



Received: 08 April 2016
Accepted: 16 October 2016
Published: 17 November 2016

*Corresponding author: Kenneth L. Rigler, Department of Applied Technology, College of Science, Technology and Mathematics, Fort Hays State University, 600 Park Street, Hays, KS 67601, USA
E-mail: klrigler@fhsu.edu

Reviewing editor:
Samuel Greiff, University of Luxembourg, Luxembourg

Additional information is available at the end of the article

EDUCATIONAL ASSESSMENT & EVALUATION | RESEARCH ARTICLE

The inclusion of self-assessment in merit evaluation

Kenneth L. Rigler^{1*}, Lorie Cook-Benjamin², Regi Wieland² and Carrie Tholstrup²

Abstract: The purpose of this survey study was to collect faculty perceptions toward changes made to the faculty merit evaluation process in a college of education at a state comprehensive university. The changes in the evaluation occurred over a two-year period, where a formative rubric and faculty self-assessment were incorporated into the merit instrument. The sampling frame for the study included the college of education faculty members at the university. The data for the study were collected in two different phases using a field-tested online survey that was created to collect the faculty perceptions of the newly developed instrument and process for completing the merit evaluation. The data collected from the surveys were analyzed using nonparametric tests and qualitative textual analysis. Findings from the study indicated favorable perceptions of the new merit evaluation and only a limited number of differences among the demographic groups.

Subjects: Higher Education; Higher Education Management; Education Policy

Keywords: merit evaluation; self-assessment; faculty review; faculty reflection

1. Introduction

Each year, post-secondary faculty are evaluated for purposes of merit, tenure, and promotion; and even though teaching, advising, and service typically absorb the majority of faculty time and resources, the primary factor considered during the review process is the quantity and quality of research publication (Adams, 2003; Diamond, 1999). There have been efforts, however, by scholars such as Boyer (1990) to make the faculty review process more equitable with the goals of the departments and universities (Diamond, 2004). As the higher education industry continues to become more complex, the faculty evaluation procedures need to advance further than typical single quantitative measures and incorporate multiple forms of data from all aspects of the professorate (Darwin, 2012; Henderson, 2009; Miller & Seldin, 2014).

ABOUT THE AUTHORS

Kenneth L. Rigler teaches instructional technology and technology and engineering education courses and has research interests in the areas of higher education leadership, organizational change, and technology and engineering education.

Lorie Cook-Benjamin teaches in the areas of curriculum, assessment, and diversity and serves as a reviewer for multiple journals with her research interest in the area of diversity.

Regi Wieland teaches in the areas of building and district leadership and has research interests in the areas of rural schools and alternative certifications.

Carrie Tholstrup is a graduate teaching assistant completing her education specialist degree and focuses in the area of education leadership.

PUBLIC INTEREST STATEMENT

The merit evaluation process has the potential to be a controversial procedure in the review of higher education faculty. This article describes the efforts by a group of faculty members to implement a new review strategy incorporating a faculty self-assessment instrument which was strongly supported by the college of education faculty. Feedback indicated that the new merit evaluation instrument allowed faculty to better articulate accomplishments and identify goals for future growth.

2. Background and purpose

The context of this study was a result of a charge by the university provost for all campus departments to review and modify their department criteria and evaluation procedures. The provost encouraged departments to consider the definition for scholarly activities and incorporate Boyer's (1990) explanation of scholarship which categorizes scholarship activities into (a) discovery, (b) integration, (c) application, (d) teaching, and (c) engagement. In response to the provost's charge, a faculty evaluation committee was formed and charged to develop a standardized process and instrument for all departments in the college education. Prior to this effort, each of the departments in the college used a separate system for faculty evaluation and merit that was summative in nature and detailed specific activities and expectations for faculty where points were awarded in each of the three categories of teaching, scholarship, and service.

The faculty evaluation committee began its work by exploring and discussing other methods of evaluation including a rubric-based assessment and formative self-assessment in order to assist faculty in reflecting and documenting activities that aligned with Boyer's model of scholarship. Following the review, the committee developed a college-wide template using a formative rubric that was aligned with the college's conceptual framework for accreditation and the merit, tenure, and promotion processes outlined in the university Memorandum of Agreement. A goal of the committee was to give faculty members the opportunity to provide a self-assessment rationale in each of the areas of teaching, scholarship, and service based upon Boyer's (1990) model. A formative rubric was field tested with selected faculty based on rank from each department, and after revisions from the field test, the formative rubric was piloted for all faculty in the college of education. During the pilot year, the committee worked with the support staff to incorporate and distribute the merit evaluation using an electronic content management system (CMS) that was also being used for accreditation. Due to feedback from faculty and administrators during the pilot year and limitations in the CMS, the committee converted the self-evaluation instrument into a PDF fillable form during the second year of the study.

