Available online at www.sciencedirect.com
ScienceDirect
Expert Systems with Applications 32 (2007) 1028–1044
Expert Systems with Applications
www.elsevier.com/locate/eswa

Evaluating intertwined effects in e-learning programs: A novel hybrid MCDM model based on factor analysis and DEMATEL

Gwo-Hshiung Tzeng [a,c], Cheng-Hsin Chiang [b], Chung-Wei Li [a,*]

[a] *Institute of Management of Technology, National Chiao Tung University, Hsinchu, Taiwan*
[b] *Applications and Services Division, National Taiwan University, Taipei, Taiwan*
[c] *College of Management, Kainan University, Taoyuan, Taiwan*

Abstract

Internet evolution has affected all industrial and commercial activity and accelerated e-learning growth. Due to cost, time, or flexibility for designer courses and learners, e-learning has been adopted by corporations as an alternative training method. E-learning effectiveness evaluation is vital, and evaluation criteria are diverse. A large effort has been made regarding e-learning effectiveness evaluation; however, a generalized quantitative evaluation model, which considers both the interaffected relation between criteria and the fuzziness of subjective perception concurrently, is lacking. In this paper, the proposed new novel hybrid MCDM model addresses the independent relations of evaluation criteria with the aid of factor analysis and the dependent relations of evaluation criteria with the aid of DEMATEL. The AHP and the fuzzy integral methods are used for synthetic utility in accordance with subjective perception environment. Empirical experimental results show the proposed model is capable of producing effective evaluation of e-learning programs with adequate criteria that fit with respondent's perception patterns, especially when the evaluation criteria are numerous and intertwined.

Keywords: E-learning; Factor analysis; Fuzzy integral; DEMATEL; Multiple criteria decision making (MCDM)

1. Introduction

Internet has significantly impacted the establishment of Internet-based education, or e-learning. Internet technology evolution and e-business has affected all industrial and commercial activity and accelerated e-learning industry growth. It has also fostered the collaboration of education and Internet technology by increasing the volume and speed of information transfer and simplifying knowledge management and exchange tasks. E-learning could become an alternative way to deliver on-the-job training for many companies, saving money, employee transportation time, and other expenditures. An e-learning platform is an emerging tool for corporate training, with many companies developing their own e-learning courses for employee on-the-job training. Employees can acquire competences and problem solving abilities via Internet learning for benefits among business enterprises, employees, and societies while at work.

Although e-learning has been developing for several years, evaluating e-learning effectiveness is critical as to whether companies will adopt e-learning systems. A considerable number of studies have been conducted emphasizing the factors to be considered for effectiveness evaluation. Several evaluation models are considered with specific aspects. The criteria used for e-learning effectiveness evaluation are numerous and influence one another.

The evaluation models however, are deficient and do not have an evaluation guideline. Effectiveness evaluation criteria must integrate learning theories, relative website design, course design, and learning satisfaction theories to form an integrated evaluation model (Allen, Russell, Pottet, & Dobbins, 1999; Hall & Nania, 1997; Hsieh, 2004). Since

* Corresponding author. Address: 7th Floor, Assembly Building 1, 1001 Ta-Hsueh Road, Hsinchu, Taiwan. Tel.: +886 3 5712121x57505; fax: +886 3 5753926.
E-mail address: samli.mt91g@nctu.edu.tw (C.-W. Li).

0957-4174/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2006.02.004

e-learning can be evaluated according to different aspects and criteria, the multi-criteria decision making (MCDM) approach is suitable for e-learning evaluation.

The purpose of this paper is to establish a new e-learning evaluation model for e-learning program effectiveness with consideration of intertwined relations and synthetic utility between criteria. Based on several evaluation criteria considered for e-learning effectiveness, this paper used several methods to establish the evaluation model. Factor analysis figures the main aspects of e-learning evaluation and generates independent factors/aspects for further evaluation using the AHP method. Criteria interrelations, and components of independent factors are usually intertwined and interaffected. Applying the DEMATEL (Decision Making Trial and Evaluation Laboratory) method (Fontela & Gabus, 1974, 1976; Warfield, 1976) illustrates the interrelations among criteria, finds the central criteria to represent the effectiveness of factors/aspects, and avoids the ''overfitting'' for evaluation. Thus, non-additive methods, fuzzy measure, and fuzzy integral, are used to calculate the dependent criteria weights and the satisfaction value of each factor/aspect for fitting with the patterns of human perception. Finally, the analytic hierarchy process (AHP) method is employed to find out the weights of factors/ aspects and obtain each e-learning program score.

The empirical experiments of this paper are demonstrated with two e-learning company-training programs. The proposed model could be used to evaluate effectiveness by considering the fuzziness of subjective perception, finding the central criteria for evaluating, illustrating criteria interrelations, and finding elements to improve the effectiveness of e-learning programs. Moreover, the results show that the effectiveness calculated by the proposed model is consistent with that from traditional additive methods.

The remainder of this paper is organized as follows. E-learning concepts, including definitions, categories, characteristics, evaluation criteria, and evaluation effectiveness models, are described in Section 2. In Section 3, a brief introduction of factor analysis, the DEMATEL method, fuzzy measure, fuzzy integral, and AHP method is given. Establishing a model using these methods is also proposed. In Section 4, empirical experiments of two real e-learning cases (Masterlink Securities Corporation training programs) are shown using the proposed evaluation model. The analysis result is discussed and compared with the traditional additive evaluation model in Section 5. Section 6 concludes the paper.

2. Environments and the effectiveness evaluation models of e-learning

E-learning combines education functions into electronic form and provides instruction courses via information technology and Internet in e-Era. The most popular definition of e-learning as defined by the American Society for Training and Development (ASTD) is a wide set of applications and processes, such as web-based learning, computer-based learning, virtual classrooms, and digital collaboration. E-learning is not an innovative education idea, since computer-aided training (CAT), computer-based training (CBT), and distance learning have been used as elements of e-learning for more than ten years. Research shows that students can be effective learners over the web, and learn as much, if not more, than in traditional courses.

E-learning is currently a burgeoning educational and training tool because of its cost saving advantages, institution reusability, and learner flexibility. World governments emphasize e-learning for social and public education, and want to enlarge it as a branch of education. The European Union in 2000, proposed the eEurope project, promoting an information society for all (Europa, 2004). Moreover, the Japanese government has proposed the eJapan project, making e-learning one of seven main application development items. E-learning has also been used with university and enterprise education. Enterprises can introduce e-learning courses and systems into the firm, which can then be used by the human resources or research development department to do on-the-job training. When companies induce e-learning courses into their organization, they can save money otherwise used for guest lecturers, and employees can learn on demand.

Each e-learning procedure, from course design to learner response or behavior measurement, will affect course performance. According to previous research, instructional system design process models are process-oriented rather than product-oriented and include built-in evaluation and revision systems (Hannum & Hansen, 1989). Systematic instructional system designs follow five learner need stages: (1) analysis, (2) design, (3) development, (4) implementation, and (5) evaluation, or the ADDIE acronym model (Hegstad & Wentligm, 2004). The ADDIE is usually used in mentoring as an intervention that can be linked to three primary functions: (1) organization, (2) training and development, and (3) career development (Mhod, Rina, & Suraya, 2004).

The basic reason for e-learning evaluation is to find out the effectiveness, efficiency, or appropriateness of a particular course of action. E-learning effectiveness evaluation intends to highlight good or bad practice, detect error and correct mistakes, assess risk, enable optimum investment to be achieved, and allow individuals and organizations to learn (Roffe, 2002). Evaluation can be most effective when it informs future decisions (Geis & Smith, 1992) and is better used to understand events and processes for future actions, whereas accountability looks back and properly assigns praise or blame.

Over the past few years, considerable studies have been undertaken primarily to find the dimensions or factors to be considered in evaluation effectiveness, however, with a specific perspective. Kirkpatrick proposed four levels of training evaluation criteria: (1) reactions, (2) learning, (3) behavior, and (4) results (Kirkpatrick, 1959a, 1959b, 1960a, 1960b). Garavaglia (1993) proposed five dimensions to evaluate e-learner change: (1) supervisory report, (2) on-the-job peer surveys, (3) action plan reports, (4) observation, and (5) self-report. Among these five methods, the observation method can avoid the possible bias a supervisor may have when reporting on a subordinate. The self-report method involves either interviews or surveys distributed or conducted two to three months after the learning session. Philips (1996) formed a logical framework to view ROI (return on investment) both from a human performance and business performance perspective. Urdan (2000) proposed four measure indicators, learner focused measures, performance focused measures, culture focused measures, and cost-return measures, to evaluate corporate e-learning effectiveness. Since web-based instruction has become the most engaging type for learning, four factors that affect the e-learning environment should also be identified: (1) efficacy studies, (2) technological advances, (3) pressures of competition and cost containment, and (4) professional responses to market influences (Miller & Miller, 2000).

Formative evaluation and summative evaluation are two common methods for evaluating e-learning course effectiveness in recent decades. Formative evaluation is used at the onset of new instructional program implementation to assess the needs and learning goals of an organization, or for program evaluation following training to revise existing programs. Several familiar formative evaluation models prescribe a four-part evaluation procedure employing expert reviews, one-to-one evaluations, small group evaluation, and field trials (Dick & Carey, 1996). Formative evaluation is typically categorized according to different processes such as design-based, expert-based, and learner-based for assessment, although.

