

University of Massachusetts Boston

From the Selected Works of Roger H. Blake

2011

A semantic analysis approach for assessing professionalism using free-form text entered online

Roger H. Blake



Available at: <https://works.bepress.com/rogerblake/6/>



A semantic analysis approach for assessing professionalism using free-form text entered online

Roger Blake^{a,*}, Oscar Gutierrez^{b,1}

^a Management Science and Information Systems Department, College of Management, University of Massachusetts Boston, 100 Morrissey Blvd, Boston, MA 02125-3393, United States

^b College of Business & Management, DeVry University, 3005 Highland Parkway, Downers Grove, IL 60515, United States

ARTICLE INFO

Article history:

Available online 12 August 2011

Keywords:

Professionalism
Computerized learning assessment
Learning outcomes
Latent semantic analysis (LSA)

ABSTRACT

This paper presents the results of an experimental study to measure professionalism for the purpose of assessing a professional development program. Soft skills such as professionalism are increasingly recognized as important, yet measuring and assessing these skills, typically best acquired experientially, has remained challenging. Following established research, we conceptualize professionalism as a construct with the dimensions of autonomy, commitment, belief in public service, self-regulation, and the use of a professional organization as a major referent. We demonstrate how these dimensions can be expressed and the professionalism of free-form text responses measured. These responses are reflections submitted online by participants in a professional development program for undergraduate business majors known as the Management Achievement Program (MAP). Latent semantic analysis is employed to measure the professionalism of these responses and to assess MAP along each of the five dimensions. The method demonstrated in this paper has several advantages over existing methods for assessment, which can be costly, require considerable time and training, and are often tied to subjective interpretation. The method demonstrated here is suitable for replication that leads to continuous improvement by “closing the loop.”

© 2011 Published by Elsevier Ltd.

1. Introduction

In an increasingly interdependent and global business world with more demand for employees having not only technical sophistication but a higher degree of professionalism, it follows that professional development programs have emerged that incorporate exposure to issues surrounding life-long self-directed professional competence and behavior. It also follows that educational institutions have recognized this need, which is reflected in the number of learning outcomes articulated for professionalism by colleges and universities alike.

The skills and knowledge that are considered key competencies are becoming increasingly complex, as is the need to assess them (Mayotte, 2010). Among these key competencies are “soft” skills, considered as “extremely important” to employers, and for which learning assessment and assurance is especially concerning (Beard, Schwieger, & Surendran, 2008). Better means of assessing “soft” skills is much needed and more complex than for “hard” skills, such as technical skills (Shuman, Besterfield-Sacre, & McGourty, 2005); research is still addressing this need. In computer science

education, for example, the literature “scarcely discusses how to assess” professional values (Fuller & Keim, 2008). The Accreditation Board for Engineering and Technology (ABET) lists specific professional skills as criteria for accrediting for bachelor’s degree programs (ABET, 2007; Siller, Rosales, Haines, & Benally, 2009). But as for robust and effective measures to assess many learning outcomes related to soft skills: “the literature remains sparse” (Shuman et al., 2005).

Soft skills, including those related to professionalism, are generally assessed using subjective methods, and because attitudes and values are integral to those assessments, behavioral observation is particularly well suited (Siller et al., 2009), and is widely used. Professional skills can be assessed from portfolio reviews, oral presentations, and interviews; these were all used to assess skills for continuous improvement, association with professional organizations, and an awareness of societal context by Hall and Bryant (2008). Assessments can be made by evaluating interviews, project work, presentations, and internships, as did Beard et al. (2008). Case studies and role playing have been used too; Fouch (2004) embedded assessments of professional skills in performance appraisals of a 6-week long tax audit case which required participants to assume specific roles.

Briedis (2002) developed an extensive set of rubrics to assess professional skills for an engineering program and incorporated them with performance appraisals of projects, presentations, and

* Corresponding author. Tel.: +1 617 287 7692.

E-mail addresses: roger.blake@umb.edu (R. Blake), ogutierrez@devry.edu (O. Gutierrez).

¹ Tel.: +1 508 638 8875.

written materials. These rubrics were used to assess skills related to an awareness of self-direction, life-long learning, association with professional organizations, initiative to pursue opportunities, and an understanding of the field in a global and societal context. Each of these is analogous to one of the dimensions of professionalism in the conceptual model of professionalism in our study.

These direct methods of skills assessment can be susceptible to influences by the assessment process itself. Many assessments of professional skills, including those cited above, are integrated with performance appraisals, and this other, primary consideration can have a confounding effect. To circumvent this issue Shuman et al. (2005) recommended the use of non-invasive, indirect methods. We follow this suggestion: the text responses used in this study were not used for performance appraisal, nor were the individuals entering those responses aware that those responses were integrated with learning outcome assessment.

Objective assessment methods can be less costly, more readily developed and administrated, and easier to interpret than subjective methods, and have also been used to assess professional skills. These methods include evaluating scores from standardized tests, the results of exit and professional certification exams, and analyzing job placement data. Scales and instruments designed and used to assess soft skills such as self-directed learning (Oddi, 1986) are also objective methods. Surveys are possibly the most often used objective method of assessment; examples include the assessments made by Beard et al. (2008), who surveyed supervisors of business school interns to assess professional skills gained during the internships, and those by Briedis (2002), who used surveys of students and alumni assess professional skills developed in an engineering program. However, objective methods are more limited than subjective methods and therefore not as well suited as for assessing soft skills (Mayotte, 2010).

The advantages and disadvantages of each form of assessment can lead institutions to what Mayotte describes as a “vicious cycle” (Mayotte, 2010). Mayotte notes that subjective assessments tend to be costly, and in addition to the training which is often required, they are complex and time consuming to design, set up, and administer. Subjective assessments require interpretation, and consequently might vary depending on the individuals conducting assessments, calling into question their consistency.

Mayotte observes that these problems can prompt the institution to adopt objective assessment methods because they are more easily interpreted, less costly, and quicker to administer. But because objective methods are not as well suited to assess in-depth or experiential knowledge, the organization is induced to move back towards subjective assessments, completing the cycle between these two forms of assessment (Mayotte, 2010).

