EVALUATION OF TEACHER DEVELOPMENT PROGRAMS: PARTICIPANT SATISFACTION AND RECOMMENDATION

Edwin King Por Wong and Alexander Seeshing Yeung

The Hong Kong Institute of Education, Hong Kong

To managers, course designers, developers and instructors of educational programs, course participants' feedback is important. The perceptions of course participants regarding the effectiveness of delivery of the program may lead the course manager, for example, to readjust the length of the program, the course designer to rethink the appropriateness of the course content, or the instructor to rearrange some of the learning activities (Decker, 1997; Fisher, 1998; Imrie, 1996; Ziehl, 1996). A commonly used approach to obtaining feedback from course participants is the survey method. Often, a survey comprises a series of Likert-type response items that are assumed to be asking course participants about their perceived effectiveness of course implementation (e.g., Marshall, 1973; Williford & Mayes, 2000). In the case of continuing education programs that emphasize life-long education, it is important to keep the length of the survey form to a minimum so that the participants are willing to complete all items. Thus it is essential to develop an evaluation instrument that is simple enough and yet has good construct validity. For the evaluation of programs of a self-funded nature, in addition to knowing the satisfaction level of the course participants, it is also important to know whether the course participants would recommend the course they have attended to their colleagues and friends. This study tests the dimensionality of a simple course evaluation instrument and examines the association between course participants' level of satisfaction and their likelihood of recommending the course to other people.
The Increasing Need for Continuing Education

Education worldwide has undergone varying degrees of reform. Since the issue of the Hong Kong Government's first Education Commission Reports in the early 1980s (e.g., Education Commission, Hong Kong, 1984), reform in the Hong Kong education sector has been vigorous and fast. Recent initiatives include a long list of items such as target-oriented curriculum, school management initiative, school-based management, the development of a general studies curriculum, the merging of special education with the main stream, the wide-spread introduction of a Putonghua (the official language in mainland China) curriculum in both primary and secondary schools after the handover of the British colony back to China in 1997, the use of information technology techniques in instruction, the mandatory submission of school development plans to the Education Department, and the formal implementation of staff appraisal system in schools (see Curriculum Development Council, Hong Kong, 2001; Curriculum Development Institute, Hong Kong, 1999, 2000; Education Commission, Hong Kong, 2002). The more recently proposed education reforms for the new millennium, which are even more innovative and comprehensive, have posed even greater challenges to frontline players such as teachers and school principals (Curriculum Development Council, Hong Kong, 2001; Education Commission, Hong Kong, 2000, 2002; UNESCO, 2000).

One of the most challenging concepts in recent reforms is continuing or life-long education (Curriculum Development Council, Hong Kong, 2001; Education Commission, Hong Kong, 2002). Whereas the emphasis of continuing education has led to the establishment of associate degrees and community colleges in the USA (Phillippe, 2000), for example, it has also led to increasing needs of the teaching profession for continuing staff development programs to update the knowledge and skills of the frontline players to better equip them for coping with the rapid changes (Association of Independent Colleges and Universities of Ohio, Columbus, 2001; Patton, 2001; Velde, 1997). In a Hong Kong context, with the encouragement from the Government of the Hong Kong Special Administration Region of China, the division in charge of continuing education programs in each of eleven tertiary education institutions joined to form a Federation for Continuing Education in Higher Education. Of the members of the Federation, the Hong Kong Institute of Education (HKIEd) specializes in organizing and implementing in-service programs for the teaching profession. The emphasis on these continuing education programs for teachers is also consistent with the institute's vision that "a culture of lifelong learning will become a reality" (HKIEd, 2000, p. 44).