Summative evaluation, one of the most popular methods focused on outcomes and used in classroom education. For example, the CIRO (contents/contexts, inputs, reactions and outcomes) model which measures learning/training effectiveness by CIRO elements, both before and after training, is currently widely used in business (Cooper, 1994). The strength of the CIRO model is consideration of objectives (contexts) and training equipment (inputs). The main emphasis of CIRO is measuring managerial training program effectiveness, but it does not indicate how measurement takes place. Adopting measures during training provides the training provider with important information regarding the current training situation, leading to improvements (Charles, Mahithorn, & Paul, 2002). Summative evaluation models lack consideration of other factors, such as individual characteristics, e-learning interface design, instructional system design, and course design, which may influence e-learning effectiveness.

Most evaluation models however, do not measure e-learning effectiveness from an overall perspective and ignore the interrelation among criteria. Most evaluation models concentrate on finding factors, aspects, or casual relationships between them. Quantitative study models mainly use traditional statistic methods or linear models (e.g. ANOVA, factor analysis and structural equation model) to find learner satisfaction or dissatisfaction via questionnaires or facial communications (Marks, Sibley, & Arbaugh, 2005; Moore, 1989; Muilenburg & Berge, 2005; Ng & Murphy, 2005; Sherry, Fulford, & Zhang, 1998). Typically, e-learning program effectiveness is evaluated by multiple intertwined and interaffected criteria, and the perceptions of utility for learners are not monotonic. Establishing a model to evaluate all available criteria and to determine central criteria, learner utility perception about these criteria, and the future improvement direction for the programs is necessary.

3. Evaluation structure model combined factor analysis and the DEMATEL method for determining the criteria weights

In this section, the concepts of establishing the evaluation structure model, combined factor analysis, and the DEMATEL method for determining the criteria weights, are introduced. In real evaluation problems, it is difficult to quantify a precise value in a complex evaluation system. However, the complex evaluation environment can be divided into many criteria or subsystems to more easily judge differences or measure scores of the divided criteria groups or subsystems. The factor analysis method is commonly used to divide criteria into groups. Although it seems logical to sum the scores of these criteria for calculating factor effectiveness, the weights between the criteria may differ and the criteria may have interdependent relationships. Assuming that criteria weights are equal may distort the results. In the proposed model, DEMATEL, fuzzy measure, and fuzzy integral are used to overcome these problems. DEMATEL is used to construct the interrelations between criteria, while fuzzy measure and fuzzy integral are used to calculate the weights and synthetic utility of the criteria. Factor weights can then be obtained via processing individual or group subjective perception by the AHP method. Then, the final effectiveness value can be obtained.

The hybrid MCDM model procedures are shown briefly in Fig. 1. Factor analysis, the DEMATEL method, fuzzy measure, fuzzy integral, AHP method, and the goals for combining these methods to evaluate e-learning effectiveness will be explained as follows.

3.1. Finding independent factors for building a hierarchical system

Based on various points of view or the suitable measuring method, the criteria can be categorized into distinct aspects. In real program problem assessment based on a general problem statement, various opinions from participants and the evaluation criteria will be setup. When the evaluation criteria in real complex problems are too large to determine the dependent or independent relation with others, using factor analysis can verify independent factors.

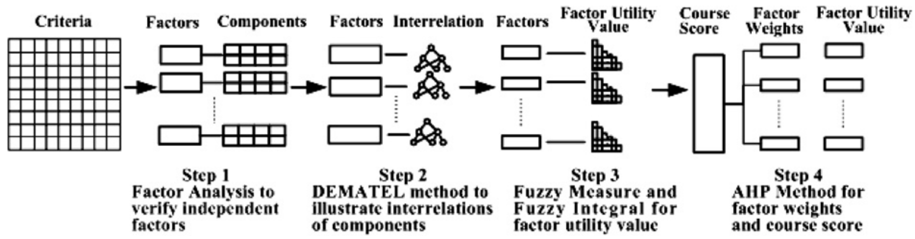

Fig. 1. Hybrid MCDM model procedures.

Another reason for using factor analysis in this paper is the conventional AHP method, which performs the final evaluation in an additive type, based on the assumption of independence among criteria within the evaluating structure systems.

Factor analysis is a dimension reduction method of multi-variate statistics, which explores the latent variables from manifest variables. Two methods for factor analysis are generally in use, principal component analysis, and the maximum likelihood method. The main procedure of principal component analysis can be described in the following steps when applying factor analysis:

Step 1: Find the correlation matrix (***R***) or variance–covariance matrix for the objects to be assessed.

Step 2: Find the eigenvalues (λ_k, $k = 1,2,\ldots,m$) and eigenvectors ($\beta_k = [\beta_{1k},\ldots,\beta_{ik},\ldots,\beta_{pk}]$) for assessing the factor loading ($a_{ik} = \sqrt{\lambda_k}\beta_{ik}$)and the number of factors (m).

Step 3: Consider the eigenvalue ordering ($\lambda_1 > \cdots > \lambda_k > \cdots > \lambda_m$; $\lambda_m > 1$) to decide the number of common factors, and pick the number of common factors to be extracted by a predetermined criterion.

Step 4: According to Kaiser (1958), use varimax criteria to find the rotated factor loading matrix, which provides additional insights for the rotation of factor-axis.

Step 5: Name the factor referring to the combination of manifest variables.

When a large set of variables are factored, the method first extracts the combinations of variables, explaining the greatest amount of variance, and then proceeds to combinations that account for progressively smaller amounts of variance. Two kinds of criteria are used for selecting the number of factors: latent root criterion and percentage of variance criterion. The former criterion is that any individual factor should account for the variance ($\mathrm{Var}(Y_k) = \lambda_k$) of at least a single variable if it is to be retained for interpretation. In this criterion only the factors having eigenvalues greater than 1 (i.e. $\lambda_k \geqslant 1$, $k = 1,2,\ldots,m$) are considered significant. The latter criterion is based on achieving a specified cumulative percentage of total variance extracted by successive factors. Its purpose is to ensure the extracted factors can explain at least a specified amount of variance. Practically, to be satisfactory the total amount of variance explained by factors should be at least 95% in the natural sciences, and 60% in the social sciences. However, no absolute threshold has been adopted for all applications (Hair, Anderson, Tatham, & Black, 1998).

3.2. Clarifying the interrelation between criteria of a factor

In a totally interdependent system, all criteria of the systems are mutually related, directly or indirectly; thus, any interference with one of the criteria affects all the others, so it is difficult to find priorities for action. The decision-maker who wants to obtain a specific objective/aspect is at a loss if the decision-maker wants to avoid disturbing the rest of the system while attaining the decision-maker's objective/aspect. While the vision of a totally interdependent system leads to passive positions, the vision of a clearer hierarchical structure leads to a linear activism which neglects feedback and may engineer many new problems in the process of solving the others.

The DEMATEL method, developed by the Science and Human Affairs Program of the Battelle Memorial Institute of Geneva between 1972 and 1976, was used for researching and solving the complicated and intertwined problem group. DEMATEL was developed in the belief that pioneering and appropriate use of scientific research methods could improve understanding of the specific *problematique*, the cluster of intertwined problems, and contribute to identification of workable solutions by a hierarchical structure. The methodology, according to the concrete characteristics of objective affairs, can confirm the interdependence among the variables/attributes and restrict the relation that reflects the characteristic with an essential system and development trend (Chiu, Chen, Tzeng, & Shyu, 2006; Hori & Shimizu, 1999; Tamura, Nagata, & Akazawa, 2002). The end product of the DEMATEL process is a visual representation—an individual map of the mind—by which the respondent organizes his or her own action in the world.

The purpose of the DEMATEL enquiry in this paper is the analysis components structure of each factor, the direction and intensity of direct and indirect relationships that flow between apparently well-defined components. Experts' knowledge is checked and analyzed to contribute to a greater understanding of the component elements and the way they interrelate. The result of DEMATEL analysis can illustrate the interrelations structure of components and can find the central components of the problem to avoid the “overfitting” for decision-making.

The steps of the DEMATEL method are described as follows:

Step 1: Calculate the average matrix. Respondents were asked to indicate the direct influence that they believe each element exerts on each of the others according to an integer scale ranging from 0 to 4. A higher score from a respondent indicates a belief that insufficient involvement in the problem of element i exerts a stronger possible direct influence on the inability of element j, or, in positive terms, that greater improvement in i is required to improve j.

From any group of direct matrices of respondents it is possible to derive an average matrix $\boldsymbol{A}$. Each element of this average matrix will be in this case the mean of the same elements in the different direct matrices of the respondents.

Step 2: Calculate the initial direct influence matrix. The initial direct influence matrix $\boldsymbol{D}$ can be obtained by normalizing the average matrix $\boldsymbol{A}$, in which all principal diagonal elements are equal to zero. Based on matrix $\boldsymbol{D}$, the initial influence which an element exerts and receives from another is shown.

The element of matrix $\boldsymbol{D}$ portrays a contextual relationship among the elements of the system and can be converted into a visible structural model—an *impact-digraph-map*—of the system with respect to that relationship. For example, as shown in Fig. 2, the respondents are asked to indicate only direct links. In the directed digraph graph represented here, element i directly affects only elements j and k; indirectly, it also affects first l, m and n and, secondly, o and q. The digraph map helps to understand the structure of elements.

