Changes over Time in Faculty Attitudes, Confidence, and Understanding as Related to Program Assessment

Don Haviland & Steve Turley
California State University, Long Beach
Seon-Hi Shin
Korea Institute for Curriculum and Evaluation

Introduction

This article examines the sustained impact of a workshop series held to support faculty engagement in program assessment in the College of Education at California State University, Long Beach. The four-part series, held monthly during spring 2008, was intended to enhance faculty understanding of their role in assessment of student learning at the program level, build confidence in their ability to participate in program assessment, and nurture attitudes that such participation was worthwhile (Haviland, Shin, & Turley, 2010).

The workshops brought together faculty from across the college and facilitated work with colleagues within academic programs. Workshop participants collaborated to identify student learning outcomes (SLOs) at the program level, determine appropriate evidence to assess each SLO, create rubrics for signature assignments measuring SLOs, and explore ways to interpret and use student performance data for program

Don Haviland is an assistant professor in the College of Education at California State University, Long Beach. His e-mail address is dhavilan@csulb.edu. Steve Turley is associate dean of the College of Education at California State University, Long Beach. His e-mail address is hsturley@csulb.edu. Seon-Hi Shin is a psychometrician at the Korea Institute for Curriculum and Evaluation in Seoul, Korea. Her e-mail address is shinseonhi@gmail.com

improvement. Workshops were sequential around focused goals. For instance, learning how to write SLOs led to identifying ways of assessing outcomes that, in turn, led to exploring ways of using data in program improvement discussions. The same outside expert presented at each workshop, leading a one-hour discussion on the topic, followed by two hours for faculty members to collaborate to apply the knowledge and skills covered by the presenter.

In this article, we explore the long-term impact of this professional development series on faculty participants' attitudes, confidence, and understanding as related to program assessment. Haviland et al. (2010) noted that the workshops had an immediate impact on these characteristics and provided the foundation for a faculty-owned system of program assessment. In the current study, we employ a hierarchical linear model to examine whether the immediate impact was sustained and how it evolved over time.

We hope this research makes at least two contributions to the field of teacher education. First, we hope that the discussion of hierarchical linear modeling will be useful to others who wish to examine the effectiveness of their programs or services and have random and/or missing data with which to work. Second, in an era in which teacher education programs are increasingly pressured to adjust how they think about and undertake the preparation of future teachers, we believe that our research offers lessons for how administrative leaders can engage faculty in colleges of education in any change initiative, not solely one related to program assessment.

Faculty, Assessment, and the Workshop Series

Program assessment is a process that calls on faculty to work together to articulate programmatic learning outcomes, collect data on student performance, and review the aggregated data to inform program improvement efforts (Palomba & Banta, 1999). It is driven by two overlapping and sometimes competing philosophies. On the one hand, for effective teaching and learning to take place, colleges must have clear programmatic learning outcomes and ways for determining how students attain those outcomes. On the other hand, colleges and their faculty must be held accountable for ensuring that students receive a quality education (Ewell, 2002).

Faculty response to program assessment can be cautious, wary, or even negative. This is ironic, given that most faculty members assess student learning in their classrooms hundreds of times a semester. Yet resistance to program assessment is typically driven by an awareness

that it is part of an accountability agenda imposed from outside the school and that it represents new work in an already busy day as well as by the belief that it poses a threat to faculty autonomy, curricular control, and academic freedom (Ewell, 2002; Wehlburg, 2008; Welsh & Metcalf, 2003; Wergin, 2005).

Such tensions are often particularly acute in colleges of education holding or pursuing accreditation from the National Council for Accreditation of Teacher Education (NCATE). Since 2002, NCATE has required member institutions to implement a unit assessment system (UAS; NCATE, 2002). The foundation of UAS is that faculty and administrators collect and use data on candidate performance and unit operations to drive continuous improvement. Among the key attributes of such assessment systems is "fairness, consistency, accuracy, and avoidance of bias" (NCATE, 2008, p. 28) and that the data be aggregated around the elements of the college's conceptual framework.