For teachers and school principals to cope with the rapid education reform in Hong Kong, the Hong Kong Institute of Education has provided numerous staff development activities. Some of these activities aim at individual development and are run on the institute's premises while others are school-based and conducted with groups of teaching staff at their own schools. Like many tertiary education institutes that have served as main providers of continuing education to the community, the HKIEd established the Division of Continuing Professional Education (CPE) in 1997. Since then, the division has been fulfilling its mission of fostering and providing opportunities for lifelong and lifewide learning (see Curriculum Development Council, Hong Kong, 2001) in response to the needs and aspirations of those working in the field of education and training so as to enable
them to contribute wisely to social and educational changes. The operation of the CPE division is self-funded in nature. In 1999-2000, a total of around 18,000 teachers and trainers in various educational fields attended different types of training courses and activities organized by the division.

**Evaluation of Continuing Education Program**

Parallel to the need for more continuing education programs is the need for a valid evaluation of the delivery of such programs. Since the 1970s, the use of student evaluation as a formal means of teaching quality evaluation has been widespread, from institutions in the USA to many other parts of the world (Williford & Mayes, 2000). Students are asked to fill in evaluation questionnaires at the end of a series of lectures for summative assessment or in the middle of a course as an interim formative assessment. In most cases, the process is mandatory. Scores and written comments obtained from the evaluation questionnaires, usually anonymous, are used to judge the performance of the lecturers and instructors and the usefulness of the course. Whereas there is support for the appropriateness of this type of evaluation process for assessing the teaching staff, some findings from research have indicated that rating scores obtained from the survey questionnaires and the effectiveness of learning may be only weakly related (Lee, 2000). Nevertheless, participant surveys are still prevalent and generally accepted as an important process of teaching quality evaluation (Williford & Mayes, 2000).

For example, Gibbs, Habeshaw and Habeshaw (1997) put together a collection of practical methods for the appraisal of teaching including the use of questionnaires, interview discussion, video and audio recording and suggestions from peers. They found that participant surveys tend to be a popular and useful method of appraisal for course effectiveness. Probably due to their efficiency and ease of administration, many institutions use questionnaires to collect information on different aspects of their educational programs.

**Satisfaction vs. Recommendation**

Currently, providers of educational programs in Hong Kong often emphasize course evaluation and they mainly use survey instruments to solicit course participants' feedback on the quality of the courses. One of the major objectives of such an evaluation process is to find out how satisfied the course participants are with the educational program. However, to managers and designers of self-funded programs in continuing professional education in particular, it is not only the level of participant satisfaction that is the focus of interest. More important for the continuance of the programs is perhaps whether the participants would recommend the program to their colleagues and friends.

The present study examined the dimensionality of a simple evaluation instrument using six Likert-type items asking about participant satisfaction and one dichotomous item asking whether course participants would recommend the program to their friends and colleagues. Because the six items asked about different aspects of the program, it would be interesting to examine whether they would form a unidimensional scale or multiple dimensions. We first scrutinized the internal consistency and dimensionality of the six items that were assumed to probe a single Satisfaction factor. Then we examined the relationship
between course participants' satisfaction and their recommendation to other people. The study involved a large sample of school teachers and the findings would have important implications for the management of self-funded educational programs.

**Method**

**The Sample**

The participants were 9,101 teachers taking part in 334 classes of in-service teacher courses organized by the Division of Continuing Professional Education of the Hong Kong Institute of Education in 1999-2000. These courses might be categorized into three major types: (a) *commissioned activities programs* — for example, professional language upgrading, certificate courses on the teaching of arts and crafts and on the teaching of music in primary schools, information technology in teaching, and certificate course for discipline teachers in primary schools that are commissioned by the Education Department or other organizations; (b) *professional development programs* that cater for a wide range of personal interests — for example, curriculum development, special education practices, remedial teaching, language proficiency programs, first aid, statistics for teachers, musical instruments, etc., and (c) *school support service programs* that are tailor-made in response to the requests of schools — for example, team building in the school and the recently popular programs on the conduct of staff development camps.

The teachers attending these programs were mostly female (over 70%), representing roughly the ratio of the teacher population in Hong Kong. They cover all subject areas in the primary and secondary schools and are representative of the whole range of teaching experience (from 1 to over 40 years of teaching). After attending each course, the teachers responded to an evaluation instrument consisting of six items asking about their level of satisfaction on the delivery of the program and a single item on whether they would recommend the program to other teachers. Response to the evaluation instrument was absolutely voluntary and anonymity was emphasized. The analysis was based on responses of teachers who had complete data ($N = 8600$).