Step 3: Derive the full direct/indirect influence matrix. A continuous decrease of the indirect effects of problems along the powers of the matrix $\boldsymbol{D}$, e.g. $\boldsymbol{D}^2, \boldsymbol{D}^3, \ldots, \boldsymbol{D}^\infty$, and therefore guarantees convergent solutions to matrix inversion. In a configuration like Fig. 2, the influence exerted by element i on element q will be smaller than influence that element i exerts on element m, and again smaller than the influence exerted on element j. This being so, the infinite series of direct and indirect effects can be illustrated. Let the (i,j) element of matrix $\boldsymbol{A}$ is denoted by a_{ij}, the matrix can be gained following Eqs. (1)–(4).

$$\boldsymbol{D} = s \cdot A, \quad s > 0 \tag{1}$$

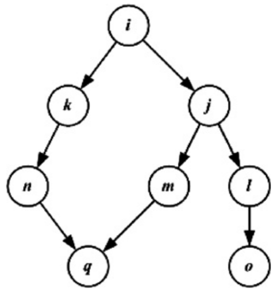

Fig. 2. An example of direct graph.

or

$$d_{ij} = s \cdot a_{ij}, \quad s > 0, \; i,j \in \{1,2,\ldots,n\} \tag{2}$$

$$0 < s < \sup, \quad \sup = \mathrm{Min}\left(\frac{1}{\max_{1\leqslant i\leqslant n}\sum_{j=1}^{n}|a_{ij}|}, \frac{1}{\max_{1\leqslant j\leqslant n}\sum_{i=1}^{n}|a_{ij}|}\right) \tag{3}$$

and

$$\lim_{m\to\infty} \boldsymbol{D}^m = [0], \quad \text{where } \boldsymbol{D} = [d_{ij}]_{n\times n}, \quad 0 \leqslant d_{ij} < 1 \tag{4}$$

The full direct/indirect influence matrix $\boldsymbol{F}$—the infinite series of direct and indirect effects of each element—can be obtained by the matrix operation of $\boldsymbol{D}$. The matrix $\boldsymbol{F}$ can show the final structure of elements after the continuous process (see Eq. (5)). Let $W_i(f)$ denote the normalized ith row sum of matrix $\boldsymbol{F}$; thus, the $W_i(f)$ value means the sum of influence dispatching from element i to the other elements both directly and indirectly. The $V_i(f)$, the normalized ith column sum of matrix $\boldsymbol{F}$, means that the sum of influence that element i receives from the other elements.

$$\boldsymbol{F} = \sum_{i=1}^{\infty} \boldsymbol{D}^i = \boldsymbol{D}(\boldsymbol{I} - \boldsymbol{D})^{-1} \tag{5}$$

Step 4: Set threshold value and obtain the impact-digraph-map. Setting a threshold value, p, to filter the obvious effects denoted by the elements of matrix $\boldsymbol{F}$, is necessary to explain the structure of the elements. Based on the matrix $\boldsymbol{F}$, each element, f_{ij}, of matrix $\boldsymbol{F}$ provides information about how element i influences to element j. If all the information from matrix $\boldsymbol{F}$ converts to the impact-digraph-map, the map will be too complex to show the necessary information for decision-making. To obtain an appropriate impact-digraph-map, decision-maker must set a threshold value for the influence level. Only some elements, whose influence level in matrix $\boldsymbol{F}$ higher than the threshold value, can be chose and converted into the impact-digraph-map.

The threshold value is decided by the decision-maker or, in this paper, by experts through discussion. Like matrix $\boldsymbol{D}$, contextual relationships among the elements of matrix $\boldsymbol{F}$ can also be converted into a digraph map. If the threshold value is too low, the map will be too complex to show the necessary information for decision-making. If the threshold value is too high, many elements will be presented as independent elements without showing the relationships with other elements. Each time the threshold value increases, some elements or relationships will be removed from the map.

After threshold value and relative impact-digraph-map are decided, the final influence result can be shown. For example, the impact-digraph-map of a factor is the same as Fig. 2 and eight elements exist in this map. Because of continuous direct/indirect effects between the eight elements, the effectiveness of these eight elements can be represented by two independent *final affected elements*: o and q. The other elements not shown in the impact-digraph-map of a factor can be considered as independent elements because no obvious interrelation with others exists.

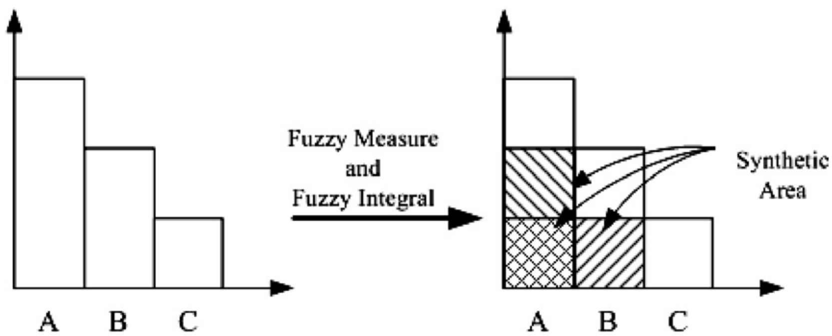

Fig. 3. Non-additive methods for finding the synthetic effect.

3.3. Determining the criteria weights and utility value of factors

The reason for applying fuzzy measure and fuzzy integral is based on the assumption that the synthetic effects of human perception exist between dependent criteria (shown as Fig. 3). Traditionally, researchers use additive techniques to evaluate the utilities of each criterion to meet the assumption of independent relationship among considered criteria. In the proposed model, the non-additive methods, or the sum between the measure of a set and the measure of its complement is not equal to the measure of space, are used to evaluate e-learning program effectiveness. Unlike the traditional definition of a measure based on the additive property, the non-additive MCDM methods, fuzzy measure and fuzzy integral, have been applied to evaluate the dependent multi-criteria problem.

The fuzzy measure was used to determine weights of dependent criteria from subjective judgment and the fuzzy integral was used to evaluate the effectiveness of the final affected elements in an e-learning program. Since Zadeh put forward the fuzzy set theory (Zadeh, 1965), and Bellman and Zadeh described the decision-making methods in fuzzy environments (Bellman & Zadeh, 1970), an increasing number of studies have dealt with uncertain fuzzy problems by applying fuzzy measure and fuzzy integral (Chiou & Tzeng, 2002; Chiou, Tzeng, & Cheng, 2005; Shee, Tzeng, & Tang, 2003; Tzeng, Yang, Lin, & Chen, 2005).

The concept of fuzzy measure and fuzzy integral was introduced by Sugeno. Fuzzy measure is a measure for representing the membership degree of an object to candidate sets (Sugeno, 1977). A fuzzy measure is defined as follows:

Definitions. Let X be a universal set and $P(X)$ be the power set of X.

A fuzzy measure, g, is a function, which assigns each crisp subset of X a number in the unit interval $[0,1]$ with three properties:

1. g: $P(X)\to [0,1]$;
2. $g(\emptyset)=0$, $g(X)=1$ (boundary conditions);
3. $A\subset B\in X$ implies $g(A)\leqslant g(B)$ (monotonicity).

In fuzzy measure, researchers always choose λ-measure to measure the relationship of each element. Sugeno proposed the so-called λ-fuzzy measure or Sugeno measure satisfying the following additional two properties:

1. $\forall A,\ B\in P(X),\ A\cap B=\phi$;
2. $g_\lambda(A\cup B)=g_\lambda(A)+g_\lambda(B)+\lambda g_\lambda(A)g_\lambda(B)$, where $\lambda\in(-1,\infty)$.

For two criteria A and B, if $\lambda>0$, i.e. $g_\lambda(A\cup B)>g_\lambda(A)+g_\lambda(B)$ implies A, B have multiplicative effect; $\lambda=0$ implies A and B have additive effect; and $\lambda<0$ imply A, B have substitutive effect. Since λ value is in the interval $(-1,\infty)$, researcher usually choose λ value as -0.99 and 1 to represent the different types of effect and to discuss the results.