The purpose of the survey study was to collect faculty perceptions toward changes made to the faculty merit evaluation process in a college of education at a state comprehensive university. The changes made to the faculty evaluation instrument and process occurred over a two-year period where both a formative rubric and faculty self-assessment were incorporated into the merit evaluation instrument. The study was designed to examine the following research questions:

1. What were the faculty perceptions towards the new faculty self-assessment instrument and process?
2. What were the differences in faculty perceptions towards the new faculty self-assessment process based on faculty classifications, years of experience, and length of time to complete the evaluation?
3. What were the differences in faculty perceptions toward the new faculty self-assessment instrument and process between the first-year CMS pilot and the second-year implementation with the PDF fillable form?

3. Review of literature

The faculty reward system, as Diamond (1999) pointed out, should be utilized to fairly recognize the strengths of individuals while meeting the needs of the department or college. This system typically examines faculty performance in the areas of teaching, research, and service (Adams, 2003). The weight each is given, however, varies (Boyer, 1990) and service and teaching have less impact on merit decisions than publication of research (Adams, 2003; Diamond, 1999). Differences among disciplines (Diamond, 1999) and among the varying types of educational institutions (Henderson, 2009) mean a single method of evaluation for promotion and tenure is not effective. Efforts are being made to make this process of recognizing merit more equitable and in line with goals of departments and universities (Diamond, 2004) due in part to the publication of Boyer's (1990) *Scholarship reconsidered: priorities of the professoriate*.

3.1. Influence of the Boyer model

Boyer's analysis of the 1989 Carnegie Foundation of the Advancement of Teaching's National Survey of Faculty revealed the merit process in tertiary education institutions failed to recognize all of the academic functions of their professors. The role of the professoriate, especially at comprehensive institutions, has evolved over the course of history in the United States (Boyer, 1990; Henderson, 2009; Youn & Price, 2009). This has necessitated changed expectations in the merit process (Miller & Seldin, 2014), which have not appropriately assessed all aspects of the merit process. According to Boyer (1990), "There is growing evidence that professors want, and need, better ways for the full range of their aspirations and commitments to be acknowledged. Faculty are expressing serious reservations about the enterprise to which they have committed their professional lives" (p. 75). In addition, Arreola (1995) asserted an evaluation system must provide information faculty deem relevant which can be accomplished, in part, by including faculty input in the construction of the system.

Since Boyer's work, university faculty have argued that other factors should be more heavily considered with respect to the current emphasis on research and publication in making tenure and promotion decisions (Braskamp & Ory, 1994; Henderson, 2009). Boyer (1990) broadened the concept of scholarship to include teaching. McKinney (2007) explained scholarly teachers approach teaching in the same way they do other scholarly activities. They "reflect on their teaching, use classroom assessment techniques, discuss teaching issues with colleagues, try new things, and read and apply the literature on teaching and learning in their discipline" (pp. 9–10). This scholarly approach to teaching leaves many concerned with how to properly assess it (Glassick, Huber, & Maeroff, 1997).

3.2. Assessing scholarly teaching

Boyer (1990) suggested gathering information for assessment from self-assessment, peer assessment, and student assessment; and the merits of multiple sources used in evaluation have been echoed by others (Arreola, 1995; Berk, 2005, 2009; Braskamp & Ory, 1994; Centra, 1979). Student evaluations of teaching, despite being viewed as faulty for the sole purpose of assessing teaching effectiveness (Darwin, 2012; Diamond, 2004; Hodges & Stanton, 2007; Paulsen, 2002), tend to be the most relied upon source for assessing teaching effectiveness (Miller & Seldin, 2014). Darwin (2012) suggested a qualitative approach to evaluation may be more appropriate in an era when higher education is rapidly growing in complexity. Research has demonstrated one such approach, self-evaluation, has steadily increased in use since 1973 (Miller & Seldin, 2014). In a 1978 study, Seldin (1980) found that over a third of deans surveyed always used self-evaluations. After surveying 410 deans of four-year liberal arts colleges, among the resources "always used" in evaluating faculty teaching performance, self-evaluations saw a nearly nine percent increase from 2000 to 2010 (Miller & Seldin, 2014).

In order for teaching to be viewed as scholarship, Glassick et al. (1997) suggested that it must be guided by six standards which they have observed in other scholarly work. The sixth standard, reflective critique, focuses on the self-evaluation component of assessment, recognizing that, when a scholar critiques his or her own work in order to improve, it reveals self-awareness and results in better scholarship in the future. It is this reflective critique that helps faculty determine best practices for the classroom and gain insight into their own teaching (LaPrade, Gilpatrick, & Perkins, 2014).