Another difficulty with assessing soft skills such as professionalism may be related to the difficulty institutions have assessing learning outcomes in general: reluctance on the part of faculty (Pringle & Michel, 2007). A survey of 420 Deans at schools accredited by the Association to Advance Collegiate Schools of Business (AACSB) found that faculty were concerned that expectations to assess learning outcomes would be too complex, difficult to understand, or too costly. They had reservations that assessments would be too time-consuming or interfere with their course assessments already in place, and they expressed worry that assessments might have an impact on their performance reviews or impinge upon their academic freedom (Kelley, Tong, & Choi, 2010).

This reluctance possibly contributes to a degree of reliance on self-assessments. Self-assessments can be valuable for self-directed learning, such as integrated portfolios and career material in a learning management outcomes system (Ducrot, Miller, & Goodman, 2008). However, self-assessments are susceptible to inaccuracies. Researchers have demonstrated that it is precisely the individuals who lack the skills they are assessing who are

more likely to have the least accurate assessments (Kruger & Dunning, 1999). Other researchers have argued that accurate self-assessments are an “illusion” and advocate they be avoided for learning assessments (Eva & Regehr, 2008). Siller et al. (2009) incorporated self-assessments in a professional learning program with experiential learning and objectives similar to the program we study. However, they expressed that direct assessments of learning in place of self-assessments “must be implemented”, and noted their plan to discontinue self-assessments as soon as possible. It is equally important for us to avoid self-assessments in this study.

A fundamental reason why professionalism in particular has proven so difficult to assess may be the lack of agreement on any operational definition for it. Professionalism is related to context and because it has a broad meaning, the term itself can be ambiguous. As a result, the conceptualization of professionalism is both difficult and important (Watson, 2002).

This research draws on a conceptual model of professionalism and presents a method of measuring professionalism, and demonstrates how that method can be used to assess the effectiveness of a particular professional development program. We use a program of activities and events called Management Achievement Program (MAP) as the context and source for our study. MAP is a program designed to develop the professionalism of undergraduate business majors through an integrated series of diverse co-curricular activities and events and activities.

On completing an event or activity (for brevity hereafter referred to simply as “events”), participants prepare and enter reflections of their experience only as free-form text. The participants are not directed or instructed for this; their only guidelines are to summarize the event and what they learned. The collection of 1621 participants’ reflections (hereby referred to as “responses”) obtained over the 2-year period of this study forms the corpus of text for this research.

We use latent semantic analysis (LSA) as the means to measure professionalism for each response. The basis for measurement is provided by Hall’s (1968) classic taxonomy of attitudinal attributes of professionalism. That research identifies five dimensions: (1) Association, (2) Perceived Societal Benefits, (3) Self-Regulation, (4) Dedication, and (5) Autonomy. We use LSA to classify and assess the extent to which reflections provided by the participants correlate to the awareness and understanding of each of the above named dimensions.

This research contributes by defining how professionalism can be measured using an established conceptual model. It presents a viable method for measuring professionalism from sources of data presented in natural language. In MAP, the free-form text reflections submitted by participants satisfied several requirements (i.e. event feedback form, compliance requirement, source of program assessment) making a second instrument using a more traditional quantitative method of data collection an additional burden. In addition, the method allows for the use of the entirety of the data, as opposed to the use of sampling techniques. This research demonstrates a practical application of how a professional development program can be assessed and practical recommendations addressing specific dimensions of professionalism offered to program administrators.

The method demonstrated in this paper can mitigate some obstacles encountered in assessing soft skills. The initial costs are only those needed to establish “gold standard” responses, as described below, and to implement the method of analysis. For the purposes of the assessment conducted for this study, the method of analysis was largely automated. This means the method can be replicated and be expected to have relatively modest costs. The algorithm used in this method produces measurements that are more objective than subjective. Therefore, the benchmarks

established from an initial application can be compared with subsequent measurements. While this paper presents the results of the first time MAP has been assessed, the method demonstrated here is intended to be incorporated into a cycle of assessment, planning, implementation, and reassessment for the continuous improvement of this professional development program.

This paper is structured as follows. The next section reviews prior work related to professionalism, elaborating on how modern thought on professionalism has been incorporated into human development programs. The MAP program is presented as a comprehensive example of such a program, and is utilized to show how it addresses different constructs of professionalism. Following that, the methodology for this research is explained, first for latent semantic analysis and then for how it is used to measure participants' understanding of the constructs of professionalism from their responses. The results from analysis are then presented along with the criteria for measurement; this section demonstrates the feasibility of using free-form text analysis for assessing the effectiveness of professional development programs. Next is a detailed discussion of MAP's effectiveness, and the paper concludes with recommendations for further research.

2. Background

2.1. Professionalism

In order to assess a program for professional development it is first necessary to develop a precise definition of what is meant by "professional demeanor" and of the criteria to measure it. A traditional modern understanding of professionalism relates to the standards of behavior that are normally determined by a professional community (Evetts, 2003; Freidson, 2001). In this understanding, an individual will be deemed a professional to the extent he or she commits to an ideology or a set of values that guide the manner by which the person will conduct himself or herself in the execution of his or her duties. Here, the professional community, normally embodied in the form of professional association or professional organization, sanctions who has achieved the status by way of licensing, certification, or membership. Boland and Tempel (2004) argue that the conception of professionalism and how it is measured have evolved over time.

These perceptions have changed since the beginning of the last century when professionalism was associated primarily with a distinct service orientation; later with the use of standards and control over work; and more recently with formal expertise being the dominant feature of professionalism. Today, however, the emphasis and pressures of lifelong learning on the modern workforce has shifted further the understanding of professionalism to become that of continuous personal development in one's area of expertise (Friedman & Phillips, 2004). Here, both the attainment and sustainment of professionalism are a matter of individual commitment and competence (Wallace & Kay, 2008) and less a matter of compliance or association (Wailes, 2003). The role of the professional community also shifts from regulator to being a knowledge base anchor for its affiliates (Karseth & Nerland, 2007). The professional commitment model (Wallace & Kay, 2008) seems to incorporate elements of this evolving view of professionalism. In this model, individuals will demonstrate greater professionalism when they are able to exercise four role behaviors that are normally associated with the ideals of what being a professional means (Wallace, 2001): control over one's own work, a sense of service, the variety, and challenging nature of work, and a sense of community.