However, as Bullough, Clark, and Patterson (2003) noted, the UAS model put forward by NCATE can conflict with certain program improvement perspectives (Powell, 2000) that may be more compelling to faculty. They argue that the NCATE push for colleges to establish uniform transition points and aggregate data across programs reflects an evaluative approach that is technocratic, mechanistic, and summative. This is particularly true, according to these authors, in large, complex colleges with myriad programs and a somewhat diffuse mission, where a uniform approach to assessment obscures the diversity and variety of an array of programs. The result can be the diminished role of faculty in the process and a compliance mentality often noted by accreditors (Wergin, 2005).

In contrast, faculty prefer a more collaborative, organic, and process-led approach (Bullough et al., 2003). In this model, faculty take ownership for a process of program self-assessment that nurtures continuous improvement through self-determination and initiative (Powell, 2000). The focus of assessment stays at the program level, the level at which faculty think and act most frequently, by providing data that are relevant and meaningful to the program. Indeed, there is evidence that faculty will embrace effectiveness activities such as assessment when these activities are focused on program improvement, rather than on external accreditation, and when faculty are seen as meaningfully involved in leading such activities (Welsh & Metcalf, 2003).

The College of Education at California State University, Long Beach confronted precisely these challenges. How, we asked, could we establish an assessment system that fulfilled the expectations of NCATE but would be embraced and led by the faculty for genuine program improvement? To work toward this goal, the college supported a workshop series, along

with other initiatives such as the creation of an assessment office, as part of its effort to embody Wehlburg's (2008) idea of transformative assessment. In this approach, assessment is practiced as a faculty-owned, faculty-driven endeavor rather than a top-down mandate, as well as one that respects the unique context and goals of individual programs, builds on existing work, and captures data that are meaningful to program faculty (Walvoord, 2004; Wehlburg, 2008).

In addition to building faculty knowledge and skills in regard to assessment, the workshops sought to shape attitudes, understanding, and confidence by recognizing that acceptance of any change is a social process (Berquist, 1992; Rogers, 2003; Sahin & Thompson, 2006). Therefore, participants were encouraged to exchange ideas, explore the feasibility of new practices, and establish a network or community focused on achieving a common goal—systematic program assessment—across multiple programs (Stigmar, 2008; Sullivan, Lakoma, Billings, Peters, & Block, 2006; Wolverton, Gmelch, & Sorenson, 1998). The workshop series reflected ongoing, coherent professional development that is regarded by researchers as more likely to influence practice than are episodic or isolated efforts (Boyle, Lamprianou, & Boyle, 2005; Murray, 1999).

The Study

Findings from a previous phase of this research showed that faculty appeared to have a better understanding of their role in the college assessment system, greater knowledge and skill in assessment-related areas, and more confidence in the value of carrying out assessment work immediately following the completion of the workshops (Haviland et al., 2010). Faculty attitudes toward program assessment immediately after the workshops, however, were ambiguous. While the findings indicated that faculty were more concerned than before the workshops that assessment might require changes in their teaching practice, there was also evidence of more positive attitudes as seen in faculty indicating a greater desire to work with colleagues on assessment (Haviland et al., 2010). Taken as a whole, the findings from the earlier study suggested that systematic professional development, together with visible administrative support, could positively shape faculty attitudes, understanding, and confidence related to program assessment and provide the foundation for a transformative system of assessment, at least in the short term.

The current study, therefore, examines whether the changes in faculty understanding of and attitudes toward assessment and confidence in their ability to do assessment work were retained over time. Data from pre- and post-workshop surveys and data from a follow-up survey

administered one year after the workshops provide the basis for this analysis. These data derive from repeated measures collected at three unevenly spaced points in time. In addition, the dataset is incomplete in that the majority of workshop participants volunteered to respond to only one or two of the three surveys. With this peculiar data structure, a conventional analysis of variance will not work. Therefore, we employed a hierarchical linear model (HLM) in our effort to ascertain whether the effects of the workshop series would be retained over time.