General fuzzy measures and fuzzy integrals, which require only boundary conditions and monotonicity, are suitable for real life. Fuzzy measures and fuzzy integrals can analyze the human evaluation process and specify decision-makers' preference structures. Following the results of Section 3.2, the impact-digraph-map and the interrelation between components of each factor are illustrated. Criteria effectiveness is affected directly/indirectly by other criteria, and can be calculated as follows:

Step 1: Calculate affected element weights using fuzzy measure. Let X be a finite criterion set, $X=\{x_1,x_2,\ldots,x_\mathrm{n}\}$, and $P(X)$ be a class of all the subsets of X. It can be noted as $g_i=g_\lambda(x_i)$. Based on the properties of Sugeno measure, the fuzzy measure $g_\lambda(X)=g_\lambda(\{x_1,x_2,\ldots,x_\mathrm{n}\})$ can be formulated as Eqs. (6) and (7) (Leszcynski, Penczek, & Grochulski, 1985).

$$
\begin{aligned}
&g_\lambda(\{x_1,x_2,\ldots,x_n\})\\
&\quad=\sum_{i=1}^{n}g_i+\lambda\sum_{i1=1}^{n-1}\sum_{i2=i1+1}^{n}g_{i1}g_{i2}+\cdots+\lambda_{n-1}g_{i1}g_{i2}\cdots g_{in}\\
&\quad=\frac{1}{\lambda}\left|\prod_{i=1}^{n}(1+\lambda g_i)-1\right|\quad\text{for }-1<\lambda<\infty \qquad (6)
\end{aligned}
$$

$$\lambda+1=\prod_{i=1}^{n}(1+\lambda g_i) \qquad (7)$$

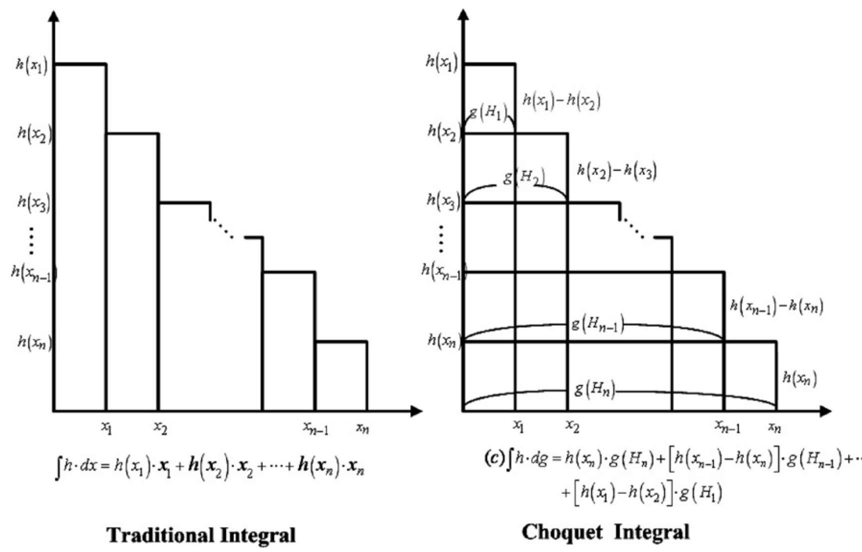

Fig. 4. Diagrams of traditional Riemann integral and non-additive fuzzy integral (Choquet integral).

Step 2: Calculate the effectiveness of final affected elements using fuzzy integral. The fuzzy integral is often used with fuzzy measure for the purpose of congregating information evaluation. The Choquet integral of fuzzy measure is the most frequently used calculation method. This paper adopts this method to calculate the effectiveness scores of final affected elements (criteria) of a factor. The basic concept of traditional integral and fuzzy integral can be illustrated in Fig. 4.

Let h is a measurable set function defined on the measurable space $(X, \aleph)$, suppose that $h(x_1) \geqslant h(x_2) \geqslant \cdots \geqslant h(x_n)$, then the fuzzy integral of fuzzy measure $g(\cdot)$ with respect to $h(\cdot)$ can be defined as Eq. (8) (Chen & Tzeng, 2001; Chiou & Tzeng, 2002; Ishii & Sugeno, 1985; Sugeno, 1974) ($(c)\int h\,dg$ means the Choquet integral). In addition, if $\lambda = 0$ and $g_1 = g_2 = \cdots = g_n$ then $h(x_1) \geqslant h(x_2) \geqslant \cdots \geqslant h(x_n)$ is not necessary. The basic concept of traditional integral and fuzzy integral can be illustrated in Fig. 4.

$$\begin{aligned}(c)\int h\,dg &= h(x_n)\cdot g(H_n) + [h(x_{n-1}) - h(x_n)] \\ &\quad \cdot g(H_{n-1}) + \cdots + [h(x_1) - h(x_2)]\cdot g(H_1) \\ &= h(x_n)\cdot[g(H_n) - g(H_{n-1})] + h(x_{n-1}) \\ &\quad \cdot [g(H_{n-1}) - g(H_{n-2})] + \cdots + h(x_1) \\ &\quad \cdot g(H_1), \quad \text{where } H_1 \\ &= \{x_1\}, H_2 = \{x_1, x_2\}, \ldots, H_n \\ &= \{x_1, x_2, \ldots, x_n\} = X \end{aligned} \tag{8}$$

Step 3: Calculate factor effectiveness. Factor effectiveness can be obtained based on the effectiveness of the final affected elements and other independent elements using the AHP method to be described in Section 3.4.

3.4. Determining factors weights and overall utility value

The analytical hierarchy procedure (AHP) is proposed by Saaty (1980). AHP was originally applied to uncertain decision problems with multiple criteria, and has been widely used in solving problems of ranking, selection, evaluation, optimization, and prediction decisions (Golden, Wasil, & Levy, 1989). Harker and Vargas (1987) stated that "AHP is a comprehensive framework designed to cope with the intuitive, rational, and the irrational when we make multi-objective, multi-criteria, and multi-factor decisions with and without certainty for any number of alternatives." The AHP method is expressed by a unidirectional hierarchical relationship among decision levels. The top element of the hierarchy is the overall goal for the decision model. The hierarchy decomposes to a more specific criteria until a level of manageable decision criteria is met (Meade & Presley, 2002). Under each criterion, subcriteria elements relative to the criterion can be constructed. The AHP separates complex decision problems into elements within a simplified hierarchical system (Shee et al., 2003).

AHP procedures to gain the weights are described as follows:

Step 1: Pairwise-compare the relative importance of factors and obtain a $n \times n$ pairwise comparison matrix; n means the number of criteria.

Step 2: Check the logical judgment consistency using the consistency index (C.I.) and consistency ratio (C.R.). The C.I. value is defined as C.I. = $(\lambda_{max} - n)/(n-1)$, and the λ_{max} is the largest eigenvalue of the pairwise comparison matrix. The C.R. value is defined as C.R. = C.I./R.I. (R.I.: random index. The R.I. value is decided by the value of n. The R.I. values from $n = 1$ to 10 be 0, 0, 0.58, 0.9, 1.12, 1.24, 1.32, 1.41, 1.45 and 1.49). In general, the values of C.I. and C.R. should be less than 0.1 or reasonably consistent.

Step 3: Use the normalized eigenvector of the largest eigenvalue (λ_{max}) as the factor weights.

The purpose of the AHP enquiry in this paper is to construct a hierarchical evaluation system. Based on the independent factors obtained in Section 3.1 and the reduced criteria derived from Section 3.2, the AHP method could gain factor weights and criteria, and then obtain the final effectiveness of the e-learning program.

4. Empirical experiment: cases of evaluating intertwined effects in e-learning

The empirical experiment in this paper was a collaborative research with MasterLink Securities Corporation, Taiwan. The empirical examples are two e-learning training programs. Program 1, a novice-training program designed to acquaint new employees with the regulations, occupational activities, and visions of a corporation, was established by Masterlink Securities. Program 2, designed by the Taiwan Academy of Banking and Finance, is a professional administration skills training program. Based on the approach constructed in Section 3, these two programs are used to explain the feasibility and features of the proposed evaluation model.

4.1. Materials

MasterLink Securities, founded in 1989, developed its core business to including brokerage, asset management, and investment banking in Taiwan, China, and Hong Kong. In Taiwan, Masterlink Securities Corporation with its 44 branches, has used e-learning as a training tool since 2003. Except for courses developed by Masterlink Securities Corporation or purchased from the Taiwan Academy of Banking and Finance, some courses are outsourcing to consulting firms. An effective e-learning evaluation model is necessary for a company designed training programs and budget allowance.

Based on the criteria and approaches from the ADDIE model, Kirkpatrick theories, CIRO model, and other theories (Bitner, 1990; Giese & Gote, 2000; Moisio & Smeds, 2004; Noe, 1986; Santos & Stuart, 2003; Wang, 2003), 58 criteria related to e-learning evaluation were chosen (shown in Appendix) and used to design Questionnaire 1. Employees in Questionnaire 1 were asked to score the importance of each element for effectiveness evaluation; then, the experiment was executed according to four stages as follows:

Step 1: The factor analysis to obtain independent criteria groups. One hundred copies of Questionnaire 1 were distributed to employees of Masterlink Securities Corporation, with 65 responses. Respondents included experts and professionals, familiar and experienced with e-learning. Respondents were evaluated using the SPSS version 11.0.5 for reliability analysis and factor analysis. According to the results of factor analysis, independent factors were obtained and named.

Step 2: The DEMATEL method to find the interrelation between entwined criteria. According to the factor analysis results, some experts were invited to discuss the relationship and influence level of criteria under the same factor, and to score the relationship among criteria based on the DEMATEL method. Factors were divided into different types, so the experts could answer the questionnaire in areas they were familiar with. In order to limit information loss from DEMATEL method results, threshold values were decided after discussion with these experts and an acceptable impact-digraph-map was found.

Step 3: The fuzzy measure approach to find out the weights of intertwined criteria and the fuzzy integral to calculate effectiveness. According to DEAMTEL results, the intertwined criteria structures of a factor were found and the fuzzy measure employed to derive central criteria weights. Based on a map of each factor, after setting the λ value as -0.99 and 1, the substitute effect and multiplicative effect, the fuzzy measure was used to calculate two different weight sets of final affected elements. Concurrently, Questionnaire 2 was designed to investigate criteria effectiveness for using the fuzzy integral method. Questionnaire 2, a web questionnaire, asked Masterlink Securities Corporation employees to score the utility value of criteria of two programs.