These views are consistent with our research requirements for assessing the effectiveness of a professional development program, as distinct from assessing the structural characteristics of

professions in business administration or information systems. Following this argument, our intention in this research was to document the participants' views of their experience with the events and activities of a specific professional development program. This was done with the objective of eliciting the extent to which the articulations of their experience would correlate with significant attitudinal characteristics associated with professionalism.

In a seminal study of professionalism, Hall (1968) identified five attitudinal attributes of the construct which reflects the manner in which individuals view their work. These five attributes were later validated by Snizek (1972), and are significantly similar to the role behaviors of the professional commitment model described by Wallace and Kay (2008). The multidimensional construct of professionalism is composed of the following attitudinal elements:

- *Association*. This is the use of a professional organization as a point of reference. This includes formal professional associations or organizations as well as the network of colleagues that provide a source of ideas, expertise, and reference.
- *Societal Benefits*. This is the sense of service, and a belief in the importance of the profession for the public and beyond oneself.
- *Self-Regulation*. This is the belief that a good form of control for the profession is through peers or colleagues, and a sense that such practice is beneficial.
- *Dedication to the profession*. This is a sense of commitment to excel in one's occupation regardless of the extrinsic rewards.
- *Autonomy*. This means having a sense of self-control over one's work responsibilities.

These are dimensions that are reflected in learning outcomes articulated by accreditation agencies. For example, ABET stipulates that engineering graduates should recognize the need for life-long learning (Dedication), function on multidisciplinary teams (Association or Self-Regulation), and understand the impact of engineering in a global and societal context (Societal Benefits).

These dimensions have also been implicitly operationalized by researchers who have found them significant influences on job satisfaction. For Information System (IS) personnel, for example, job satisfaction was found to be most highly related to job characteristics and professionalism (Chen, 2008). Among the elements comprising these were the desire to finish jobs perfectly and to take on new challenges (analogous to Dedication), good relationships with colleagues and compatibility with the organization (analogous to Association), the importance of the future direction set by top management (analogous to Self-Regulation), and exercising responsibility using one's own judgment (analogous to Autonomy).

2.2. The MAP professional development program

The impetus to integrate professionalism into the curriculum of business schools has accelerated starting in the early years of this decade (Trank & Rynes, 2003) due to a perceived de-professionalization of business education (Gioia, 2002; Gioia & Corley, 2002; Pfeffer & Fong, 2002). Emphasis on assurance of learning (Gardiner, Corbitt, & Adams, 2009; Kelley et al., 2010) and increased regulatory standards (Christie, 2009; Glennie, 2002; Johnson, 2010) have impelled many institutions to recognize professionalism as a key learning objective. As a result, the need to develop curricula and programs aimed at professionalism, and consequently the need to measure their success.

This research draws on a professional development program called the Management Achievement Program (MAP). MAP was designed as a mechanism to develop and enhance the professional demeanor of its participants. The rationale for introducing MAP was that professionalism can best be developed and demonstrated when participants engage in co-curricular activities, as opposed to

Table 1
MAP events and the dimensions of professionalism targeted.

Dimension	Meaning in the context of MAP	Types of MAP events or activities contributing to the dimension	
Reference or association	Students attend events or will be involved in activities that reinforce the values of their discipline or reinforce their identity as future professionals in their field of choice. They will participate in activities that contribute to their better understanding of their field of study	<ul style="list-style-type: none"> • Meet the firms • Company tours 	<ul style="list-style-type: none"> • Senior executive forums • Info sessions
Perceived Societal Benefits	Students engage in activities that reinforce the sense of value that their program of study has instilled in them, projecting themselves as future contributors to society and to their profession	<ul style="list-style-type: none"> • Volunteering • CM forums 	<ul style="list-style-type: none"> • Senior executive forms
Self-Regulation	Students participate in those MAP activities and events which they trust will involve people who can judge their performance or potential	<ul style="list-style-type: none"> • Master classes • Workshops 	<ul style="list-style-type: none"> • Career fairs • Meet the firms
Dedication/Sense of Calling	Students show commitment to the MAP program by behaving according to existing participation policies, and by actively pursuing engagement in one's field of choice	<ul style="list-style-type: none"> • Master classes • Company tours 	<ul style="list-style-type: none"> • CM forums • Senior executive forums
Autonomy	Students engage in MAP events that will contribute to their individual sense of professional realization	<ul style="list-style-type: none"> • Career fairs • Workshops 	<ul style="list-style-type: none"> • Seminars • Student clubs

3. Methodology

This section describes our methodology for measuring professionalism from free-form text MAP responses. Although a single aggregate measure of professionalism might be feasible, it would have limited value. Such a measure would lend little guidance to diagnose a program's specific strengths and weaknesses, for develop plans for improvement, or evaluate the results of implementing those plans. It is therefore essential that our methodology produce valid measures for each of the five dimensions of professionalism. It is for that purpose we turn to latent semantic analysis (LSA).

3.1. Latent semantic analysis

MAP responses are entered online by participants are the free-text reflections of specific activities or events they have recently attended. To evaluate the level of professionalism inherent in these responses it would be preferable, but highly impractical, to employ a panel of subject-matter experts to judge the participants' awareness of each dimension of professionalism for each response. LSA offers a similar capability, and is a technique that is widely used to represent knowledge and simulate human judgment (Landauer & Dumais, 1997). In applications similar to ours, it has been used for successfully predicting subjective ratings of essays by human readers (Landauer, Foltz, & Laham, 1998), and to produce ratings of essays that are well-correlated with ratings of the same essays made by expert graders (Foltz, 1996). Provided that a context of an analysis is given in the form of benchmarks, LSA can mimic human judgment and provide automated support as well, such as to determine whether a given e-mail is, or is not, a customer complaint (Coussement & Van den Poel, 2008).

For judging essays, the benchmarks for the context of analysis might be taken from a selected set of high-scoring essays; for e-mails, they might be e-mails deemed to be complaints. In our case, the benchmarks were responses selected as exemplifying a high level of professionalism along one of the dimensions. These benchmarks are also usually known as "gold standards."