The nature of a self-assessment calls to question the reliability of such evaluation for administrative purposes (Centra, 1979). However, as a qualitative approach to the evaluation process, self-evaluation allows for triangulation and elevation of its value (Jick, 1979) when combined in the professional portfolio with more quantitative data, providing a "frame of reference or context" for the other items (Diamond, 2004, p. 34). It is this triangulation, Seldin (2004) noted, that allows decision-makers to develop an accurate picture of a faculty member's effectiveness.

3.3. Self-assessment

Centra (1979) explained, “in self-assessment teachers rate their effectiveness on a scaled form or provide brief written evaluations of their teaching performance” (p. 48). This form of evaluation, which requires faculty to judge and evaluate their own performance, can be conducted formally with checklists or surveys, while informal self-assessments take the form of strength and weakness inventories and collegial conversations sharing ideas and perceptions. Self-evaluation serves as a “logical first step” (Centra, Froh, Gray, & Lambert, 1987, p. 17) in faculty improving their teaching and allows faculty insight other sources cannot provide (Berk, 2005; Miller & Seldin, 2014). This form of evaluation, while not heavily relied upon for tenure and merit decisions (Arreola, 1995; Centra, 1979), does foster necessary self-reflection (Braskamp & Ory, 1994; Dochy, Segers, & Sluijsmans, 1999) which can shape instruction (Miller & Seldin, 2014).

3.4. Self-reflection

Self-reflection, as Bullock and Hawk (2001) explained, is accurately describing and analyzing what a teacher’s practice and resulting student achievement has been, then considering the implications of these while planning future teaching. Dewey (1933) emphasized the role of reflective thinking as a means of making meaning and understanding relationships; reflection allows teachers to develop theory which guides practice (Rodgers, 2002). Reflection develops naturally as a social construct allowing teachers to engage with colleagues and share what they learn from reflecting on practice (Hoffman-Kipp, Artiles, & López-Torres, 2003).

4. Methods

In this study, a survey method was used to collect faculty perceptions toward changes made to the faculty merit evaluation process. The survey provided a mechanism to collect participants’ demographics, attitudes, beliefs, and behaviors to study relationships between variables and changes over time. The changes made to the faculty evaluation instrument and process occurred over a two-year period, where both a formative rubric and faculty self-assessment were incorporated into the merit evaluation instrument. During the first year, the new instrument and process were piloted using a CMS and changes were made based on feedback from the college constituents. During the second year, the modified instrument was again implemented using a new distribution method (i.e. a fillable PDF form). Following each of the two distribution phases, an online survey was sent to the faculty with an invitation to participate in the study and provide their perceptions of the instrument and process for completing the merit evaluation. The quantitative survey method was an appropriate way to collect faculty perceptions over the two-year period (Babbie, 2009).

The survey was developed by the faculty evaluation committee and included three sections: (a) demographics, (b) Likert-type items relating to faculty perceptions, and (c) open-ended questions for feedback. The faculty demographics were included in order to analyze how classification, years of experience, and time to completion impacted perceptions. The survey was validated via a field test with an expert panel of three professionals in the college and reviewed by the university Institutional Research Board prior to distribution. Feedback from the faculty and administrators during the pilot year indicated an overall positive perception of the evaluation instrument but also technical difficulties and limitations with the CMS. During the second year, the merit evaluation instrument was modified to incorporate a greater emphasis on self-assessment and the completion process was changed from the CMS to a fillable PDF form. The survey instrument was modified to align with the changes and was again validated by the expert panel of three professionals in the college and reviewed by the university Institutional Research Board.

4.1. Participants

In this study, a purposive sampling technique was used with faculty members in a college of education who had participated in the new faculty merit evaluation procedures. The purposive sampling technique was appropriate for this study as it targeted a population with a specific experience and generalizability to a larger population was not a primary concern (Trochim & Donnelly, 2008). The sampling frame for the pilot year was 37 with 19 participants (51%) and the sampling frame for the

Table 1. Faculty classification

Classification	Pilot with CMS (N = 19)		Implementation with PDF (N = 15)	
	f	%	f	%
Non-tenure track	4	21	5	33
Tenure track	9	47	3	20
Tenured	6	32	7	47

Table 2. Years of experience

Experience	Pilot with CMS (N = 19)		Implementation with PDF (N = 15)	
	f	%	f	%
1-4 years	4	21	3	20
5-9 years	5	26	4	27
10-19 years	4	21	4	27
20+ years	6	32	4	27

implementation year was 39 with 15 participants (38%). The faculty classification and years of experience were collected and analyzed (see Tables 1 and 2), but age and gender were not collected in order to maintain the anonymity of the study.

4.2. Data collection and analysis

The data for this study was collected over a two-year period through an anonymous online survey with the purpose of collecting faculty perceptions toward a new evaluation instrument and the process of completing the process electronically. The survey was distributed in two phases, first, during the pilot year after the faculty completed and submitted the self-assessment using the electronic CMS, and one year later, during the implementation year after the faculty completed and submitted the self-evaluation using the electronic PDF fillable form. With both cycles, the data was collected using an online survey system and was available for four weeks.