Step 4: The AHP method to find the weights and derive e-learning program effectiveness. A further goal for Questionnaire 2 was to use a pair-comparing method to find the factor weights and reduced criteria by AHP methods, and ask employees to score the satisfaction utility of criteria. The score is based on the Likert five-point scale; 1 stands for very dissatisfied, 2 for dissatisfied, 3 for neither dissatisfied or satisfied, 4 for satisfied, 5 for very satisfied. Because there were two different program types and objectives, Questionnaire 2 was delivered to different employee groups. Twenty-six and 28 e-learning questionnaire surveys were returned, after which, factor weights and criteria were obtained and program effectiveness calculated.

4.2. Results

4.2.1. Result of Stage 1

Questionnaire reliability analysis was analyzed following responses received. According to reliability analysis results, Cronbach's α value is higher than 0.8 and the standardized element α value is 0.977 showing questionnaire reliability to be significant and effective (reliability analysis results shown in Table 1).

KMO and Bartlett's test was used to measure the appropriate usage of factor analysis. According to Kaiser's research, KMO > 0.7 is middling to do factor analysis, and KMO > 0.8 is meritorious. The KMO value of this paper is 0.737 (Bartlett's test of sphericity: approximately

Table 1
Reliability analysis results

Source of variance	Sum of sq.	d.f.	Mean square	*F*-test	Probability
Between people	4541.418	65	69.868		
Within people	6210.810	376	1.651		
Between measures	308.001	57	5.404		
Residual	5902.809	371	1.593	3.392	0.000
Total	10752.229	383	2.81		
Grand mean	6.973				
Alpha	0.977				
Standardized element alpha = 0.978					

$\chi^2 = 4740$, d.f. = 1653, significance = 0.000); therefore, it is suitable for factor analysis. This method uses a correlation coefficient to test whether it is suitable and significant to use factor analysis. According to the results of KMO and Bartlett's test, this questionnaire is suitable to use factor analysis.

The principle component analysis was used to extract factors from 58 criteria and the varimax method was used for factor rotation. Then, nine factors whose eigenvalue was more than 1.0 were chosen. Nine factors were named based on the loading of each factor: "Personal Characteristics and System Instruction," "Participant Motivation and System Interaction," "Range of Instruction Materials and Accuracy," "Webpage Design and Display Of Instruction Materials," "E-Learning Environment," "Webpage Connection," "Course Quality and Work Influence," "Learning Records" and "Instruction Materials" (Shown in Table 2).

4.2.2. Result of Stage 2

According to factor analysis results, some experts and professionals were invited to discuss and scored the relation between criteria of each factor based on the DEMATEL approach. Experts and professionals included system designers, webpage designers, instructors, managers, and human resources experts. Factors 1 and 2 were discussed with managers and human resources experts. Factor 4 was discussed with webpage designers. Factors 5 and 6 were discussed with system designers. Instructors were responsible to factors 3, 7, 8, and 9.

Thus, after experts and professionals scored the relation of criteria, the full direct/indirect influence matrix and the impact-digraph-map of each factor was calculated and drawn. According to the results of DEMATEL, the threshold value of each factor was decided by the experts. The threshold value of each factor from factors 1 to 9 is 0.85, 0.47, 1.5, 2.1, 1.6, 6.5, 2.1, 3.8 and 3.5. The impact-digraph-maps of DEMATEL method results were obtained and shown as Fig. 5.

4.2.3. Result of Stage 3

According to Fig. 5, the intertwined structures of several criteria, affected by other criteria, were illustrated. Therefore, the fuzzy measure for the final affected elements of each factor could be calculated out. Using factor 1 as an example, the criteria, "Rewards" and "Learning Expectations," are two final affected elements affected by other criteria, but they did not influence other criteria. "Rewards" was affected by "Personal Motivation," "Self-Efficacy," "Career Planning," and "Ability;" "Learning Expectations" was affected

Table 2
Factor analysis result: names and components (criteria) of factors

	Factor	Components	λ[a]	A[b]	B[c]
1	Personal Characteristics and System Instruction	Personal Motivation, Rewards, Work Attitude, Learning Expectation, Work Characteristics, Self-Efficacy, Ability, Career Planning, Organization Culture, Instruction Goals, System Functions, System Instructions	25.98	44.8	44.8
2	Participant Motivation and System Interaction	Operating Skills, Solving Solutions, Mastery, Managerial Skills, Professional Skills, Inspire Originality, Supervisor's Support, Colleagues, Work Environment, Causes of Problem, Understanding Problems, Pre-Course Evaluation, Multi-Instruction, Communication Ways	4.926	8.494	53.3
3	Range of Instruction Materials and Accuracy	Accuracy, Range of Instruction Materials, Sequence of Instruction Materials, Usage of Multimedia	3.945	6.802	60.1
4	Webpage Design and Display of Instruction Materials	Text & Title, Display of Webpages, Sentence Expression, Length of Webpages, Graphs and Tables, Colors of Webpages	2.533	4.368	64.5
5	E-Learning Environment	Browser Compatibility, Browsing Tool, Path of Webpages, Transferring Time, Available, Reflection of Opinions	1.956	3.372	67.83
6	Webpage Connection	Underconstructing Webpages, System Prompts, Connecting to Main Page, Connection of Webpages	1.846	3.183	71.02
7	Course Quality and Work Influence	Course Arrangement, Course Design, Personal Satisfaction, Technical Evaluation, Course Contents, ROI/Work Influence	1.667	2.874	73.9
8	Learning Records	Learning Records, Instruction Activities, Course Subject	1.505	2.596	76.5
9	Instruction Materials	Level of Instructional Materials, Update Frequency, Readable	1.282	2.21	78.7

Extraction method: principal component analysis.
Rotation method: Varimax with Kaiser normalization.
[a] Eigenvalue.
[b] Percentage of variance.
[c] Cumulative %.

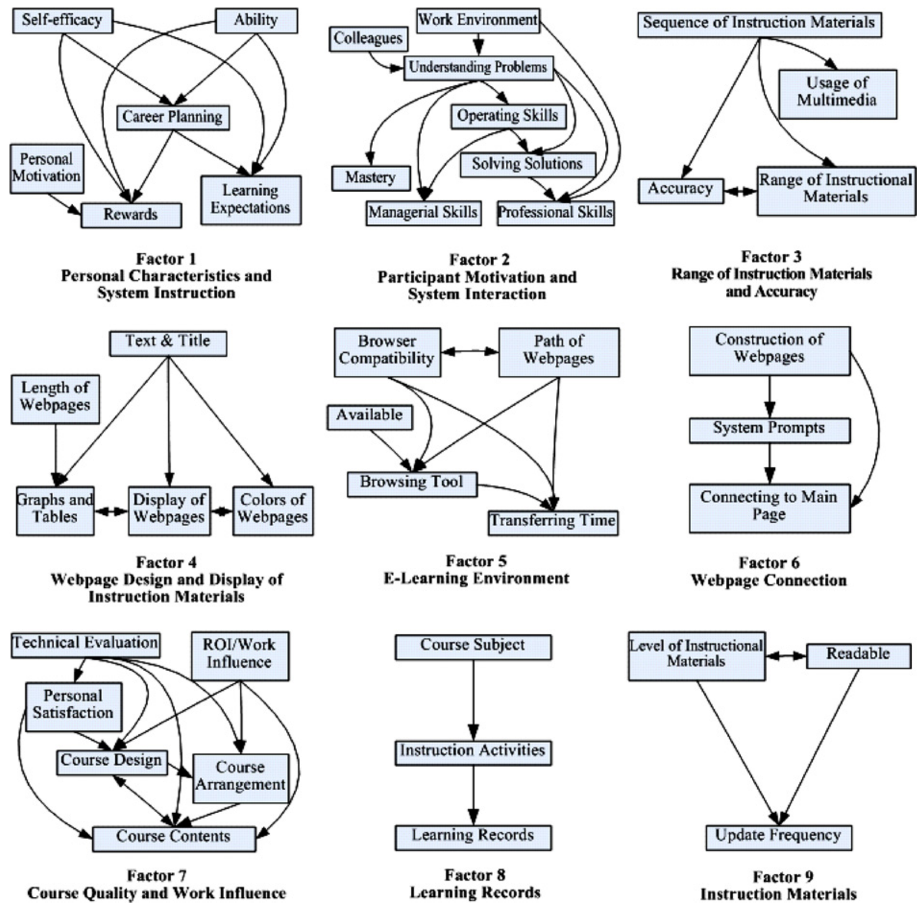

Fig. 5. The impact-digraph-maps of nine factors derived by DEMATEL method.

by "Career Planning," "Ability," and "Self-Efficacy." Since these criteria have an influential relationship, the fuzzy measure should be employed to evaluate the weights of "Rewards" and "Expectations." The λ value was set as 1 and -0.99, indicating different synthetic effects of criteria. Fuzzy measure results of final affected elements of factor 1 are listed in Table 3. The e-learning satisfaction survey could then be implemented to calculate the fuzzy integral value of each factor. For example, the satisfaction value of the criteria, "Personal Motivation," "Self-Efficacy," "Ability," and