LSA is used to measure the similarity of an individual document (or essay, e-mail, or in our case, response) with a "gold standard." We use LSA to measure the professionalism of each response against an established gold standard for each of the five dimensions. Fig. 2 provides an overview of how we used LSA to process the free-form text responses and produce five cosine scores for each, one for each dimension. These scores range from 0 to 1; where higher scores indicate more similarity with the benchmark gold standard. Our application of LSA and the measures produced

from it are very similar to other applications of LSA to analyze free-form text such as automated essay grading systems (e.g., as for essay grading developed by Duwairi (2006). Looking at the cosines produced by LSA for the responses shown in Fig. 2, the first and third measure highly for one of the dimensions, Association and Autonomy, respectively, while the second does not measure highly for any dimension.

LSA is an appropriate technique for emulating human judgments of professionalism because it can overcome many of the inherent difficulties an algorithm has inferring semantics from text. Many English words are ambiguous until the context of their usage makes it clear. For instance, the word "networking" appears frequently in MAP responses. In some cases, it could refer to networks of computers, in others to networks of professionals. The word "course", another word often used in responses, could refer to a course taken at a university, the course of a career, or simply part of the conjunctive "of course". LSA estimates the context of usage by analyzing co-occurrences of words. In the case of "networking", LSA might differentiate between a response in which "networking" appears along with "bandwidth" and "Internet", and a response in which "networking" appears along with the words "event" and "job". The context provided by the words "career", "direction", and "plans" appearing in the same response as "course" might enable LSA to differentiate that response from another in which "course" appears along with the words "grade" and "university". Conversely, LSA can also infer that two documents are dissimilar by a lack of co-occurrences. The measures of similarity produced by LSA compare favorably to measurements from other types of techniques such as those that are explicitly knowledge-based (Mihalcea, Corley, & Strapparava, 2006), as well as to other methods used for automated essay scoring (Dikli, 2006).

The data available to assess MAP consists of relatively short text responses entered online by participants. The average length of a MAP response is 53 words with 299 characters. Out of all responses, 23.5% have 25 or fewer words, and 89.9% have 100 or fewer. LSA is used often to analyze short texts such as these. For example, this technique was used to analyze texts from the abstracts of technical memos in order to select those that were most likely to be of interest to a particular reader (Foltz & Dumais, 1992). Foltz (1996) found LSA to compare favorably with human judgment to select the body of literature that individual sentences were most close to. Pilato, Pirrone, and Rizzo (2008) designed an intelligent tutoring system to assess how correct answers given by students through a "chat" type interface were. Using answers shorter than our responses, LSA measured how close each was to the correct answer, that measure determining the subsequent question to be directed to the student.

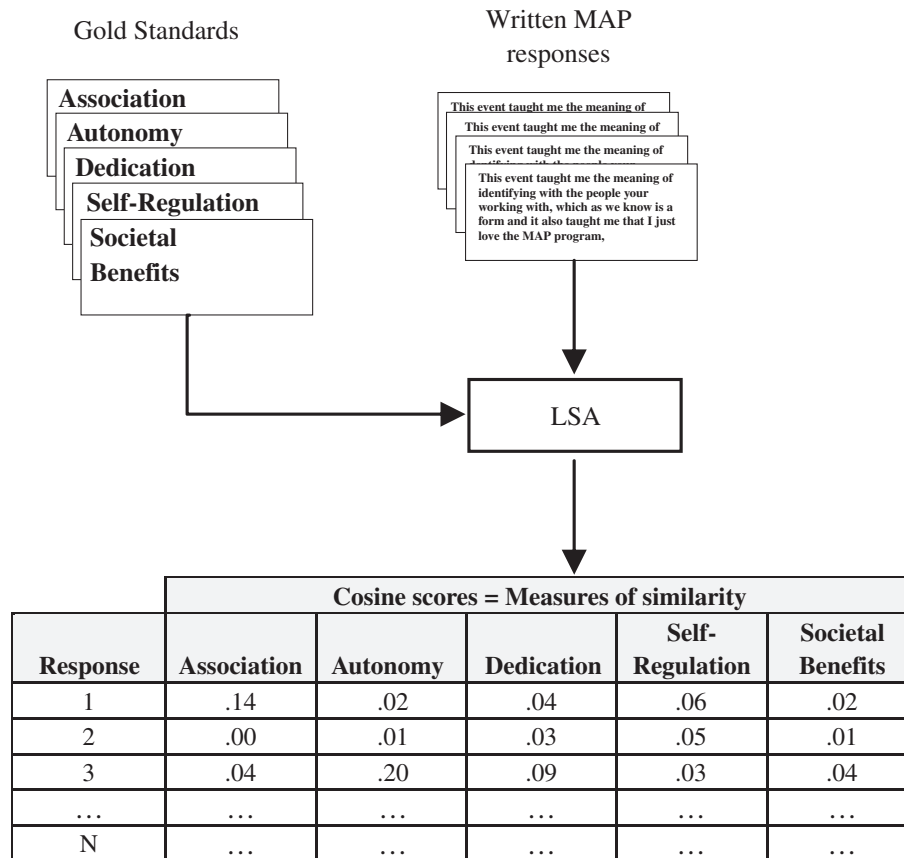


Fig. 2. Application of LSA to measure MAP responses.

In addition to its known ability to discern the context of short text documents, LSA is a very appropriate technique for our study because of its ability to evaluate the similarities between documents in a way consistent with judgments made by human experts.

3.2. Gold standards

The three studies cited above each measured the similarity between a single document and a gold standard derived from documents possessing a known quality, such as the correct answers in the case of the intelligent tutoring system. For MAP the gold standards were responses judged to represent excellence in a specific dimension of professionalism by a panel of three non-faculty industry experts with significant managerial experience. To qualify as a gold standard, a response needed a consensus from the panel. A total of 46 documents were thus selected as the gold standards. A representative gold standard response for each of the five dimensions of professionalism is shown in Table 2.