Thus, in addition to examining the retention of changes in faculty understanding, confidence, and attitudes related to assessment, this paper demonstrates how to formulate an HLM to fit such data and how to interpret analysis results. In particular, we examined two survey items: *It is not clear to me how the new system is better than what we currently have* and *I am concerned about assessment requiring changes in my teaching practice.* Both of these items showed statistically significant changes between the pre-workshop survey and the post-workshop survey (Haviland et al., 2010).

Methods

Data Sources

The assessment workshops were held in each of four consecutive months in spring 2008 in the College of Education at California State University, Long Beach, a public, urban, master's-granting university enrolling more than 35,000 students. The college serves approximately 3,000 students each year, has roughly 75 full-time faculty, and is composed of more than 20 academic programs. Five affiliated credential programs, whose faculty participated in the assessment workshops, are housed in other colleges on campus.

Data for the study were collected from surveys administered at three points in time. Pre- and post-workshop series surveys were administered before and after the series in spring 2008. The follow-up survey took place in April 2009. Two authors (Haviland and Turley) developed the surveys, using as a model the *Stages of Concern Questionnaire* created by George, Hall, and Stiegelbauer (2006). Three main categories of concerns were identified that might facilitate or impede faculty assessment work (*attitudes* regarding the value of assessment, *understanding* of assessment and the way the system would work in the college, and *confidence* in the ability to carry out assessment work). Each survey contained approximately 30-items, 18 of which dealt with faculty perceptions regarding assessment practice. Participants responded to these items on a 7-point Likert scale, with *0=This statement seems irrelevant* and

7 =This statement is very true of me at this time. All the surveys were administered online.

One approach to handling repeated measures similar to those in this study involves traditional univariate and multivariate procedures under the classical linear model framework. However, this approach requires that subjects have no missing observations. It also assumes that covariates are constant within a subject and that they have a certain pattern of covariance structure, such as Huynh-Feldt or Type H for univariate data or an unstructured matrix for multivariate data (Stevens, 1996). Our data were not of this type.

An alternative for analyzing repeated measures is to use HLM techniques. This approach allows data to be missing at random and covariates to vary within a subject. Various patterns of a covariance matrix for random error within a subject are also permitted (Bryk & Raudenbush, 1992; Verbeke & Molenberghs, 2000). These features are desired when unequally spaced repeated measures are analyzed.

The sample size for this study was relatively small. Thus, the ability of HLM to tolerate cases having missing responses at random permitted a test of statistical significance. Moreover, the intervals between different time points at which the repeated measures were obtained varied. The pre-workshop survey was administered three months before the administration of the post-workshop survey; the follow-up survey was approximately 12 months later. When repeated measures are unequally spaced in time, a matrix called SP(POW) is often used to properly represent the within-subject covariance structure (Littell, Milliken, Stroup, & Wolfinger, 1996). In this structure, the correlation between a pair of repeated measures is parameterized to be stronger as the distance in time between the two measures becomes smaller.

The surveys were anonymous, with respondents using a self-generated code to allow researchers to match individual responses. If participants responded to the survey item *It is not clear to me how the new system is better than what we currently have* (U6) or the item *I am concerned about assessment requiring changes in my teaching practice* (A2) at, minimally, two points in time, they were included in the analysis. The HLM adopted for this study was formulated such that the outcome value at the lowest level (Level 1) was a participant's rating score on survey item U6 or A2 at a certain point in time. The regression model was a simple polynomial of degree two at this level. That is, two covariates for different time effects were introduced: The regression coefficient of linear time effect (*Time*) indicated the linear trend of the attribute of interest, for instance, understanding, whereas the coefficient of quadratic time effect (*Time*Time*) determined acceleration of the linear trend. For

example, if the time effect was positive and the time-by-time effect was negative for the *understanding* item (U6), this suggested that faculty understanding of assessment significantly increased from the first survey to the second survey but decreased significantly one year later.