**Material**

There were a total of seven response items. Six items asked the teachers about their perceptions of the various aspects of program delivery. Teachers responded on a Likert-type scale from 1 = *very dissatisfied* to 6 = *very satisfied*. The six aspects were:

- Achievement of the course objectives.
- Usefulness of the programs to professional needs.
- Length of course/service.
- Instructor's preparation of information.
- Interaction between instructor and course participants.
- Learning materials/handouts used or provided.

A single item asked whether the teachers would recommend the course they had
attended to other teachers. This dichotomous item was coded such that 1 = no and 2 = yes.

Statistical Analyses

Preliminary Analysis

We first examined the alpha reliability of a single scale comprising six response items on various aspects of the programs. Then we conducted a principal components analysis to test the ability of the six items to form a single Satisfaction factor.

Confirmatory Factor Analysis

The critical concern of the confirmatory factor analysis was to examine the correlation between the Satisfaction factor derived from the six response items and the Recommendation factor from a single item. Because the Satisfaction responses were based on a 6-point scale and the Recommendation response was dichotomous, all items were treated as ordinal variables. The conduct of confirmatory factor analysis has been described elsewhere (e.g., Bentler, 1990; Bollen, 1989; Byrne, 1998; Joreskog & Sorbom, 1993; Marsh, Balla, & Hau, 1996; Marsh & Hocevar, 1985; Pedhazur & Schmelkin, 1991) and is not further detailed here. All analyses throughout this paper were conducted with the SPSS version of PRELIS and LISREL (Joreskog & Sorbom, 1988). The goodness of fit of models is evaluated on the basis of suggestions of Marsh, Balla, and McDonald (1988) and Marsh, Balla, and Hau (1996) with an emphasis on the Tucker-Lewis index (TLI). The CFA models were based on a 7 x 7 covariance matrix. Support for a strong relationship between teachers' perceptions of the program and their recommendation to other teachers would require a high correlation between the Satisfaction and Recommendation constructs.

Results

Preliminary Analysis

The alpha reliability estimates for the Satisfaction construct comprising six response items was good (α = .92). Principal components analysis of the six items with varimax rotation (SPSS, 1993) revealed a single factor with an Eigen value of 4.35 explaining 72.5% of the total variance. As expected, the six items formed a good measure of teacher perception of the in-service training programs with very good internal consistency.

Confirmatory Factor Analysis

Based on the principal components analysis, a single factor comprising six items was posited in the confirmatory factor analysis and a separate Recommendation factor based on a single dichotomous response item was included in the model to examine the relationship between the Satisfaction and the Recommendation constructs. The means, standard deviations, and correlations of the seven items are presented in Table 1. These
correlations show that the six items presumably pertaining to the Satisfaction construct were more closely related to each other (correlation coefficients ranging from .56 to .78.) than to the teachers' level of recommendation (correlation coefficients ranging from .30 to .35). This result provided preliminary support for a distinction between the Satisfaction and Recommendation constructs.

Table 1: Means, Standard Deviations, Covariances and Correlations of Variables

<table>
<thead>
<tr>
<th>Item 1</th>
<th>Item 2</th>
<th>Item 3</th>
<th>Item 4</th>
<th>Item 5</th>
<th>Item 6</th>
<th>Recommend</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>4.21</td>
<td>4.06</td>
<td>4.06</td>
<td>4.52</td>
<td>4.26</td>
<td>4.35</td>
</tr>
<tr>
<td>SD</td>
<td>1.07</td>
<td>1.07</td>
<td>1.17</td>
<td>1.12</td>
<td>1.13</td>
<td>1.09</td>
</tr>
</tbody>
</table>

Covariances (diagonal and below) & Spearman Correlations (above diagonal)