The sample gold standards shown in Table 2 exemplify professionalism in each of Hall's dimensions. For example, the author of the gold standard response selected for Association describes the value of connecting with supportive co-workers; this demonstrates the perceived importance of identifying with one's organization. Autonomy is related to an individual's ability to define and pursue a chosen career path; the sample gold standard response for Autonomy in Table 2 shows the respondent to recognize the importance of setting a career direction by identifying "the jobs I am interested in" as well as "the industries I am interested in". The advice to "love what you do" and "have the will to work hard" is from a participant who clearly sees the significance of Dedication. The recommendation to "Get opinions from as many people as you can" in the gold standard for Self-Regulation is an

acknowledgment of the need to accept the judgments of other professionals. Societal Benefits is a dimension that involves the belief that one's profession will produce benefits to society. In the sample gold standard response for that dimension the writer reflected on the interaction between ethical behavior in the business world and the quality of life in society at large.

With gold standards established for each dimension, LSA can measure the degree of similarity between any individual response and those gold standards as described next.

3.3. Using LSA to measure professionalism

Latent semantic analysis (LSA) works through singular value decomposition (SVD), a form of factor analysis. As a context for the discussion which follows, Fig. 3 illustrates the procedure for using LSA to produce measurements of similarity between responses and each of the five respective gold standards.

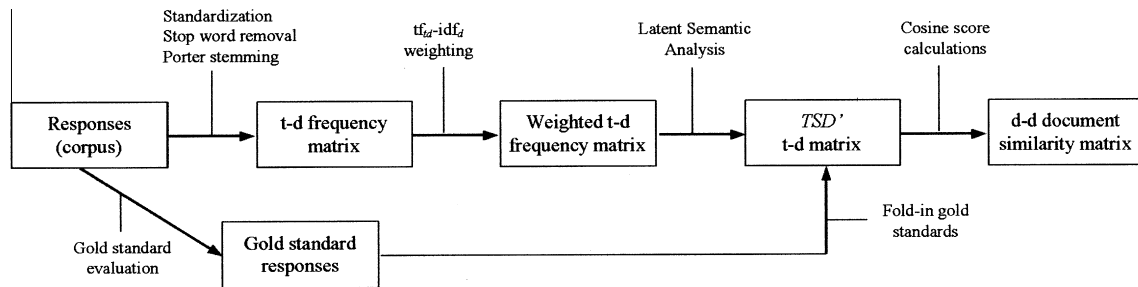
As shown in Fig. 3, the procedure begins with the main corpus of responses to be assessed and the set of responses defined to be the gold standards for each of the five dimensions. The same pre-processing steps are applied to standardize all responses, including the gold standard responses, after which LSA is applied to create a semantic space using the main corpus of responses. To avoid any confounding effects that could occur if the semantic space were to be built with both gold standard responses and the responses to be assessed, the gold standard responses are held separate and added into the semantic space after it is created. After these are added, a step known as "folding-in", the cosine measures of similarity between each response and the respective gold standards for each dimension of professionalism can be calculated directly.

The following discussion examines the entire process in more detail. That process starts by developing a term-document matrix

Table 2

Sample gold standard responses for each dimension of professionalism.

Dimension of professionalism	MAP outcome	Sample gold standard response
Association	Plan one's identity as a future professional	"I learned that identifying with a group can help you feel confident and supported in the work place. What I learned through this event can be directly applied to a career path, by taking advantage of employee resources, in order to connect with co-workers on a deeper personal level. It helps form stronger bonds"
Autonomy	Develop a sense of professional realization	"I have learned for myself that it is very important to know what I want to do in my life (what job I am interested in). I should identify what industries I am interested in and in what field (sales, advertising, investment). After I can start searching for companies that can offer me the job I am looking for and prepare to the interview as best as I can"
Dedication	Reinforce the importance of a sense of calling to one's field and one's career	"To take the opportunities that life and your working careers give you and make the best of it. Love what you do when it comes to working. Have the will to work hard and just because you come from a small town and background does not mean you cannot be a big success"
Self-Regulation	Integrate the judgments of other professionals in one's self-evaluation	"Get advice from multiple experts/professionals on what to include on your resume. Research what to include in resumes in related field of interest. Get opinions from as many people as you can on your resume"
Societal Benefits	Appraise one's values and attitudes towards their profession's contributions to society	"I have learned the importance of sustainability for business and society. Companies can become more ethical by improving the quality of life in society. This topic could be applied to my career path because environmental and ethical issues are addressed in the business setting every day. Managers also have to deal with sustainability risk management"

**Fig. 3.** Procedure to use LSA for measuring the similarity of responses and the gold standards for professionalism.

from the documents in a corpus. Each entry in a t by d term-document matrix represents the number of occurrences of term t in document d , and as Fig. 3 shows, separate term-document matrices are built for the main corpus of responses and for the gold standard responses. Prior to LSA, four pre-processing steps are applied to both matrices in the same way.

The first step is to standardize each response by removing proper names, numeric and any special characters, and converting the text to lower case. In the second step, stop words are removed. Stop words are commonly words such as “the”, “an”, and “is”. Because these are frequently occurring words likely to appear in many, if not most, documents, retaining them can artificially inflate the apparent similarity between documents without improving the ability to distinguish between them. The third step is to stem each word. Stemming transforms similar words to a common root, the transformed word is called a term. For example “apply”, “applies”, “applied”, and “applying” all would be stemmed to the root “appli”, and “application” would be stemmed to “applic.” For our analysis, we used the Porter stemming algorithm (Porter, 1993), which is an algorithm frequently employed for LSA. The fourth and final pre-processing step is to remove terms either less than three characters in length or those appearing less than five times throughout the corpus; these are terms that can influence the analysis but are less likely to contribute to a document's context. After these pre-processing steps were applied the resultant term-document matrix in the main corpus for our study consisted of 313 terms and 1621 documents.

LSA uses a term-document matrix as input and its effectiveness is well known to yield improved results if a weighting function is

applied to the terms in the raw term-document matrix first (Manning, Raghavan, & Schütze, 2008). Weighting functions are based on two factors. The first factor increases the weight of a term in proportion to the number of occurrences within a document. The second increases the weight of a term in inverse proportion to the number documents in which it appears. The logic for using weighting the number of documents in which a term appears inversely is the same as the logic for eliminating stop words; if a term appears in a high percentage of documents, it is very unlikely to add to the ability to develop a semantic structure to differentiate between the contexts of those documents.