Therefore, the level-1 model for participant *j* at time point *i* can be written as

$$Y_{ij}=\pi_{0j} + \pi_{1j}\,(Time)_{ij} + \pi_{2j}\,(Time*Time)_{ij} + e_{ij}, \quad (1)$$

where π_{1j} is the model intercept and π_{1j} and π_{2j} are the regression coefficients of linear trend and acceleration of linear trend, respectively. The time variable was coded 0, 3, or 15 in the unit of months to reflect three survey times in chronological order. Consequently, the intercept is the rating score of participant $_j$ at the time of the first survey. Random error within a subject is denoted by e_{ij}. It was assumed that the random error was normally distributed with mean zero and had an SP(POW) covariance structure. The SP(POW) covariance structure involves two parameters: overall variance (σ^2) and autoregressive parameter (ρ) (Littell et al., 1996). Further, the two time effects were assumed not to vary between participants. Hence the Level-2 model is formulated as

$$\begin{aligned}\pi_{0j} &= \beta_{00}\\ \pi_{1j} &= \beta_{10}\\ \pi_{2j} &= \beta_{20},\end{aligned} \quad (2)$$

where β_{00} indicates the average rating score across participants at the time of the first survey. The variance between subjects is assumed to be negligible as compared to the variance within subjects in the repeated measures of the study. Thus, the random effect terms at this level were dropped. It is not unusual to drop random effects when the contribution of the between-subject effects is negligible in repeated measures experiments (Littell et al., 1996). As a result, the combined model becomes

$$Y_{ij} = \beta_{00} + \beta_{10}\,(Time)_{ij} + \beta_{20}\,(Time*Time)_{ij} + e_{ij}. \quad (3)$$

This combined model is the one utilized in this study. As shown in the model formulation above, the intercept β_{00} and regression coefficients β_{10} and β_{20} remain the same over time and across persons. In other words, the linear trend and the acceleration will be the same across participants, and the average rating score in the first survey will not vary across participants or over time. Equivalently, it means that there are no random effects at the person level (Level 2). All tests used the .05 level of significance.

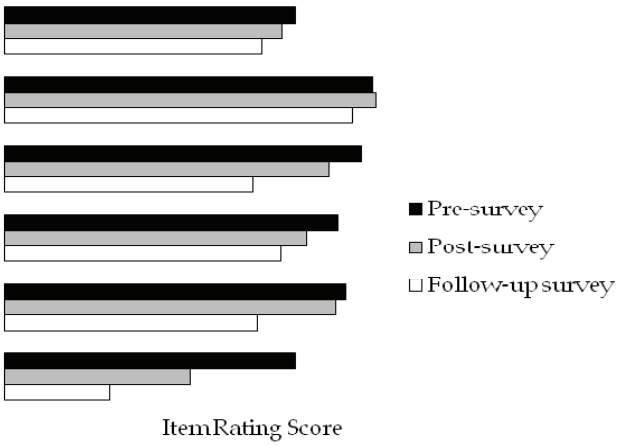
Results

A total of 36 (67%) participants completed the pre-workshop survey, and 24 (44%) completed the post-workshop survey. Further, 22 (41%) participants responded to the follow-up survey one year later. Using the respondent-created numeric identifiers, we were able to identify 18 participants for the *understanding* HLM and 20 for the *attitude* HLM.

Although the response rates to the second and third surveys were relatively low, the order of the contribution of each academic program to the total participants in all three surveys was unchanged over time. This suggests that the data attrition or missing data occurred in a random fashion, and there seemed no reason to believe that missing data happened in a systematic fashion. Therefore, identified participants were included in the HLM analysis insomuch as they responded to the item under investigation in at least two surveys. It is worth recalling that the HLM permits missing data at random without biasing the results.

By and large, there was a steady decline over time in the average rating scores of six survey items grouped as *understanding* (Figure 1), which asked respondents about their need for better understanding on a variety of dimensions. Faculty understanding of how the new system of program assessment would work, what resources would be available, and what roles they should play improved over time. Among the six items,

Figure 1
Average Self-Ratings of Faculty Understanding of Assessment Over Time



only *I am not clear how the new system is better than what we currently have* (item U6) showed a statistically significant decline between the pre- and post-workshop surveys (Haviland et al., 2010).

When the data from all three surveys were analyzed simultaneously, the HLM results showed that the decline was maintained for the interval from the pre-workshop survey through the follow-up survey. The coefficient of linear time effect (β_{10}) was significantly negative, suggesting that faculty self-reported need for understanding decreased over time, on average, continuing the increase in understanding noted in the pre- and post-workshop survey analysis (Haviland et al., 2010). The coefficient of quadratic time effect (β_{20}) was not significant, which indicated that the linear trend of the decrease was retained over time until the follow-up survey.

