-55.

Koepke, K., & O'Brien, A. (2012). Advancing pedagogy: Evidence for the role of online instructor training in improved pedagogical practices. *Journal of Asynchronous Learning Networks, 12*(1): 73-84.

Kucsera, J.V., & Svinicki, M. (2010). Rigorous evaluations of faculty development programs. *Journal of Faculty Development, 24*(2): 5-18.

Kukulska-Hulme, A. (2012). How should the higher education workforce adapt to advancements in technology for teaching and learning? *Internet and Higher Education*, *15*(4), 247-254.

Lackey, K. (2011). Faculty development: An analysis of current and effective training strategies for preparing faculty to teach online. *Online Journal of Distance Learning Administration, 14*(5). Retrieved from http://www.westga.edu/~distance/ojdle/winter144/lackey144.html

Lavoie, D., & Rosman, A.J. (2007). Using active student-centered learning-based instructional design to develop faculty and improve course design, delivery, and evaluation. *Issues in Accounting Education, 22*(1): 105-118.

Lindman, J.M., & Tahamont, M. (2006). Transforming selves, transforming courses: Faculty and staff development and the construction of interdisciplinary diversity courses. *Innovative Higher Education, 39*(4): 289-304.

Liu, O.L. (2012). Student evaluation of instruction: In the new paradigm of distance education. *Research in Higher Education, 53*: 471-486.

Longanecker, D.A. (2008). Mission differentiation vs. mission creep: Higher education's battle between creationism and evolution. Retrieved from http://www.wiche.edu/info/gwypf/dal_mission.pdf.

Lowenthal, P.R., & Thomas, D. (2010). Digital campfires: Innovations in helping faculty explore the online learning wilderness. *MERLOT Journal of Online Learning and Teaching, 6*(3). Retrieved from http://jolt.merlot.org/vol6no3/lowenthal_0910.htm

Maxwell, W., & Kazlauskas, E. (1992). Which faculty development methods really work in community colleges? *Community College Journal of Research and Practice, 16*(4): 351-360.

McQuiggan, C.A. (2012). Faculty development for online teaching as a catalyst for change. *Journal of Asynchronous Learning Networks, 12*(1), 27-62

Meyer, K.A. (2008a). The "virtual face" of institutions: What do home pages reveal about higher education? *Innovative Higher Education*, *33*(3), 141–157.

Meyer, K.A. (2008b). The "virtual face" of institutions: Why legislators and other outsiders view higher education as aloof. *The Internet and Higher Education, 11*(3-4), 178-185.

Meyer, K.A. (2014). An analysis of the research on faculty development for online teaching and identification of new directions. *Journal of Asynchronous Learning Networks, 17*(4), 93-112.

Meyer, K.A. & Murrell, V.S. (2014a). A national study of training content and activities for faculty development for online teaching. *Journal of Asynchronous Learning Networks*, *18*(1). Retrieved from http://olj.onlinelearningconsortium.org/index.php/jaln/article/view/389/80.

Meyer, K.A., & Murrell, V.S. (2014b). A national study of theories and their importance for faculty development for online teaching. *Online Journal of Distance Learning Administration,* 17, 2.

O'Meara, K.A. (2007). Striving for what? Exploring the pursuit of prestige. In Smart, J.D. (Ed.), Higher education: Handbook of theory and research, 22 (pp, 121-179). New York: Springer.

Online Learning Consortium. (n.d.) Quality framework narrative, the 5 pillars. Retrieved from http://onlinelearningconsortium.org/Quality_Framework_Narrative_5_pillars.

Orozcco, M., Fowkles, J.K., Jerzak, P., & Musgrove, A. (2012). Zero to sixty plus in 108 days: Launching a central elearning unit and its first faculty development program. *Journal of Asynchronous Learning Networks, 12*(1), 177-192.

Potter, R., & Meisels, G. (2005). Enhancing teacher preparation and improving faculty teaching skills: Lessons learned from implementing "Science That Matters" a standards based interdisciplinary science course sequence. *Journal of Science Education and Technology, 14*(2), 191-204.

Puzziferro, M., & Shelton, K. (2008). A model for developing high-quality online courses: Integrating a systems approach with learning theory. *Journal of Asynchronous Learning Networks, 12*(3-4), 119-136.

Ragan, L.C., Bigatel, P.M., Kennan, S.S., & Dillon, J.M. (2012). From research to practice: Towards the development of an integrated and comprehensive faculty development program. *Journal of Asynchronous Learning Networks, 16*(5), 71-86.

Reilly, J., Vandenhouten, C., & Gallagher-Lepak, S. (2012). Faculty development for e-learning: A multi-campus community of practice (COP) approach. *Journal of Asynchronous Learning Networks, 12*(1), 99-110.

Rogers, E.M. (2003). *Diffusion of Innovation.* New York: Free Press.

Schrum, L., Burbank, M.D., Engle, J., Chambers, J.A., & Glassett, K.F. (2005). Post-secondary educators' professional development: Investigation of an online approach to enhancing teaching and learning. *Internet and Higher Education, 8*, 279-289.

Skeff, K.M., Stratos, G.A., Mygdal, W., DeWitt, T.A., Manfred, L., Quirk, M., Roberts, K., Greenberg, L., & Bland, C.J. (1997). Faculty development. *Journal of General Internal Medicine, 12*, 56-63.

Southern Regional Education Board. (2009). Guidelines for professional development for online teachers. Southern Regional Education Board: Atlanta, GA. Retrieved from http://publications.sreb.org/2009/09T01_Guide_profdev_online_teach.pdf

Stehle, S., Spinath, B., & Kadmon, M. (2012). Measuring teaching effectiveness: Correspondence between students' evaluations of teaching and different measures of student learning. *Research in Higher Education, 53*, 888-904.

Steinert, Y., Mann, K., Centeno, A., Dolmans, D., Spencer, J., Gelula, M., & Prideaux, D. (2006). A systematic review of faculty development initiatives designed to improve teaching effectiveness in medical education. *Medical Teacher, 28*(6), 497-526.

Storandt, B.C., Lacher, A.P., & Dossin, L.C. (2012). Toward an understanding of what works in professional development for online instructors: The case of PBS Teacherline. *Journal of Asynchronous Learning Networks, 12*(1), 121-162.

Vaill, A.L., & Testori, P.A. (2012). Orientation, mentoring and ongoing support: A three-tiered approach to online faculty development. *Journal of Asynchronous Learning Networks, 12*(1), 111-120.

Wilson, J. L., & Meyer, K.A. (2009). Higher education websites: The "virtual face" of diversity. *Journal of Diversity in Higher Education*, *2*(2), 91-102.