The weighting function for our analysis is based on a binary transformation of a term's frequency within a document, and a logarithmic transformation of the number of times a term appears throughout all documents. More specifically, the weighting function used is:

$tf-idf_{t,d} = tf_{t,d}idf_t$ in which :

$tf-idf_{t,d}$ is the weight to be applied to element t,d in the raw term-document matrix,

$tf_{t,d} = 1$ if term t appears in document d , and 0 otherwise, and $idf_t = 1 + \log(N/df_t)$ where N is the number of documents in the corpus and df_t is the number of documents in which t appears. A logarithmic function is used to dampen the weighting for terms occurring in many documents throughout the corpus.

LSA, which is first applied directly on the weighted term-document matrix built from the main corpus of responses, uses singular value decomposition (SVD). SVD is a dimension reduction

technique similar to factor analysis and is used to reduce the term-document matrix to k dimensions. Decomposition produces three matrices, T , S , and D . The matrix S is a k by k matrix; T is a t by k matrix for terms, and D a k by d matrix for documents. The value of k is determined by the analyst and has a significant influence on the results of the analysis. The goal is to choose a value of k that is large enough to represent the underlying semantic structure of a corpus, but small enough to avoid overfitting that could be caused by any spurious terms present in the term-document matrix. There are several methods used to determine the value of k for dimension reduction. We use the well-known Kaiser criterion (Kaiser, 1960) which retains dimensions having eigenvalues of 1 or greater. A detailed description of LSA, SVD, and dimension reduction is provided by Landauer, Foltz, and Laham (1998).

The gold standard responses are next “folded-in” to the T , S , and D matrices; this step is done to avoid potential confounding effects that could occur if they were included in the SVD to begin with. With the addition of gold standard documents a new term-document matrix W is created by the products of TSD' and then used to measure similarities between documents. Cosine scores are used because they normalize the effects of different document lengths. The cosine between any two documents i and j can be calculated from column vectors i and j in W , and is defined as:

$$\frac{W_i W_j}{|W_i| |W_j|}, \text{ equivalent to: } \frac{\sum_t W_{ti} W_{tj}}{\sqrt{\sum_t W_{ti}^2} \sqrt{\sum_t W_{tj}^2}}$$

The cosines between each pair of documents are calculated and these values are used to create a d' by d' matrix M of cosine similarity measures (note that M includes folded-in gold standard documents and therefore $d' > d$). M is a symmetric matrix with values of 1 along the diagonal; higher cosine scores in M indicate greater document similarity. While M will have cosine scores for all document pairs, it is the cosine scores between each response document and the five gold standards that are of interest.

For this procedure we used R version 2.12.0 in conjunction with the LSA, Snowball, and RStem packages for the latent semantic analysis and calculations of cosine similarity scores, and PASW 18.0 for the statistical analysis which is discussed in the next section.

4. Analysis

4.1. Measurements of professionalism from LSA

Table 3 provides some examples of cosine scores produced by LSA which measure the similarity between an individual response and the gold standards the five dimensions. This table also provides quotes taken directly from each of the responses.

Table 3
Measures of professionalism for each of the five dimensions for selected responses.

Cosine similarity of response with dimensions of professionalism						
Response	Association	Autonomy	Dedication	Self-Regulation	Societal Benefits	Quote from response
2041	0.1883	0.0414	0.0968	0.0344	0.0102	“Notice the importance of building a network, even more important than your skills sometimes”
4074	0.0690	0.1425	0.1223	0.0236	0.0531	“This event helped me see how wide of a scope of jobs an MIS student can choose and made me want take more classes”
6277	0.0451	0.0220	0.1430	0.0258	0.1012	“I have learned that success has a direct relationship with hard work”
5536	0.0558	0.0000	0.0582	0.2322	0.0259	“I knew the interview process was a sales process but I did not realize how much you really have to sell yourself to the interviewer”
1928	0.0489	0.0274	0.0507	0.0531	0.2717	“I have learned that founding a firm does not only mean money but having a way where we can help the community”

Each of these selected responses has a high degree of professionalism along a dominant dimension of professionalism, as can be seen from their cosine scores in the results from semantic analysis. For instance, the participant who wrote the first response, 2041, offers the insight that networking can be more important than skills; networking is one way that Association, which involves establishing one's own professional identity, can be built. The writer of the second response showed an awareness that professionals can choose and direct their own careers, which related to the dimension of Autonomy, by stating an intention to choose among career and study options. Response 6277 reflects Dedication by a statement linking “hard work” directly to success. By understanding the importance of “selling yourself to the interviewer”, the writer of the fourth response, 5536, spoke to the importance of the judgment of other professionals, which corresponds to Self-Regulation. The author of the last response, 1928, demonstrates an awareness of the dimension of Societal Benefits through the realization of the opportunities for helping a community that can arise after forming a successful company.

4.2. Summary statistics

Table 4 presents descriptive statistics for the cosine scores produced by LSA for the 1621 responses in our analysis. This table shows that the lowest cosine scores are 0.0, indicating that there are responses having no similarity to one of the gold standards, which can be expected. The maximum scores are approximately .50, indicating responses with a high similarity to a gold standard and therefore a high measure of professionalism.

It is of note that the mean and maximum cosine scores (or measures) differ by dimension; the mean for Dedication (.0670) is more than double that of Autonomy (.0316). But before determining whether these differences might be significant, it is important to evaluate the validity of these measures, as done next.

4.3. Discriminant validity

As shown in Table 1 there are different types of MAP events and each is anticipated to develop professionalism along one, or

Table 4
Descriptive statistics for the cosine scores produced by LSA.

Dimension	N	Minimum	Maximum	Mean	Std. deviation
Association	1621	0.0000	0.3399	0.0640	0.0469
Self-Regulation	1621	0.0000	0.4822	0.0556	0.0483
Autonomy	1621	0.0000	0.3031	0.0316	0.0429
Dedication	1621	0.0000	0.3527	0.0670	0.0424
Societal Benefits	1621	0.0000	0.5203	0.0519	0.0482
All dimensions	1621	0.0000	0.5203	0.0540	0.0475

possibly two, of dimensions of professionalism. From only the descriptive statistics in Table 4, it is conceivable that if individual responses were examined that each type of MAP event could be equally effective along all five dimensions. As a result, taking the measures as a whole, the five measures of professionalism would all tend to vary together and be higher or lower for any given response. In that case, we would expect the measures to be well-correlated.