The covariance structure within a subject was assumed to be SP(POW) in the analysis. The two parameters determining the structure, o^2 and ρ, turned out significantly different from zero at the significance level of .01 according to the Wald *Z* tests. However, the Wald test is known to be unreliable with small samples. Accordingly, the likelihood ratio test, which is preferred for small samples, was conducted to compare two competing covariance structures: SP(POW) and Simple. The Simple structure is the simplest pattern and assumes an equal error variance and no correlations within a subject, thus having only one parameter, o^2. The results were not significant ($\chi 2 = 1.96$, $df = 1$, $p = .16$). This means that the Simple structure is preferred to the SP(POW) to fit the HLM model to understanding data. Table 1 contains the results of the HLM analysis under the assumption of the covariance structure as Simple.

There were no consistent trends over time in the attitude questions. Figure 2 shows that, on average, faculty members were slightly less interested in participating in assessment over time (item A4), whereas

Table 1
HLM Results from Understanding Model

Fixed Effect	*Coefficient*	*se*	*t ratio*	*p value*
Intercept, β_{00}	4.6875	0.4734	9.90	< .0001
Time, β_{10}	-0.6532	0.2694	-2.42	0.0226
Time*Time, β_{20}	0.0302	0.0168	1.80	0.0834
Variance Component	Estimate	*se*	*Z*	*p value*
Variance, $\acute{o}^2$	3.5862	0.7734	4.64	< .0001

they were more positive about working with colleagues to do assessment (item A3) than they were before the workshops. Haviland et al. (2010) reported that faculty concern about changes in teaching practice under the new system of program assessment increased significantly after the workshop series (item A2).

When data from the follow-up survey were added to the analysis, the increase in concern was not retained; moreover, the level of concern dropped significantly, to a point even below reported concern at the time of the pre-workshop survey. The HLM analysis for item A2 showed that

Figure 2
Average Self-Ratings of Faculty Attitudes toward Assessment over Time

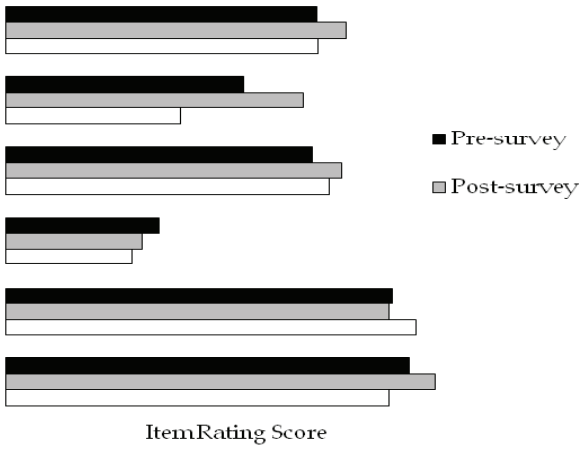


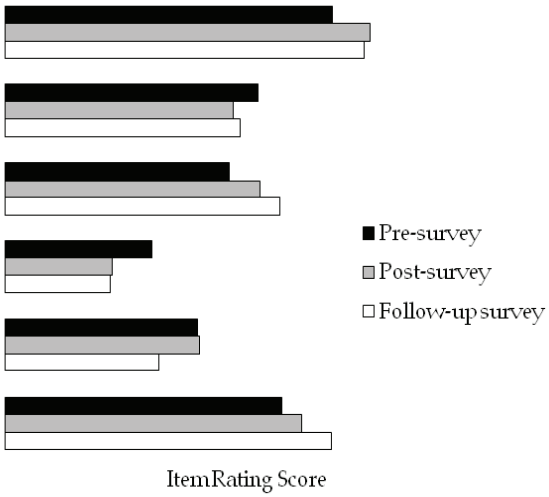
Table 2
HLM Results from Attitude Model

Fixed Effect	Coefficient	*se*	*t ratio*	*p value*
Intercept, β_{00}	2.9806	0.4078	7.31	< .0001
Time, β_{10}	0.3746	0.1506	2.49	0.0191
Time*Time, β_{20}	-0.0280	0.0096	-2.92	0.0069
Variance Component	Estimate	*se*	*Z*	*p value*
Variance, σ^2	3.1176	0.7118	4.38	< .0001
Autoregressive, ρ	0.8591	0.0579	14.85	< .0001