The quantitative data were analyzed using descriptive statistics in order to provide an overview of the characteristics of the participants and their perceptions of the merit evaluation process. Medians and interquartile ranges were calculated on the Likert-type items to provide summaries of distribution of scores for the variables. Nonparametric statistics (i.e. Kruskal-Wallis ANOVA) were used for comparisons across the groups due to the ordinal nature of the Likert data collected and the small number of participants. The qualitative data from the open-ended questions were analyzed for patterns, coded based on similar themes, and utilized for the improvement of the merit evaluation instrument.

4.3. Limitations

The study results have limitations due to the nature of the study within one specific college at a university where a new method of merit evaluation was piloted, modified, and implemented. The primary purpose of the study was not to generalize results to a larger population but, instead to provide evaluation data following a change to the merit evaluation process within a college. Therefore, the study results are limited to the faculty members who participated in the study and not generalizable to a larger population.

4.4. Results

Nineteen faculty members participated in the first-year CMS survey and 15 faculty members participated in the second-year PDF survey. In the first-year CMS survey, the majority of respondents were tenure track faculty (47%); and in the second-year, the majority of respondents were tenured faculty

Table 3. Length of time to complete

Time	Pilot with CMS (N = 19)		Implementation with PDF (N = 15)	
	f	%	f	%
1-3 h	12	67	11	73
4-6 h	4	22	4	27
7-9 h	1	5	0	0
10+ h	1	5	0	0

(47%) (see Table 1). During both years, the number of years of experience of the faculty member participants were similar (see Table 2) and the majority of faculty completed the merit evaluation in three hours or less (see Table 3).

Kruskal-Wallis tests were used to compare distributions of self-rated satisfaction levels for variables relating to tenure, years of experience, and time of completion during both the first-year CMS pilot and second-year PDF implementation (research questions 1 and 2). For the first-year CMS pilot, the median satisfaction levels were highest among the non-tenure/temporary faculty (see Table 4) and faculty with 10–19 years of experience (see Table 5), but there were no significant differences across the groups. In the time for completion group, 11 participants (67%) completed the evaluation in three hours or less (see Table 6) and only two participants (17%) needed more than six hours for completion.

For the second year PDF implementation, the median satisfaction levels were again highest among the non-tenure/temporary faculty (see Table 7) and faculty with 10–19 years of experience (see Table 8) and all participants completed the evaluation in six hours or less (see Table 9). The reflection on growth in teaching ($p = .016$) for faculty classification and the process for completing the PDF ($p = .024$) in time to completion were the only variables with data sufficient to reject the null hypotheses across groups (i.e. $p < .05$).

In order to analyze the change in faculty perceptions from the first year to the second year relating to the evaluation instrument and process for completing the evaluation (research question 3), Wilcoxon-Mann-Whitney tests were conducted to compare the differences in the underlying

Table 4. Descriptive statistics and p values comparing faculty classification for CMS pilot

Item	p	Non-tenure/temporary (n = 4)	Tenure track (n = 9)	Tenured (n = 6)
		Mdn (IQR)	Mdn (IQR)	Mdn (IQR)
The self-evaluation allowed me to document my teaching activities	.473	4.5 (1)	4.0 (1)	4.0 (2)
The self-evaluation allowed me to document my scholarly activities	.494	4.5 (1)	4.0 (2)	4.0 (2)
The self-evaluation allowed me to document my service activities	.494	4.5 (1)	4.0 (2)	4.0 (2)
My overall perception of the evaluation instrument (not CMS)	.056	5.0 (1)	4.5 (1)	4.0 (1)
The process of completing the electronic self-evaluation in CMS	.196	5.0 (1)	4.0 (1)	4.0 (2)

Notes: N = 19, Mdn = median, IQR = interquartile range.

Table 5. Descriptive statistics and p values comparing faculty years of experience for CMS pilot

Item	p	1–4 years (n = 4)	5–9 years (n = 5)	10–19 years (n = 4)	20+ years (n = 6)
		Mdn (IQR)	Mdn (IQR)	Mdn (IQR)	Mdn (IQR)
The self-evaluation allowed me to document my teaching activities	.430	4.0 (2)	4.0 (2)	4.5 (1)	4.0 (1)
The self-evaluation allowed me to document my scholarly activities	.457	4.0 (2)	4.0 (2)	4.5 (1)	4.0 (2)
The self-evaluation allowed me to document my service activities	.457	4.0 (2)	4.0 (2)	4.5 (1)	4.0 (2)
My overall perception of the evaluation instrument (not CMS)	.950	4.0 (3)	4.0 (2)	4.0 (2)	4.0 (1)
The process of completing the electronic self-evaluation in CMS	.994	4.0 (2)	4.0 (2)	4.5 (2)	4.0 (1)

Notes: N = 19, Mdn = median, IQR = interquartile range.