We would also expect the measures to be correlated if our methodology is unable to distinguish among dimensions. Another reason the measures for the five dimensions might vary together would be a lack of discriminant validity.

On the other hand, if our measures are valid and we are in fact measuring five distinct dimensions, we would expect that the measures from individual responses for the dimensions would not have a high degree of correlation. Correlation has long been known as a means to test discriminant validity (Campbell & Fiske, 1959), and for this purpose Table 5 shows the coefficients of correlation between each pair of dimensions from the corpus of 1621 responses.

This table shows that the correlations between dimensions are low; the average is .1371. The other nine correlations range from .030 to .189; all are under .20. The highest correlation is .49, which is possible if the degree of professionalism along two dimensions is raised by certain types of events that are offered frequently which is not infeasible. However, even this correlation is within guidelines established by researchers and we therefore conclude that our measures are valid and represent separate and distinct dimensions of professionalism. Given discriminant validity, we can proceed to examine the differences between dimensions.

4.4. Measurements of professionalism for each dimension

While the descriptive statistics presented in Table 5 show that the measures (as mean cosine scores) vary by dimension, they do not show whether differences are significant. If the means are not significantly different, one cause could be that our measures are not reliable and therefore cannot detect differences that in fact do exist. It could also mean that MAP has been equally effective for all five dimensions, although we would not expect this to be true. On the other hand, finding significant differences would mean the measures are reliable and the assessment valid.

We used a one-way ANOVA to test whether the mean scores between dimensions are different. If these differences exist, then planned comparisons can reveal which dimensions are most closely related. Finding associated dimensions would be desirable; it would suggest that efforts to improve MAP in one dimension may simultaneously yield improvements in another. The results from the ANOVA are shown in Table 6, and from these we conclude that there are significant differences in the mean cosine scores between dimensions ($N = 1621$, $p < .05$).

To proceed with planned comparisons we used the Tukey-B test. The results from these tests showed that there are three distinct homogeneous subsets, which are presented in Table 7. The validity tests, ANOVA, and planned comparisons demonstrate that professionalism can be measured along each of five established dimensions, and they reveal that professionalism in MAP has been

Table 6

One-way ANOVA to evaluate cosine score differences between dimensions of professionalism.

Source	Type III sum of squares	df	Mean square	F	p
Dimensions of Professionalism	24.90	5	4.98	2251.85	.000
Error	17.91	8100	.002		
Total	42.81	8105			

Table 7

Homogeneous subsets of the dimensions of professionalism ($\alpha = .05$).

Dimension of professionalism	N	Subset		
		a	b	c
Autonomy	1621	0.0316		
Societal Benefits	1621		0.0519	
Self-Regulation	1621		0.0556	
Association	1621			0.0640
Dedication	1621			0.0670

highest for Association and Dedication as a distinct subset separate from the next highest subset of Self-Regulation and Societal Benefits. With a mean score lower than either of these, Autonomy stands alone in its own subset.

Finding confirmation of significant differences, and that distinct subsets of dimensions exist, we turn to finding if there have been trends in the differences of those subsets over time.

4.5. Trends

While there are significant differences between dimensions, we do not know how they may have been changing over the period of this study. To examine this, we plotted the mean cosine scores for the three homogeneous subsets as a time series, as shown in Fig. 4. The plot shows the relative order among these three subsets has

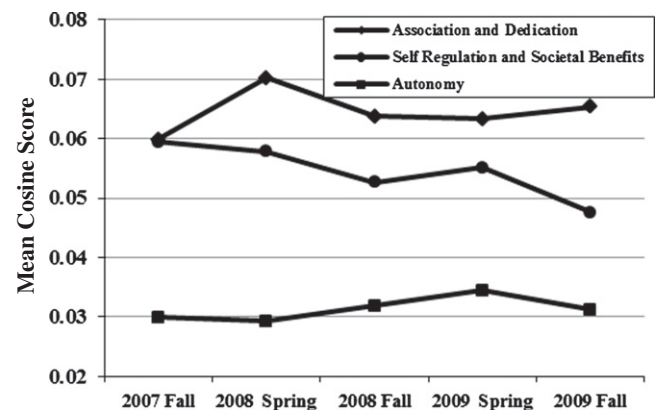


Fig. 4. Measures of professionalism for the three sets of homogeneous dimensions over time.

Table 5

Coefficients of correlation between measures (cosine scores) of professionalism dimensions.

Dimension	Association	Autonomy	Dedication	Self-Regulation	Societal Benefits
Association	1.000	.060	.164	.490	.044
Autonomy	.060	1.000	.189	.084	.041
Dedication	.164	.189	1.000	.058	.134
Self-Regulation	.490	.084	.058	1.000	.030
Societal Benefits	.044	.041	.134	.030	1.000

remained unchanged. It shows Association and Dedication to have consistently the highest measures. The next highest dimensions have been Self-Regulation and Societal Benefits, but for these two dimensions, measures have been declining and over the period of the study have decreased by some 20%. Autonomy is of particular concern because that dimension is not only perennially lowest, but the measures are roughly half those of the other four dimensions.

The statistically significant differences coupled with trends that seem evident led us to consider these questions:

1. With Autonomy consistently the dimension for which MAP has been the least successful as indicated by these measures, what can be done to improve professionalism along this dimension?
2. What should be done to reverse the decreasing sense of professionalism for Self-Regulation and Societal Benefits?
3. Although MAP has been consistently most successful for Association and Dedication, how could the sense of professionalism along these dimensions be reinforced, or even improved?

It should be noted that while statistically significant differences do exist between dimensions, and there are clear patterns in those differences over time, we have not established with certainty the existence of causal relationships. This research presents a methodology to measure professionalism and to assess a professionalism development program. A start to evaluating causality can be made by using this method to re-assess the program after suggestions from this research are selected, prioritized, and implemented. To develop those suggestions we expand on the results of this analysis with the following discussion.