the coefficient of linear time effect was significantly positive but that the coefficient of quadratic time effect was significantly negative, meaning that the initial increase of the concern about changes was not maintained over time (Table 2). The increase in concern seen in the post-workshop survey became a decrease in concern by the time of the follow-up survey (Figure 2). The SP(POW) covariance structure turned out preferable for the attitude data, according to the likelihood ratio test ($\chi 2 = 7.35$, $df = 1$, $p = .01$).

All the average rating scores on confidence items at the time of the follow-up survey indicated that faculty members became more confident in the value of carrying out assessment work as compared to their confidence level prior to the workshops (Figure 3). Specifically, the average scores of items C3 (*I can help my colleagues figure out how to do assessment in their programs*) and C6 (*I think that we'll be able to make changes to our programs and practices once we have data*) steadily increased from the pre-workshop survey through the post-workshop survey to the follow-up survey. These findings indicate that faculty members developed confidence over time in helping colleagues figure out how to do program assessment and were increasingly confident that they would be able to make changes to their programs and practices when data were available. These improvements were mirrored by the fact that they became less

Figure 3
Average Self-Ratings of Faculty Confidence Regarding Assessment over Time



worried about developing an assessment plan (C4) over time, probably because plans were completed at the end of the workshop series.

Discussion

The research presented here shows that the positive impact of ongoing, focused professional development in program assessment on faculty understanding, confidence, and attitudes related to program assessment can be sustained and even improved over time. The professional development gave faculty participants opportunities to work together during a four-part workshop series to develop program assessment plans built around shared learning goals. Since that time, all programs have completed those plans, and all have developed and used rubrics to assess student attainment of program learning outcomes.

Faculty understanding of program assessment grew over time, both during the workshops and then in the year of implementation following the workshops. Moreover, as faculty engaged in collaborative program assessment activities and practices, they reported greater understanding and confidence in parallel. Indeed, earlier research (Haviland et al., 2010) supports the interpretation that the more that faculty understood expectations and roles regarding assessment, the more confident they felt about doing it. This study shows that faculty confidence remained even or became stronger one year after the workshops.

Attitudes also appear to have improved over time. Faculty were more concerned about assessment requiring changes in their teaching practice immediately after the workshops than before the experience, suggesting the workshops might have contributed to more negative attitudes (Haviland et al., 2010). However, this concern dropped significantly one year after the workshop series.

The results from this study align with literature indicating the importance of the social environment of professional development as a central element in changing practice (Stigmar, 2008; Sullivan et al., 2006; Wolverton et al., 1998). Moreover, they also reflect theory and research suggesting that the adoption of change and innovation is a social process, driven by interactions among colleagues (Berquist, 1992). Such interactions give individuals the opportunity to build knowledge of the innovation, explore its comparative advantages, and determine its suitability for their practices, values, and culture (Rogers, 2003). The assessment workshops, together with ongoing support and implementation, provided the opportunity for faculty in the college of education to do just that and to consequently augment early gains in understanding and confidence, while also relieving concerns (such as those about im-

pact on practice) that may have otherwise dampened attitudes toward assessment.

These findings suggest that giving faculty the opportunity and support to carry out assessment activities in the year following the workshops was important in shaping the improvements in understanding, confidence and attitudes. It seems likely that the chance to engage in ongoing assessment practice, along with better understanding of the assessment system, helped reduce faculty concern about change, built confidence in the value of doing assessment, and facilitated acceptance of the new system as their own.