<table>
<thead>
<tr>
<th></th>
<th>Item 1</th>
<th>Item 2</th>
<th>Item 3</th>
<th>Item 4</th>
<th>Item 5</th>
<th>Item 6</th>
<th>Recommend</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1</td>
<td>.91</td>
<td>.78</td>
<td>.59</td>
<td>.70</td>
<td>.63</td>
<td>.67</td>
<td>.35</td>
</tr>
<tr>
<td>Item 2</td>
<td>.73</td>
<td>.92</td>
<td>.60</td>
<td>.66</td>
<td>.60</td>
<td>.64</td>
<td>.35</td>
</tr>
<tr>
<td>Item 3</td>
<td>.56</td>
<td>.57</td>
<td>.93</td>
<td>.59</td>
<td>.57</td>
<td>.56</td>
<td>.30</td>
</tr>
<tr>
<td>Item 4</td>
<td>.65</td>
<td>.61</td>
<td>.55</td>
<td>.91</td>
<td>.69</td>
<td>.77</td>
<td>.32</td>
</tr>
<tr>
<td>Item 5</td>
<td>.60</td>
<td>.58</td>
<td>.54</td>
<td>.65</td>
<td>.92</td>
<td>.67</td>
<td>.31</td>
</tr>
<tr>
<td>Item 6</td>
<td>.63</td>
<td>.61</td>
<td>.53</td>
<td>.72</td>
<td>.63</td>
<td>.91</td>
<td>.32</td>
</tr>
</tbody>
</table>

Note: N = 8,600

Table 2: Solution of Confirmatory Factor Analysis Model

<table>
<thead>
<tr>
<th>Variables</th>
<th>Satisfaction</th>
<th>Recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor coeff.</td>
<td>Unique</td>
<td>Factor coeff.</td>
</tr>
<tr>
<td>Item 1</td>
<td>.86*</td>
<td>.26*</td>
</tr>
<tr>
<td>Item 2</td>
<td>.83*</td>
<td>.31*</td>
</tr>
<tr>
<td>Item 3</td>
<td>.71*</td>
<td>.49*</td>
</tr>
<tr>
<td>Item 4</td>
<td>.86*</td>
<td>.26*</td>
</tr>
<tr>
<td>Item 5</td>
<td>.79*</td>
<td>.37*</td>
</tr>
<tr>
<td>Item 6</td>
<td>.84*</td>
<td>.30*</td>
</tr>
</tbody>
</table>

Factor Correlation

Satisfaction -
Recommendation .40*

Note: N = 8,600. Parameters estimates are completely standardized. *p < .05. The respective null model had a $\chi^2$ value of 38151.60 (21 df). Factor Coeff. = Factor coefficient. Unique = uniqueness.

The solution of the confirmatory factor analysis model is presented in Table 2. The model provided a good fit to the data ($\chi^2 = 2003.49$, $df = 14$, TLI = .92, RNI = .95). The respective null model had a $\chi^2 = 39151.60$, $df = 21$. All factor coefficients were large (from .71 to .86) and the item measurement uniquenesses were small (all < .5). This result provided good support for the six items to form a single Satisfaction factor. Inspection of the correlation between the Satisfaction and Recommendation constructs revealed that these two constructs
were only moderately correlated ($r = .40$). Together with Table 1, showing the correlations of individual items, the results show that teachers' favorable perceptions of an in-service training program may not be strongly related to their level of recommendation of the program to other teachers.

**Discussion**

The effective delivery of relevant training programs for teachers to cope with education reforms in Hong Kong may contribute significantly toward the successful implementation of the changes incurred (see Curriculum Development Council, Hong Kong, 2001; Curriculum Development Institute, Hong Kong, 1999, 2000; Education Commission, Hong Kong, 2002). However, because the mode of operation for the provision of such training courses is mainly self-funded, the dimension of financial management of such courses may impose considerable difficulties on their organization. That is why course providers are usually very concerned about both the quality and popularity of their training courses. Although classroom observation is generally considered a useful method for teacher supervision so as to assure teaching quality (Hazi, 1998), this can be an unduly expensive process to operate in self-funded training programs. No wonder the more efficient and cost-effective approach of collection of students' comments through the use of survey techniques is still very common (Williford & Mayes, 2000). Therefore, the development of a simple and valid instrument is important for obtaining participant feedback and imposing controlling measures.