Table 3
Fuzzy measure for two final affected elements of factor 1

Factor	Element	λ	Fuzzy measure
1	Rewards	1	$g_{1-1}=0.192$, $g_{1-6}=0.190$, $g_{1-7}=0.190$, $g_{1-8}=0.189$ $g_{(1-1,1-6)}=0.416$, $g_{(1-1,1-7)}=0.416$, $g_{(1-1,1-8)}=0.417$, $g_{(1-6,1-7)}=0.411$, $g_{(1-6,1-8)}=0.412$, $g_{(1-7,1-8)}=0.412$, $g_{(1-1,1-6,1-7)}=0.683$, $g_{(1-1,1-7,1-8)}=0.683$, $g_{(1-1,1-6,1-8)}=0.683$, $g_{(1-6,1-7,1-8)}=0.678$ $g_{(1-1,1-6,1-7,1-8)}=1$
		-0.99	$g_{1-1}=0.696$, $g_{1-6}=0.689$, $g_{1-7}=0.689$, $g_{1-8}=0.690$ $g_{(1-1,1-6)}=0.910$, $g_{(1-1,1-7)}=0.910$, $g_{(1-1,1-8)}=0.910$, $g_{(1-6,1-7)}=0.908$, $g_{(1-6,1-8)}=0.910$, $g_{(1-7,1-8)}=0.908$, $g_{(1-1,1-6,1-7)}=0.978$, $g_{(1-1,1-7,1-8)}=0.978$, $g_{(1-1,1-6,1-8)}=0.978$, $g_{(1-6,1-7,1-8)}=0.978$ $g_{(1-1,1-6,1-7,1-8)}=1$
	Learning Expectations	1	$g_{1-6}=0.260$, $g_{1-7}=0.260$, $g_{1-8}=0.260$ $g_{(1-6,1-7)}=0.587$, $g_{(1-6,1-8)}=0.588$, $g_{(1-7,1-8)}=0.588$, $g_{(1-6,1-7,1-8)}=1$
		-0.99	$g_{1-6}=0.792$, $g_{1-7}=0.792$, $g_{1-8}=0.793$ $g_{(1-6,1-7)}=0.963$, $g_{(1-6,1-8)}=0.963$, $g_{(1-7,1-8)}=0.963$, $g_{(1-6,1-7,1-8)}=1$

Elements: 1–1: "Personal Motivation"; 1–6: "Self-Efficacy"; 1–7: "Ability"; 1–8: "Career Planning".

Table 4
Fuzzy integral results of each element in different programs

Factor	Elements of factor	λ value	Integral value		Directive impact elements	Indirective impact elements
			Program 1	Program 2		
1	Rewards	1	2.475	3.589	Self-Efficacy, Ability, Career Planning, Personal Motivation	
		−0.99	2.552	3.753		
	Learning Expectations	1	2.447	3.593	Self-Efficacy, Ability, Career Planning	
		−0.99	2.476	3.764		
2	Managerial Skills	1	2.529	3.641	Understanding Problems, Operating Skills	Work Environment, Colleagues
		−0.99	2.548	3.693		
	Professional Skills	1	2.507	3.609	Work Environment, Understanding Problems, Solving Solutions	Colleagues, Operating Skills
		−0.99	2.623	3.761		
	Mastery[a]		2.585	3.684	Understanding Problems	Work Environment, Colleagues
3	Accuracy	1	2.671	3.626	Sequence of Instruction Materials, Range of Instruction Materials	
		−0.99	2.763	3.682		
	Range of Instruction Materials	1	2.641	3.604	Sequence of Instruction Materials, Accuracy	
		−0.99	2.696	3.678		
	Usage of Multimedia[a]		2.484	3.745	Sequence of Instruction Materials	
4	Display of Webpages	1	2.537	3.697	Text & Title, Graphs and Tables, Colors of Webpages	Length of Webpages
		−0.99	2.645	3.740		
	Graphs and Tables	1	2.471	3.688	Text & Title, Length of Webpages, Display of Webpages	Colors of Webpages
		−0.99	2.577	3.739		
	Colors of Webpages	1	2.508	3.736	Text & Title, Display of Webpages	Length of Webpages, Graphs and Tables
		−0.99	2.601	3.745		
5	Transferring Time	1	2.360	3.602	Browser Compatibility, Browsing Tool, Path of Webpages	Available
		−0.99	2.413	3.643		
6	Connect To Main Page	1	2.498	3.608	Construction of Webpages, System Prompts	
		−0.99	2.498	3.620		
7	Course Contents	1	2.604	3.718	Technical Evaluation, ROI/Work Influence,	
		−0.99	2.676	3.771	Personal Satisfaction, Course Design, Course Arrangement	
8	Learning Records[a]		2.318	3.658	Instruction Activities	Course Subject
9	Update Frequency	1	2.520	3.720	Level of Instructional Materials, Readable	
		−0.99	2.546	3.741		

[a] Without synthetic effect, the element did not use the fuzzy measure and fuzzy integral for evaluation.

Table 5
Final score of each program

Factor	AHP weight (factor)	AHP weight (criterion)	Elements of factor	Fuzzy integral			
				$\lambda = -0.99$		$\lambda = 1$	
				Program 1	Program 2	Program 1	Program 2
1	0.105	0.249	Rewards	2.552	3.753	2.475	3.589
		0.249	Learning Expectations	2.476	3.764	2.447	3.593
		0.086	Work Attitude[a]	2.438	3.729	2.438	3.729
		0.082	Work Characteristics[a]	2.517	3.666	2.517	3.666
		0.084	Organization Culture[a]	2.451	3.537	2.451	3.537
		0.085	Instruction Goals[a]	2.186	3.703	2.186	3.703
		0.086	System Functions[a]	2.362	3.640	2.362	3.640
		0.082	System Instructions[a]	2.258	3.615	2.258	3.615
2	0.115	0.183	Managerial Skills	2.548	3.693	2.529	3.641
		0.183	Professional Skills	2.623	3.761	2.507	3.609
		0.180	Mastery[b]	2.585	3.684	2.585	3.684
		0.077	Inspire Originality[a]	2.281	3.518	2.281	3.518
		0.077	Supervisor's Support[a]	2.578	3.799	2.578	3.799
		0.078	Causes of Problem[a]	2.475	3.597	2.475	3.597
		0.073	Pre-Course Evaluation[a]	2.498	3.495	2.498	3.495
		0.074	Multi-Instruction[a]	2.592	3.729	2.592	3.729
		0.074	Communication Ways[a]	2.438	3.684	2.438	3.684
3	0.109	0.378	Accuracy	2.763	3.682	2.671	3.626
		0.378	Range of Instruction Materials	2.696	3.678	2.641	3.604
		0.245	Usage of Multimedia[b]	2.484	3.745	2.484	3.745
4	0.109	0.284	Display of Webpages	2.645	3.740	2.537	3.697
		0.276	Graphs and Tables	2.577	3.739	2.471	3.688
		0.278	Colors of Webpages	2.601	3.745	2.508	3.736
		0.167	Sentence Expression[a]	2.601	3.719	2.601	3.719
5	0.114	0.835	Transferring Time	2.413	3.643	2.360	3.602
		0.165	Reflection of Opinions[a]	2.331	3.631	2.331	3.631
6	0.111	0.679	Connect To Main Page	2.498	3.620	2.498	3.608
		0.321	Underconstructing Webpages[a]	2.498	3.597	2.498	3.597
7	0.109	1	Course Contents	2.676	3.771	2.604	3.718
8	0.104	1	Learning Records[b]	2.318	3.658	2.318	3.658
9	0.110	1	Update Frequency	2.546	3.741	2.520	3.720
Final score				2.489	3.644	2.452	3.610

[a] The criteria whose *influence level* did not reach the threshold value were considered independent criteria.
[b] Without synthetic effect, the element did not use the fuzzy measure and fuzzy integral for evaluation.

"Career Planning" in program 2 are 3.597, 3.792, 3.719 and 3.370, and the integral value of "Rewards" at $\lambda = 1$ is 3.589. The fuzzy integral values of the final affected elements are shown in Table 4. These results could be implemented to calculate final results of each program.

4.2.4. Result of Stage 4

The weights of nine factors and the reduced criteria were calculated out and used to find the effectiveness of each program. The final score for each program is shown in Table 5.

5. Discussions

The proposed novel hybrid MCDM method should be a useful model for evaluating e-learning program effectiveness. Based on our empirical experiments of the Masterlink Securities Corporation's e-learning program survey, factor analysis was used to classify each element into nine different independent factors. Those criteria under the same factor had some interrelations with each other. The direct/indirect influential relationship of criteria was figured using the DEMATEL method. Affected criteria effectiveness was determined with the fuzzy integral value. Then, program effectiveness values were calculated by considering independent criteria effectiveness results, fuzzy integral value of intertwined criteria, and AHP factor weights. The hybrid MCDM model proposed in this paper contains the following properties:

5.1. The key elements found and improvement alternatives illustrated

Using the proposed model, a company may find factors that improve e-learning effectiveness. This paper also used the DEAMTEL method to find the direct/indirect influential relationship of criteria that helps reduce the number of criteria and find factor improvement direction. Therefore, interactive effects accurately reflect in the final evaluation.