Table 6. Descriptive statistics and p values comparing faculty time for completion for CMS pilot

Item	p	1–3 h (n = 12)	4–6 h (n = 4)	7–9 h (n = 1)	10+ hours (n = 1)
		Mdn (IQR)	Mdn (IQR)	Mdn (IQR)	Mdn (IQR)
The self-evaluation allowed me to document my teaching activities	.186	4.0 (1)	4.0 (1)	*	*
The self-evaluation allowed me to document my scholarly activities	.208	4.0 (1)	4.0 (1)	*	*
The self-evaluation allowed me to document my service activities	.208	4.0 (1)	4.0 (1)	*	*
My overall perception of the evaluation instrument (not CMS)	.151	4.0 (2)	5.0 (1)	*	*
The process of completing the electronic self-evaluation in CMS	.195	5.0 (1)	5.0 (1)	*	*

Notes: N = 18, Mdn = median, IQR = interquartile range.

*Insufficient data for calculation.

Table 7. Descriptive statistics and p values comparing faculty classification for PDF

Item	p	Non-tenure/ temporary (n = 5)	Tenure track (n = 3)	Tenured (n = 7)
		Mdn (IQR)	Mdn (IQR)	Mdn (IQR)
The self-reflection instrument allowed me to reflect on my growth in teaching activities from the past year	.016*	5.0 (1)	4.5 (**)	4.0 (1)
The self-reflection instrument allowed me to reflect on my growth in scholarly activities from the past year	.061	5.0 (1)	4.5 (**)	4.0 (0)
The self-reflection instrument allowed me to reflect on my growth in service activities from the past year	.063	5.0 (1)	4.5 (**)	4.0 (0)
The self-reflection instrument will impact changes I make in my future teaching activities	.083	5.0 (1)	3.5 (**)	4.0 (0)
The self-reflection instrument will impact changes I make in my future scholarly activities	.090	5.0 (1)	3.5 (**)	4.0 (1)
The self-reflection instrument will impact changes I make in my future service activities	.090	5.0 (1)	3.5 (**)	4.0 (1)
The process of completing the electronic self-reflection instrument using the PDF template	.588	4.0 (1)	3.5 (**)	3.5 (3)
My overall perception of utilizing the self-reflection instrument for the process of merit evaluation	.279	5.0 (1)	**	4.0 (1)

Notes: N = 15, Mdn = median, IQR = interquartile range.

*Significant at the p < .05 level.

**Insufficient data for calculation.

Table 8. Descriptive statistics and p values comparing faculty years of experience for PDF

Item	p	1–4 years (n = 3)	5–9 years (n = 4)	10–19 years (n = 4)	20+ years (n = 4)
		Mdn (IQR)	Mdn (IQR)	Mdn (IQR)	Mdn (IQR)
The self-reflection instrument allowed me to reflect on my growth in teaching activities from the past year	.549	5.0 (*)	4.0 (1)	4.0 (2)	4.0 (2)
The self-reflection instrument allowed me to reflect on my growth in scholarly activities from the past year	.535	5.0 (*)	4.0 (1)	4.5 (1)	4.0 (2)
The self-reflection instrument allowed me to reflect on my growth in service activities from the past year	.640	5.0 (*)	4.0 (1)	4.5 (1)	4.0 (1)
The self-reflection instrument will impact changes I make in my future teaching activities	.564	4.0 (*)	4.0 (1)	4.5 (2)	4.0 (1)
The self-reflection instrument will impact changes I make in my future scholarly activities	.318	4.0 (*)	4.0 (1)	4.5 (2)	3.5 (1)
The self-reflection instrument will impact changes I make in my future service activities	.318	4.0 (*)	4.0 (1)	4.5 (2)	3.5 (1)
The process of completing the electronic self-reflection instrument using the PDF template	.431	4.0 (*)	4.0 (2)	4.0 (2)	2.5 (3)
My overall perception of utilizing the self-reflection instrument for the process of merit evaluation	.257	4.0 (*)	4.5 (1)	4.5 (1)	4.0 (1)

Notes: N = 15, Mdn = median, IQR = interquartile range.

*Insufficient data for calculation.