In the year between the workshops and the follow-up survey, faculty had the opportunity to use the data that they collected for program improvement. Moving beyond the initial act of simple measurement to purposefully using data for program improvement can generate faculty buy-in (Palomba & Banta, 1999; Walker, Golde, Jones, Bueschel, & Hutchings, 2007; Wright, 2002), and such experience may have contributed to more positive attitudes about assessment in this instance as well. Spending time discussing assessment and developing skills for carrying out the work with colleagues, along with the availability of systematic organizational supports, appears to have given faculty greater clarity about their role and the resources that they could access for support.

Together with the findings from our previous study (Haviland et al., 2010), the results of this study also confirm the findings of others (Boyle et al., 2005; Murray, 1999; Sahin & Thompson, 2006) that professional development is meaningful when it takes place over time, focuses on a topic with clear and attainable goals for learning and growth, and integrates collaboration with colleagues for support. The growth in assessment skills and knowledge, improvements in understanding, and gains in confidence suggest that faculty participants benefited from receiving a manageable amount of information about assessment over an extended period. The workshops provided a means for conveying consistent messages on multiple occasions (Hall & Hord, 2001). It is worth recalling that, in interviews subsequent to the workshop series, participants commented on greater clarity and alignment regarding what the college would be doing with assessment, what the role of faculty members would be, and how the system would work (Haviland et al, 2010).

It is important to note that the findings of the study were based on a relatively small sample that had missing data. A repeated measure design allows fewer subjects than does a conventional linear model such as an analysis of variance. For instance, 15 subjects per group can be used in a repeated measure design (Stevens, 1996). However, the input data must be complete if they are to be analyzed within the conventional

model framework. The number of participants who completed all three surveys was less than 15 due to missing data on one or more occasions at random. Further, the time intervals across three surveys were not evenly spaced, requiring that the covariance structure among repeated measures or within a subject should be formulated properly during the estimation of model parameters. Thus, an HLM is better suited for analyzing such a data structure.

Conclusion

Program assessment is not the only initiative about which schools of education are being pressured by external stakeholders to adopt. As the Obama administration seeks to transform education in the United States, teacher preparation programs are coming under increasing and sustained pressure from alternative certification pathways (Foderaro, 2010) and confront calls to change or update their preparation practices, both generally and in a host of specific areas (Darling-Hammond, 2010; Garcia, Arias, Murri & Serna, 2010; Gay, 2010; Lieberman & Mace, 2010). While it is unclear how the response to these pressures will play out, it seems likely that further changes in philosophy and practice, much like program assessment, lie ahead for teacher education programs.

The very features that made the professional development series and ongoing implementation discussed here effective also hold the potential to help schools of education and teacher preparation programs manage the changes appearing on the horizon. Change forced upon an organization from the outside is difficult to manage and often met with a certain compliance mentality. However, educational leaders who support clearly defined and ongoing professional development, use it as a forum to promote faculty ownership of proposed changes, and present a clear and sustained message may find it easier to elicit faculty support and engagement in the process of change. Doing so may be a way to turn external pressures into internal opportunities for creativity and renewal that can transform teacher preparation.

References

Berquist, W. H. (1992). *The four cultures of the academy: Insights and strategies for improving leadership in collegiate organizations*. San Francisco: Jossey-Bass.

Boyle, B., Lamprianou, I., & Boyle, T. (2005). A longitudinal study of teacher change: What makes professional development effective? Report of the second year of the study. *School Effectiveness and School Improvement, 16*(1), 1-27.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models*. Newbury Park, CA: Sage.

Bullough, R. V., Clark, C., & Patterson, R. S. (2003). Getting in step: Accountability, accreditation and the standardization of teacher education in the United States. *Journal of Education for Teaching, 29*(1), 35-51.

Darling-Hammond, L. (2010). Teacher education and the American future. *Journal of Teacher Education, 61*(1-2), 35-47.

Ewell, P. T. (2002). Perpetual movement: Assessment after twenty years. Speech given to the American Association of Higher Education Annual Meeting, Boston, MA. Retrieved from www.teaglefoundation.org/learning/pdf/2002_ewell.pdf

Foderaro, L. W. (2010, April 18). Alternate path for teachers gains ground. *The New York Times*. Retrieved from http://www.nytimes.com

Garcia, E., Arias, M. B., Murri, N. J. H., & Serna, C. (2010). Developing responsive teachers: A challenge for a demographic reality. *Journal of Teacher Education, 61*(1-2), 132-142.