The present study has shown that a simple survey instrument with six response items is capable of measuring a Satisfaction factor from course participants' perceptions. This six-item scale measuring Satisfaction has high internal consistency. Together with the dichotomous item for the measurement of course participants' inclination to recommend the course to other people, course providers are able to obtain useful information on the quality of the training courses and their popularity. For programs in the mode of self-funding operation, the latter is of particular importance. Thus for the evaluation of a self-funded program, both the Satisfaction and Recommendation constructs should be included. The simple and valid instrument described here seems to be useful for the purpose. It is cost-effective and efficient because it only requires two to three minutes to complete. The simplicity of design is also important because course participants usually do not like to spend much time filling in evaluation instruments at the close of the training programs.

An interesting finding that is also a focus of the present study is the low correlation between the Satisfaction factor and the Recommendation factor. Intuitively we might have expected a high correlation between the Satisfaction and Recommendation factors, as one would expect a course participant to recommend a course to others if he or she is satisfied with it. However, it seems that there may be some other factors affecting course participants' recommendation of a training course in addition to their level of satisfaction. This implies that participants' level of satisfaction may not guarantee their recommendation of the course to other people. This also implies an even more difficult job for providers of self-funded programs. Course providers most likely cannot rely too much on the degree of satisfaction of course participants merely in aspects of achievement of program objectives, usefulness of the programs to professional needs, appropriateness of
duration, instructors' preparation of information, interaction between instructors and program participants, and usefulness of learning materials used or provided, for the planning and provision of new training courses.

Further studies should be carried out to explain the observed phenomenon. In particular, we should explore why a course participant's level of satisfaction is not closely associated with an inclination to recommend it to other people. Interviews with course participants may shed light on the results found here. Also, because the present sample consisted of both primary and secondary school teachers and results were obtained from a wide range of training courses, we need to interpret the results with care and with reference to the context of the data collected. Future research could further investigate the multidimensionality of the various aspects of program evaluation. Although the present study has provided support for the unidimensionality of the Satisfaction factor, if each item is expanded to include multiple indicators representing a specific dimension, it may be possible to develop a more comprehensive multidimensional measure that is useful for the evaluation of specific dimensions of each program. However, this is beyond the scope of the present study which aims at developing a simple, easy-to-administer scale for evaluation purpose.

Further research should also study the differential perceptions of course participants in different course types and participants with various teaching experiences and curriculum areas. Gender differences will also be an interesting issue to investigate. In essence, multigroup structural equation modeling techniques should provide an even stronger validation of the scales developed here.

Another limitation of the study is the use of a single dichotomous item to infer a Recommendation factor. Thus even though the Satisfaction factor has demonstrated good reliability and strong construct validity, we recommend a stronger validation of the Recommendation factor by using multiple items for the construct in future research. When strong measures of both the Satisfaction and Recommendation constructs are established, a further differentiation between groupings may provide valuable information and guidance to course providers for the design and management of self-funded training courses to assist teachers to cope with changes resulted from education reforms in Hong Kong and elsewhere.

Acknowledgement

The authors thank Wally Y. W. Fung and Rachel Y. H. Hon for constructive comments on earlier versions of this paper.

References


The Authors

EDWIN K. P. WONG is the head of the Division of Continuing Professional Education of the Hong Kong Institute of Education in Hong Kong. Dr. Wong obtained his Ph.D. from the University of New South Wales, Australia. His expertise and research interests are in the fields of educational management and leadership and continuing education, particularly in the areas of quality assurance, staff development and strategic planning.

ALEXANDER S. YEUNG is a lecturer and senior program developer of the Division of Continuing Professional Education of the Hong Kong Institute of Education. His specializations include evaluation studies, self-concept and motivation, cognition and instruction, and research methodologies.

Correspondence: <kpwong@ied.edu.hk>