Table 9. Descriptive statistics and *p* values comparing faculty time for completion for PDF

Item	<i>p</i>	1–3 h (n = 11)	4–6 h (n = 4)	7–9 h (n = 0)	10+ h (n = 0)
		Mdn (IQR)	Mdn (IQR)	Mdn (IQR)	Mdn (IQR)
The self-reflection instrument allowed me to reflect on my growth in teaching activities from the past year	.563	4.0 (1)	4.0 (2)	***	***
The self-reflection instrument allowed me to reflect on my growth in scholarly activities from the past year	.270	4.0 (1)	4.0 (2)	***	***
The self-reflection instrument allowed me to reflect on my growth in service activities from the past year	.490	4.0 (1)	4.0 (1)	***	***
The self-reflection instrument will impact changes I make in my future teaching activities	.551	4.0 (1)	***	***	***
The self-reflection instrument will impact changes I make in my future scholarly activities	.281	4.0 (1)	4.0 (1)	***	***
The self-reflection instrument will impact changes I make in my future service activities	.281	4.0 (1)	4.0 (1)	***	***
The process of completing the electronic self-reflection instrument using the PDF template	.024*	4.0 (1)	2.5 (3)	***	***
My overall perception of utilizing the self-reflection instrument for the process of merit evaluation	.050	4.0 (1)	4.0 (1)	***	***

Notes: *N* = 15, Mdn = median, IQR = interquartile range.

*Significant at the *p* < .05 level.

***No data for calculation.

Table 10. Descriptive statistics and *p* values comparing CMS pilot to PDF fillable form

Item	<i>p</i>	First year CMS (N = 19)	Second year PDF (N = 15)
		Mdn (IQR)	Mdn (IQR)
My overall perception of the evaluation instrument	.585	4.0 (2)	4.0 (1)
The process of completing the evaluation	.283	4.0 (1)	4.0 (2)

Notes: Mdn = median, IQR = interquartile range.

distributions from the first year to the second (see Table 10). In both instances, perception of the instrument from year one to year two ($z = -.546$, $p = .585$) and perception of the process in completing the evaluation from year one to year two ($z = -1.074$, $p = .283$), there were no significant differences indicated in the analysis (see Table 10).

The first-year CMS survey included one open-ended question relating to the overall effectiveness of the merit instrument. Eight (42%) of the participants provided textual responses that were coded into two themes: technical and form content. Four (50%) of the eight responses related to the technical limitations of completing the evaluation using the CMS. For example, one respondent noted, “make sure the chair can access the faculty qualifications” and another respondent noted, “an actual submission button where the form can no longer be changed.” Three (37.5%) of the eight responses related to the actual form content and recommended changes included, “I think it might be a good idea to have a comments section for the grade distributions, course evaluations, and advisory evaluations” and more “genuinely self-reflective open ended questions.”

The second-year PDF survey included three open-ended questions with the first question related to how the faculty perceived the process to differ from the first-year CMS evaluation to the second-year PDF evaluation. Thirteen (86.7%) of the 15 participants provided textual responses for the first question and 10 (77%) of the 13 responses were related to the user-friendliness of the PDF evaluation. For example, the participants noted that “entering the information was much easier than the [CMS],” that the PDF “form was easy to fill out and clear in terms of directions,” and “the template worked very well—changes were easy to make and it minimized the time completing the information.”

The second open-ended question asked the participants to identify how the instrument influenced their reflection on their accomplishments over the past year. Twelve (80%) participants provided textual responses to the second question and two primary themes emerged from the responses: identify accomplishments and goal setting. The participants identified that the evaluation provided a “summary document” which allowed the faculty to clearly see their accomplishments. One participant noted, “I reflect all the time ... but it does allow me to see it in print and talk about it with the chair which solidifies some aspects of my reflection.” Other comments noted how the evaluation “helped articulate strengths and weaknesses” and served as a “reminder each year that I need to continually improve in certain areas” and “helps me realize I have lots of room to grow.” Goal setting was the other common theme from the responses as participants identified, “It holds me accountable and provides an opportunity to reflect on last year goals and create new goals” and “I use the instrument as a goal setting piece to meet my statement of responsibilities.”

The third and final open-ended question is related to the improvement of the PDF merit instrument. Seven (46.7%) participants provided textual responses to the third question and all of the responses related to minor technical deficiencies in the PDF form such as the need for “additional boxes,” additional space in the reflective areas, and the need to identify minimum technology requirements to fill out the PDF form.

5. Discussion

The purpose of this quantitative survey study was to collect faculty perceptions toward changes made to the faculty merit evaluation process in a college of education at a state comprehensive university and the study was guided by three research questions: (a) overall faculty perceptions, (b) perception differences based on faculty classifications, years of experience, and length of time to complete, and (c) perception differences from year one to year two. For research question 1, the descriptive analysis indicated the faculty had favorable perceptions of the new self-assessment instrument and the process for completing the merit evaluation using both the CMS during the first year and the PDF fillable form during the second year. In alignment with evidence from Berk (2009) and Miller and Seldin (2014), the faculty reported value in self-assessment for the merit review process.