According to weights derived by the AHP, central factors, which are more important and will affect e-learning effectiveness, could be found. Therefore, the evaluator could determine the score of one e-learning program. After using this e-learning effectiveness evaluation model, evaluators found the aspects needing improvement, for e-learning effectiveness to increase. Although the difference of each factor weight is not significant, as shown in Table 5, factor 5, "E-Learning Environment", with the highest weight (0.114) should be given more attention to effectiveness. The performance of factor "E-Learning Environment" will affect the entire program effectiveness.

Using the DEMATEL can reduce the number of criteria for evaluating factor effectiveness; concurrently, a company can improve the effectiveness of a specific factor based on the impact-digraph-map. For example, the effectiveness of factor "Personal Characteristics and System Instruction," can be represented by the effectiveness of central criteria "Rewards" and "Learning Expectations," but the key element for improving factor "Personal Characteristics and System Instruction" are "Self-Efficacy" and "Ability." It is easier for a company to find the exact department or persons responsible for improvement using results from the proposed model.

5.2. The fuzziness in effectiveness perception considered

The non-additive multi-criteria evaluation techniques, fuzzy measure and fuzzy integral, are employed to refine the situations which conform to the assumption of independence between criteria. The λ value used in the fuzzy measure and fuzzy integral affords another viewpoint for evaluating how to remove the mechanical additive evaluating method. This means improving individual criterion performance by considering the effect from the others if the synthetic effect exists. In other words, if the evaluator investigates the types of synthetic effects of learners, designer, managers, and other respondents, program effectiveness can be improved on the dependent criteria with a multiplicative effect.

Moreover, the concepts of the fuzzy measure and fuzzy integral approach used in the proposed model will make evaluation more practical and flexible by using different λ values. For example, the original satisfaction value of criterion "Rewards" of factor, "Personal Characteristics and System Instruction" in program 1 is 2.416. According to Table 4, the synthetic effect comes from "Personal Motivation," "Self-Efficacy," "Ability," and "Career Planning" criteria. After calculating the effectiveness using fuzzy measure ($\lambda = -0.99$) and fuzzy integral, the effectiveness value of element "Rewards" changed to 2.552. This also conforms to the situation that "Rewards" is not the single criterion for a learner to express the satisfaction on factor "Personal Characteristics and System Instruction." If the criteria are independent, the λ value can be set to 0.

5.3. The result of hybrid MCDM model is consistent with the traditional additive model

According to Table 5, the effectiveness of the general administration training (program 2) is better than the novice training (program 1). Whether from substitutive effects ($\lambda = -0.99$) or multiplicative effects ($\lambda = 1$), the effectiveness (satisfaction) of novice training is less than general administration training. The main reason for this result is that new employees go through novice training for the first time and are not familiar with e-learning type training. Therefore, they may not feel comfortable using this system and attending these kinds of programs. Furthermore, general administration training is an e-learning program relative to daily work. The program consists of professional skills helpful to work; hence, employee satisfaction is high.

Comparing the proposed hybrid MCDM model with the traditional additive models, the results are consistent. Program effectiveness is calculated by the traditional AHP method and the scores for programs 1 and 2 are 2.451 and. 3.617. Another survey, which asked employees to score the programs according to the Likert five-point scale for program satisfaction using the simple additive weighting (SAW) method, showed scores for programs 1 and 2 at 2.697 and. 3.828. These results show novice training to be less satisfactory than general administration training which is consistent with results from the proposed model. The results also mean that the hybrid MCDM model is a reasonable tool to evaluate e-learning programs.

6. Concluding remarks and future perspectives

E-learning evaluation is still deficient and does not have evaluation guidelines. Since e-learning could be evaluated in numerous and intertwined facets and criteria, a multicriteria decision making model should be more suitable for e-learning evaluation. This paper outlines a hybrid MCDM evaluation model for e-learning effectiveness.

Based on several aspects of e-learning effectiveness evaluation, this paper integrated several methods to make the proposed model, the hybrid MCDM model, much closer to reality. According to the results of empirical study, the hybrid MCDM model should be a workable and useful model to evaluate e-learning effectiveness and to display the interrelations of intertwined criteria. As a result, if the effectiveness of an e-learning program is deficient we could find out the problem based on AHP weights and interrelation based on the impact-digraph-map of each factor. After using this e-learning effectiveness evaluation model, the evaluators could find the aspects needing improvement, so that e-learning program effectiveness could increase. Compared with traditional e-learning evaluation, this model considers more aspects and criteria which may affect e-learning program effectiveness.

Though this paper establishes a new model to evaluate e-learning effectiveness, some interesting points may be worth investigating for further researches. This paper did not concentrate on the fuzzy measure (λ). Therefore, further research may take real situations and the effects of each factor into consideration. Moreover, this paper takes one corporation as a case to implement the e-learning effectiveness evaluation model. This merely presents one case of the industry. Further research is recommended to look into the corporations of one industry adopting e-learning, and compare e-learning effectiveness among those corporations.

Appendix. Fifty-eight criteria for emperical e-learning programs

No.	Criteria	Description
1	Browser Compatibility	Learning materials could be read by different browsers
2	Browsing Tool	Browsing tool means the tools that could let users know how to go front page, next page and enlarge or contraction pictures. Browsing tool design, menu button design and interface design are consistency and easy to use
3	Path of Webpages	System provides suitable function of learner control. Display the path of learning materials
4	Transferring Time	When a learner is learning online, the waiting time for transferring data is appropriate
5	Available	Learning materials are easy to access and always be available
6	Reflection of Opinions	Instruction website could let instructors to know the opinions of learners
7	Underconstructing Webpages	Webpage won't connect to under-construction Webpages and each links are work
8	System Prompts	When something should be described and give some instructions, system will provide appropriate system prompt. System prompts and instructions match up with learning materials
9	Connecting to Main Page	Every webpage could link back to main page
10	Connection of Webpages	Relative Webpages could connect to each other
11	Text & Title	The size of text and headline are appropriate. The row spacing and spacing are appropriate
12	Display of Webpages	To display in screen size, general appearance are regularity and adjustment
13	Sentence Expressions	The reading sequence, paragraphs, erratum and expression of sentence are appropriate
14	Length of Webpage	The classification of webpage contents and webpage length are comfortable to read
15	Graphs And Tables	Graphs and tables are suitable expressed, the integration and composition and background are displayed appropriately
16	Colors of Webpages	Media display skills and usage of color could let learners feel comfortable. The colors of webpage design consider contrast of colors, systematic usage of colors and harmony of colors
17	Accuracy	The accuracy of learning materials or cited terminology is appropriately used
18	Range of Instruction Materials	The contents of learning material, such as range, depth, integration and structure are properly display
19	Sequence of Instruction Materials	The display of learning materials is ordinal. The instruction materials integrate relative subjects and the structures of instruction material contents are appropriate
20	Usage of Multimedia	Multimedia design is appropriate. The usage of voice and image could attract learners' attention
21	Course Arrangement	Course arrangement is proper. And course arrangement will affect the intention and the level of learners' transfer what they have learned into their daily work
22	Course Design	Course design provides what learners want to learn. According to course design principle, the level of transference of implementing what learners have learned into daily work
23	Personal Satisfaction	Personal satisfaction affects the level of transference of what workpeople have learned into work
24	Technical Evaluation	Personal attitude toward the reflection of technical evaluation feedback affect the level of transference of what workpeople have learned into work
25	Course Contents	According to course contents, the level of transference of implementing what workpeople has learned into work

(*continued on next page*)

Appendix (*continued*)

No.	Criteria	Description
26	ROI/Work Influence	After participating e-learning courses, the affective level of spending time, investment and the return on investment
27	Learning Records	System could record learners' learning behavior and evaluate the learning performance
28	Instruction Activities	Each instructional activity matches up with e-learning. Instruction activities are properly used
29	Course Subject	The range and subject of course is appropriate
30	Level of Instruction Materials	The level of instruction materials is suitable for learners. The learning materials contain their uniqueness
31	Update Frequency	The update date of learning materials, the contents, the subjects and the items are fit in with trend and different time or places
32	Readable	Learning materials are readable. They contain theories and practical issues
33	Personal Motivation	Personal motivations of participating e-learning affect the level of transference of what learners have learned into work
34	Rewards	Merit system and rewards affect the transference of what learners have learned into work
35	Work Attitude	Work attitude affect the level of transference of what learners have learned into work
36	Learning Expectation	Personal expectations toward e-learning affect the level of transference of what learners have learned into work
37	Work Characteristics	Personal work characteristics affect the level of transference of what learners have learned into work
38	Self-Efficacy	Self-efficacy affects the level of transference of what learners have learned into work
39	Ability	Personal abilities affect the level of transference of what learners have learned into work
40	Career Planning	Career planning and objectives setting affect the level of transference of what learners have learned into work
41	Organization Culture	Organization climate and organization culture encourage learners applying what knowledge they have learned to workforce
42	Instruction Goals	Learners realize the instruction goal of e-learning website
43	System Functions	Provide the functional label of system operating interface. Provide search function of learning materials
44	System Instructions	Provide instructions of system software and hardware. Provide the functions of download and print. Provide system menu
45	Operating Skills	After learning, learners could increase the level of operating skills
46	Solving Solutions	After learning, learners could find the way to solve problems
47	Mastery	After learning, learners could master what they have learned during e-learning courses
48	Managerial Skills	After learning, learners could increase the level of managerial skills
49	Professional Skills	After learning, learners could increase the level of professional skills
50	Inspire Originality	After learning, learners could inspire originality
51	Supervisor's Support	Supervisors support affect learners implement what they have learned into work
52	Colleagues	Colleagues could discuss and implement what they have learned into work
53	Work Environment	Working environment encourages learners apply what they have learned to work
54	Causes of Problem	After learning, learners could know the real reason which leads to occurrence
55	Understanding Problems	After learning, learners could increase the understanding level of problems which they want to know
56	Pre-Course Evaluation	According to learners' background, provide pre-course assessment. Attract the motivation and interests of learners
57	Multi-instruction	E-learning courses use multi-instructional ways to express
58	Communication Ways	The communication ways of instruction website are convenient to use

References

Allen, T. D., Russell, J. E. A., Pottet, M. L., & Dobbins, G. H. (1999). Learning and development factors related to perceptions of job content and hierarchical plateauing. *Journal of Organizational Behavior, 20*(12), 1113–1137.