Gay, G. (2010). Acting on beliefs in teacher education for cultural diversity. *Journal of Teacher Education, 61*(1-2), 143-152.

George, A. A., Hall, G. E., & Stiegelbauer, S. M. (2006). *Measuring implementation in schools: The Stages of Concern questionnaire*. Austin, TX: Southwest Educational Development Laboratory/SEDL.

Hall, G. E., & Hord, S. M. (2001). *Implementing change: Patterns, principles, and potholes*. Needham Heights, MA: Allyn & Bacon.

Haviland, D., Shin, S., & Turley, S. (2010). Now I'm ready: The impact of a faculty development program on faculty concerns with program assessment. *Innovative Higher Education, 35*(4), 61-75.

Lieberman, A., & Mace, D. P. (2010). Making practice public: Teacher learning in the 21st century. *Journal of Teacher Education, 61*(1-2), 77-88.

Littell, R. C., Milliken, G. A., Stroup, W. W., & Wolfinger, R. D. (1996). *SAS system for mixed models*. Cary, NC: SAS Institute.

Murray, J. P. (1999). Faculty development in a national sample of community colleges. *Community College Review, 27*(3), 47-64.

National Council for Accreditation of Teacher Education (NCATE). (2002). *Professional standards for the accreditation of schools, colleges, and departments of education*. Washington, DC: Author

National Council for Accreditation of Teacher Education (NCATE). (2008). *Professional standards for the accreditation of teacher preparation institutions*. Washington, DC: Author.

Palomba, C. A., & Banta, T. W. (1999). *Assessment essentials: Planning, implementing, and improving assessment in higher education*. San Francisco: Jossey-Bass.

Powell, L. A. (2000). Realising the value of self-assessment: The influence of the business excellence model on teacher professionalism. *European Journal of Teacher Education, 23*(1), 37-48.

Rogers, E. M. (2003). *The diffusion of innovations* (4th ed.). New York: Free Press.

Sahin, I.. & Thompson, A. (2006). Using Rogers' theory to interpret instructional

computer use by COE faculty. *Journal of Research on Technology and Education, 39*(1), 81-104.

Stevens, J. (1996). *Applied multivariate statistics for the social sciences* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Stigmar, M. (2008). Faculty development through an educational action programme. *Higher Education Research and Development, 27*(2), 107-120.

Sullivan, A. M., Lakoma, M. D., Billings, J. A., Peters, A. S., & Block, S. D. (2006). Creating enduring change: Demonstrating the long-term impact of a faculty development program in palliative care. *Journal of General Internal Medicine, 21*(9), 907-914.

Verbeke, G., & Molenberghs G. (2000). *Springer series in statistics: Linear mixed models for longitudinal data*. New York: Springer-Verlag.

Walker, G. E., Golde, C. M., Jones, L., Bueschel, A. C., & Hutchings, P. (2007, December 14). The importance of intellectual community. *The Chronicle of Higher Education*, pp. B6-B8.

Walvoord, B. A. (2004). *Assessment clear and simple: A practice guide for institutions, departments, and general education*. San Francisco: Jossey-Bass.

Wehlburg, C. M. (2008). *Promoting integrative and transformative assessment: A deeper focus on student learning*. San Francisco: Jossey-Bass.

Welsh, J. F., & Metcalf, J. (2003). Faculty and administrative support for institutional effectiveness activities. *The Journal of Higher Education, 74*(4), 445-468.

Wergin, J. F. (2005). Taking responsibility for student learning: The role of accreditation. *Change: The Magazine of Higher Learning, 37*(1), 30-33.

Wolverton, M., Gmelch, W. H., & Sorenson, D. (1998). The department chair as double agent: The call for department change and renewal. *Innovative Higher Education, 22*(3), 203-215.

Wright, B. D. (2002). Accreditation and the scholarship of assessment. In T. W. Banta (Ed.), *Building a scholarship of assessment* (pp. 240-258). San Francisco: Jossey-Bass.