For research question 2, there were only two items during the second-year data, where differences were identified among the groups and for research question 3, there were no differences from the first year to the second year. The limited number of significant differences among the groups indicated a general consensus across the faculty participants and no particular demographic of faculty resisted the change. The median satisfaction levels were highest among the non-tenure/temporary faculty for both years suggesting the possibility that the faculty appreciated the opportunity to document, present, and reflect on their accomplishments from the previous year. Further research specific to the non-tenure/temporary faculty would be needed to test this hypothesis.

For research question 3, though there were no significant differences in the quantitative data from the first-year CMS evaluation to the second-year PDF evaluation ($p = .585$ and $.283$, see Table 10), the qualitative data from the open-ended questions indicated that the faculty participants found the second-year PDF evaluation to be much easier to complete and more reflective. These were positive comments for the research team as the driving motivation to transition away from the CMS and

implement the PDF was to increase the user-friendliness of the evaluation. Further, the team modified the instrument after the first year to place a greater emphasis on the reflective practices of the faculty members.

The overall positive perceptions of the new merit evaluation process may have partially resulted from the minimal time requirement to complete the evaluation and the alignment with the tenure and promotion processes at the university—two of the primary target goals of the committee for the new merit evaluation instrument. During the first year, only one participant indicated the evaluation required more than six hours to complete and during the second year, no participants reported needing more than six hours. Though there was an increase in the self-assessment narratives in the second-year PDF evaluation, the narratives closely aligned with content faculty would document and present as part of the tenure and/or promotion process at the university.

The results of the study are important for the academic community due to the increasing emphasis on educational accountability. As faculty are assessed, the evaluation procedures need to expand beyond simple quantitative measures and provide educators the opportunity to provide qualitative self-reflection and self-assessment components. As demonstrated in this study, the inclusion of self-reflection allowed the educators to better document their achievements and provides a better dialog for future self-growth. Though the results are limited to one college, the study provides a platform for further research for greater generalization to the larger population of educators.

6. Conclusion

The limitations in the study indicated the need for multiple areas of additional practice and research relating to the faculty merit evaluation process. First, the literature review demonstrated a lack in current research relating to the topic of faculty merit evaluation and self-assessment. Further, the literature was unclear on the distinction between self-evaluation, self-assessment, and self-reflection. Greater research is needed to clarify and operationalize each of these constructs. There is also a need for additional research relating to administrators' and supervisors' perceptions of the merit evaluation instruments and processes. Finally, additional research is needed within different colleges and universities (i.e. larger samples) to test the perceptions of other faculty members in order to generalize the findings to a larger population of faculty members.

The results of the study suggest that the faculty within the college examined were supportive of the inclusion of self-assessment in the merit evaluation process and no particular group of faculty were overly resistant to the change. The process of conducting the study identified a need within the college to provide orientation and training for new and returning faculty members on the annual merit evaluation process and the inclusion of the self-assessment in the evaluation. As other departments and colleges review their merit evaluation processes, it is a goal of the research team that the developed merit instrument, processes employed, and results of the study can be used by other scholars to help guide continuous improvement in the faculty merit evaluation process. The most recent version of the merit evaluation instrument may be viewed at the following link: <https://goo.gl/jUQo4Q>.

Funding

The authors received no direct funding for this research.

Author details

Kenneth L. Rigler¹
E-mail: klrigler@fhsu.edu
ORCID ID: <http://orcid.org/0000-0002-3089-9788>
Lorie Cook-Benjamin²
E-mail: lcookbenjamin@fhsu.edu
ORCID ID: <http://orcid.org/0000-0002-8860-3978>
Regi Wieland²
E-mail: rlwieland@fhsu.edu

Carrie Tholstrup²

E-mail: cbtholstrup@fhsu.edu

¹ Department of Applied Technology, College of Science, Technology and Mathematics, Fort Hays State University, 600 Park Street, Hays, KS 67601, USA.

² College of Education and Technology, Fort Hays State University, 600 Park Street, Hays, KS 67601, USA.

Citation information

Cite this article as: The inclusion of self-assessment in merit evaluation, Kenneth L. Rigler, Lorie Cook-Benjamin, Regi Wieland & Carrie Tholstrup, *Cogent Education* (2016), 3: 1250490.