Bellman, R. E., & Zadeh, L. A. (1970). Decision making in a fuzzy environment. *Management Science, 17*(4), 141–164.

Bitner, M. J. (1990). Evaluating service encounters: The effects of physical surroundings and employee responses. *Journal of Marketing, 54*(2), 69–82.

Charles, T., Mahithorn, B., & Paul, A. B. R. (2002). The design of a training porgramme measurement model. *Journal of European Industrial Training, 26*(5), 230–240.

Chen, Y. W., & Tzeng, G. H. (2001). Fuzzy integral for evaluating subjectively perceived travel costs in a traffic assignment model. *European Journal of Operational Research, 130*(3), 653–664.

Chiou, H. K., & Tzeng, G. H. (2002). Fuzzy multiple-criteria decision-making approach for industrial green engineering. *Environmental Management, 30*(6), 816–830.

Chiou, H. K., Tzeng, G. H., & Cheng, D. C. (2005). Evaluating sustainable fishing development strategies using fuzzy MCDM approach. *Omega, 33*(3), 223–234.

Chiu, Y. J., Chen, H. C., Tzeng, G. H., & Shyu, J. Z. (2006). Marketing strategy based on customer behavior for the LCD-TV. *International Journal of Management and Decision Making, 7*(2/3), 143–165.

Cooper, M. (1994). Evaluating professional training. *Training and Development, 10*(10), 26–31.

Dick, W., & Carey, L. (1996). *The systematic design of instruction*. New York: Harper Collins.

Europa (2004). The eContent programme: Stimulating the production of digital content and promoting linguistic diversity. Available from http://europa.eu.int/scadplus/leg/en/lvb/l24226d.htm.

Fontela, E., & Gabus, A. (1974). DEMATEL, innovative methods, Report no. 2, Structural analysis of the world problematique. Battelle Geneva Research Institute.

Fontela, E., & Gabus, A. (1976). *The DEMATEL observer*. Battelle Institute, Geneva Research Center.

Garavaglia, P. L. (1993). How to ensure transfer of training. *Training & Development, 47*(10), 57–69.

Geis, G. L., & Smith, M. E. (1992). The function of evaluation. In H. D. Stolovitch & E. J. Keeps (Eds.), *Handbook of human performance technology*. San Francisco: Jossey-Bass.

Giese, J. L., & Gote, J. A. (2000). Defining consumer satisfaction. *Academy of Marketing Science Review.* Available from http://www.amsreview.org/articles/giese01-2000.pdf.

Golden, B. L., Wasil, E. A., & Levy, D. E. (1989). Applications of the analytic hierarchy process: A categorized, annotated bibliography. In B. L. Golden, E. A. Wasil, & P. T. Harker (Eds.), *The analytic hierarchy process*. Berlin: Springer-Verlag.

Hair, J. J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis*. Englewood Cliffs, NJ: Prentice-Hall.

Hall, M. L., & Nania, S. (1997). Training design and evaluation: An example from a satellite based distance learning program. *Public Administration Quarterly, 21*(3), 370–385.

Hannum, W., & Hansen, C. (1989). *Instructional systems development in large organizations*. Englewood Cliffs, NJ: Educational Technology Publications.

Harker, P., & Vargas, L. (1987). The theory of ratio scale estimation: Saaty's analytic hierarchy process. *Management Science, 33*(11), 1383–1403.

Hegstad, C. D., & Wentlign, R. M. (2004). The development and maintenance of exemplary formal mentoring programs in fortune 500 companies. *Human Resource Management, 15*(4), 421–448.

Hori, S., & Shimizu, Y. (1999). Designing methods of human interface for supervisory control systems. *Control Engineering Practice, 7*, 1413–1419.

Hsieh, P. Y. (2004). Web-based training design for human resources topics: A case study. *TechTrends, 48*(2), 60–68.

Ishii, K., & Sugeno, M. (1985). A model of human evaluation process using fuzzy integral. *International Journal of Man-Machine Studies, 22*(1), 19–38.

Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika, 23*(1), 187–200.

Kirkpatrick, D. L. (1959a). Techniques for evaluating training programs. *Journal of ASTD, 13*(11), 3–9.

Kirkpatrick, D. L. (1959b). Techniques for evaluating training programs: Part 2—learning. *Journal of ASTD, 13*(12), 21–26.

Kirkpatrick, D. L. (1960a). Techniques for evaluating training programs: Part 3—behavior. *Journal of ASTD, 14*(1), 13–18.

Kirkpatrick, D. L. (1960b). Techniques for evaluating training programs: Part 4—results. *Journal of ASTD, 14*(2), 28–32.

Leszcynski, K., Penczek, P., & Grochulski, W. (1985). Sugeno's fuzzy measure and fuzzy clustering. *Fuzzy Sets and Systems, 15*(2), 147–158.

Marks, R. B., Sibley, S. D., & Arbaugh, J. B. (2005). A structural equation model of predictors for effective online learning. *Journal of Management Education, 29*(4), 531–563.

Meade, L. M., & Presley, A. (2002). R&D project selection using the analytic network process. *IEEE Transactions on Engineering Management, 49*(1), 59–66.

Mhod, T. L., Rina, A., & Suraya, H. (2004). Teaching and learning of e-commerce courses via hybrid e-learning model in unitar. *Journal of Electronic Commerce in Organizations, 2*(2), 78–94.

Miller, S. M., & Miller, K. L. (2000). Theoretical and practical considerations in the design of web-based instruction. In B. Abbey (Ed.), *Instructional and cognitive impacts of web-based education*. IDEA Group Publishing.

Moisio, A., & Smeds, R. (2004). E-learning: A service offering. *Knowledge and Process Management, 11*(4), 252–260.

Moore, M. G. (1989). Three types of interaction. *The American Journal of Distance Education, 3*(2), 1–6.

Muilenburg, L. Y., & Berge, Z. L. (2005). Student barriers to online learning: A factor analytic study. *Distance Education, 26*(1), 29–48.

Ng, K. C., & Murphy, D. (2005). Evaluating interactivity and learning in computer conferencing using content analysis techniques. *Distance Education, 26*(1), 89–109.

Noe, R. A. (1986). 'trainees' Attributes and attitudes: Neglected influences on training effectiveness. *Academy of Management Review, 11*(4), 736–749.

Philips, J. (1996). *Accountability in human resource management*. Oxford: Butterworth-Heinemann.

Roffe, I. (2002). E-learning: Engagement, enhancement and execution. *Quality Assurance in Education, 10*(1), 40–50.

Saaty, T. L. (1980). *The analytical hierarchy process: Planning priority setting, resource allocation*. New York: McGraw-Hill.

Santos, A., & Stuart, M. (2003). Employee perceptions and their influence on training effectiveness. *Human Resource Management Journal, 13*(1), 27–45.

Shee, D. Y., Tzeng, G. H., & Tang, T. I. (2003). AHP, fuzzy measure and fuzzy integral approaches for the appraisal of information service providers in Taiwan. *Journal of Global Information Technology Management, 6*(1), 8–30.

Sherry, A. C., Fulford, C. P., & Zhang, S. (1998). Assessing distance learners' satisfaction with instruction: A quantitative and a qualitative measure. *The American Journal of Distance Education, 12*(3), 4–28.

Sugeno, M. (1974). *Theory of fuzzy integrals and its applications*. Tokyo: Tokyo Institute of Technology.

Sugeno, M. (1977). *Fuzzy measures and fuzzy integrals: A survey*. New York: North-Holland.

Tamura, M., Nagata, H., & Akazawa, K. (2002). Extraction and systems analysis of factors that prevent safety and security by structural models. In *41st SICE annual conference, Osaka, Japan*.

Tzeng, G. H., Yang, Y. P. Ou, Lin, C. T., & Chen, C. B. (2005). Hierarchical MADM with fuzzy integral for evaluating enterprise Intranet web sites. *Information Sciences, 169*(3–4), 409–426.

Urdan, T. A. (2000). *Corporate e-learning: Exploring a new frontier*. San Francisco: W.R. Hambrecht.

Wang, Y. C. (2003). Assessment of learner satisfaction with asynchronous electronic learning systems. *Information & Management, 41*(2), 75–86.

Warfield, J. N. (1976). *Societal systems, planning, policy and complexity*. New York: John Wiley & Sons.

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control, 8*(2), 338–353.