References

- Adams, J. (2003). Assessing faculty performance for merit: An academic accomplishment index. *Journalism & Mass Communication Educator*, 58, 240–250.
- Arreola, R. A. (1995). *Developing a comprehensive faculty evaluation system: A handbook for college faculty and administrators on designing and operating a comprehensive faculty evaluation system*. Bolton, MA: Anker.
- Babbie, E. (2009). *The practice of social research*. Belmont, CA: Cengage.
- Berk, R. A. (2005). Survey of 12 strategies to measure teaching effectiveness. *International Journal of Teaching and Learning in Higher Education*, 17, 48–62. Retrieved from <http://www.isetl.org/ijtlhe/pdf/IJTLHE8.pdf>
- Berk, R. A. (2009). Using the 360° multisource feedback model to evaluate teaching and professionalism. *Medical Teacher*, 31, 1073–1080. doi:10.3109/01421490802572775
- Boyer, E. L. (1990). *Scholarship reconsidered: Priorities of the professoriate*. Princeton, NJ: Carnegie Foundation for the Advancement of Teaching.
- Braskamp, L. A., & Ory, J. C. (1994). *Assessing faculty work: Enhancing individual and institutional performance*. San Francisco, CA: Jossey-Bass.
- Bullock, A. A., & Hawk, P. P. (2001). *Developing a teacher portfolio: A guide for preservice and practicing teachers*. Upper Saddle River, NJ: Merrill Prentice Hall.
- Centra, J. (1979). *Determining faculty effectiveness: Assessing teaching, research, and service for personnel decisions and improvement*. San Francisco, CA: Jossey-Bass.
- Centra, J., Froh, R. C., Gray, P. J., & Lambert, L. M. (1987). *A guide to evaluating teaching for promotion and tenure*. Littleton, MA: Copley.
- Darwin, S. (2012). Moving beyond face value: Re-envisioning higher education evaluation as a generator of professional knowledge. *Assessment & Evaluation in Higher Education*, 37, 733–745. doi:10.1080/02602938.2011.565114
- Dewey, J. (1933). *How we think*. Buffalo, NY: Prometheus Books.
- Diamond, R. M. (1999). *Aligning faculty rewards with institutional mission: Statements, policies, and guidelines*. Bolton, MA: Anker.
- Diamond, R. M. (2004). *Preparing for promotion, tenure, and annual review: A faculty guide* (2nd ed.). Bolton, MA: Anker.
- Dochy, F., Segers, M., & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: A review. *Studies in Higher Education*, 24, 331–350. <http://dx.doi.org/10.1080/03075079912331379935>
- Glassick, C. E., Huber, M. T., & Maeroff, G. I. (1997). *Scholarship assessed: Evaluation of the professoriate*. San Francisco, CA: Jossey-Bass.
- Henderson, B. B. (2009). The work of the people's university. *Teacher-Scholar*, 1, 5–29.
- Hodges, L., & Stanton, K. (2007). Translating comments on student evaluations into the language of learning. *Innovative Higher Education*, 31, 279–286. doi:10.1007/s10755-006-9027-3
- Hoffman-Kipp, P., Artiles, A. J., & López-Torres, L. (2003). Beyond reflection: Teacher learning as praxis. *Theory into Practice*, 42, 248–254. doi:10.1207/s15430421tip4203_12
- Jick, T. D. (1979). Mixing qualitative and quantitative methods: Triangulation in action. *Administrative Science Quarterly*, 24, 602–611. doi:10.2307/2392366
- LaPrade, K., Gilpatrick, M., & Perkins, D. (2014). Impact of reflective practice on online teaching performance in higher education. *Journal of Online Learning & Teaching*, 10, 625–639.
- McKinney, K. (2007). *Enhancing learning through the scholarship of teaching and learning: The challenges and joys of juggling*. Bolton, MA: Anker.
- Miller, J. E., & Seldin, P. (2014). Changing practices in faculty evaluation. *Academe*, 100, 35–38.
- Paulsen, M. B. (2002). Evaluating teaching performance. *New Directions for Institutional Research*, 2002, 5–18. [http://dx.doi.org/10.1002/\(ISSN\)1536-075X](http://dx.doi.org/10.1002/(ISSN)1536-075X)
- Rodgers, C. (2002). Defining reflection: Another look at John Dewey and reflective thinking. *Teachers College Record*, 104, 842–866. <http://dx.doi.org/10.1111/tcre.2002.104.issue-4>
- Seldin, P. (1980). *Successful faculty evaluation programs: A practical guide to improve faculty performance and promotion/tenure decisions*. Crugers, NY: Coventry Press.
- Seldin, P. (2004). *The teaching portfolio: A practical guide to improved performance and promotion/tenure decisions* (3rd ed.). San Francisco, CA: Jossey-Bass.
- Trochim, W. M. K., & Donnelly, J. P. (2008). *The research methods knowledge base* (3rd ed.). Mason, OH: Cengage Learning.
- Youn, T. I. K., & Price, T. M. (2009). Learning from the experience of others: The evolution of faculty tenure and promotion rules in comprehensive institutions. *The Journal of Higher Education*, 80, 204–237. <http://dx.doi.org/10.1353/jhe.0.0041>



© 2016 The Author(s). This open access article is distributed under a Creative Commons Attribution (CC-BY) 4.0 license.

You are free to:

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made.

You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

No additional restrictions

You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