5. Discussion

5.1. Assessing professionalism and MAP

Different types of MAP events are offered because each type is expected to increase participants' awareness of professionalism along one or more specific dimensions. For example, student groups are expected to increase a sense of identity and Association with a professional discipline; volunteer activities to increase participants' awareness of potential Societal Benefits in their careers. By examining different types of MAP events for individual dimensions, MAP can be assessed at a granular level and tangible recommendations offered to the institution. At this level of detail, recommendations can be made to target individual dimensions by actions such as changing the number, frequency, incentives, and contents of specific types of events, all of which are readily implemented.

Table 8 presents the measures of professionalism for each type of MAP event and dimension, expressed as mean cosine scores. The bottom line provides an indication of the overall success for each type of event, showing that in aggregate the most successful events have been career fairs, workshops, student clubs, and seminars, and that the least successful have been company tours, meet the firm events, and senior executive forums.

Table 8
Mean cosine scores from each type of MAP event for the five dimensions.

Dimension	Career fairs	CM forums	Company tours	Information sessions	Master classes	Meet the firms	Seminars	Senior executive forums	Student clubs	Volunteering	Workshops
Association	0.064	0.053	0.017	0.063	0.049	0.054	0.065	0.048	0.051	0.069	0.071
Self-Regulation	0.073	0.045	0.021	0.056	0.048	0.034	0.056	0.046	0.070	0.038	0.059
Autonomy	0.037	0.030	0.000	0.019	0.029	0.030	0.029	0.020	0.040	0.034	0.034
Dedication	0.062	0.075	0.136	0.062	0.067	0.047	0.062	0.046	0.062	0.068	0.069
Societal Benefits	0.061	0.060	0.027	0.040	0.055	0.036	0.061	0.049	0.052	0.054	0.048
All dimensions	0.059	0.053	0.040	0.048	0.050	0.040	0.055	0.042	0.055	0.053	0.056

The variation in mean scores across all event types is low, ranging from a high of .059 for career fairs to a low of .040 for meet the firm events. This suggests that no single type of event has dominated and that each has been at least somewhat effective.

However, the success of various types of events has varied considerably among the dimensions. For example, the lowest mean score for information system events is .0198 for Autonomy and the highest is .0636 for Association, more than a factor of three. A similar pattern is present for other types of events; differentials such as this indicate specific types of events have been more effective for some dimensions than for others. In the next sections we expand these results and bring forth suggestions for the institution consider to improve the effectiveness of MAP.

5.2. Assessing the effectiveness of MAP for dimensions of professionalism

Table 9 shows each of the different types of events and their effectiveness, as measured by their mean cosine scores, for Autonomy. The effectiveness of each type of event is categorized in one of three levels: high, medium, or low. An event is categorized as having a medium level of effectiveness if the mean score falls within 20% of the overall mean score for all events; as high if it exceeds that overall mean by 20%, and as low if it is 20% or more less than that mean. For example, the mean overall score for Autonomy across all events is .032. With a score of .019, information session events are more than 20% below that mean and therefore considered to have a low level of effectiveness. Student clubs, by having a score of .040, exceed that mean by 20% and therefore considered highly effective for this dimension.

In addition to arraying each type of MAP event by their effectiveness, the types of events expected to be most effective for each dimension are appear in italic font. The events so highlighted are those initially expected to be most effectiveness, and are the same as listed in Table 1. The same categorization is used for the other two homogenous subsets of dimensions which appear in Tables 11 and 12, with the same method of highlighting. Through understanding the effectiveness of each type of MAP event as compared to expectations, suggestions for improvement are developed in the following subsections.

5.2.1. Autonomy

The most effective types of events for Autonomy were student clubs, career fairs, volunteering, and workshops, as shown in Table

Table 9
MAP events with high, medium, and low levels of effectiveness for Autonomy.

	High	Medium	Low
Autonomy	<i>Student clubs</i> Career fairs	CM forums, meet the firm, master classes, <i>Seminars</i>	Information sessions Senior exec forums
	Volunteering <i>Workshops</i>		Company tours

9. Student clubs were anticipated to be among the most effective for Autonomy, as shown in Table 1 and the results show they were. Of all types of events, student clubs require particular initiative because they are largely self-organized and self-directed, reflecting Autonomy.

Workshops met the expectation that they would be among the most effective events for increasing the sense of Autonomy. But because Autonomy is of particular concern, we examined individual responses from workshops and found them to have considerable variability depending on the nature of the workshop. Workshops held for exam preparation, such as for the CPA exam, tended to have lower measures for this dimension and therefore less effectiveness. On the other hand, workshops focusing on career preparation, entrepreneurship, and networking had high measures and were quite effective. Some responses reflected how these workshops led respondents to be “more assured that my idea can be developed into a small business”, know “how to meet and connect with people to build a career to opportunity”, and understand the importance of “building a career network.” Finding that particular workshops may influence an awareness of Autonomy suggests that the institution consider increasing the number of workshops focusing on that content.

The scores for Autonomy were consistently much lower than for all other dimensions. More alarming is that out of the five dimensions, Autonomy has the lowest mean scores for every individual type of MAP event, although relative to others, several types of events were shown to have more effective. Without a methodology capable of assessing MAP at this level of detail, these facts would have remained unknown and the opportunity to improve professionalism in this dimension unrealized.

One strategy to improve Autonomy might be to increase the incentive for participation in the types of events found to yield the highest measures, such as career fairs or volunteering. One readily available means to do this might be to increase the award of MAP miles to participate in these types of events. To investigate whether or not this might be successful we analyzed the correlations between the MAP miles awarded for each type of MAP event (generally the award is based on the type of event) and the measures of professionalism from responses. The results are presented in Table 10 and shows there is little or no correlation between the MAP miles awarded and professionalism along four of the five dimensions. This provides little evidence that increasing MAP mile awards would improve professionalism for those four dimensions. Interestingly, Autonomy stands out as the one dimension for which there is a significant correlation between MAP miles and professionalism, a result we did not expect. However, correlation does not demonstrate causality. But it does suggest that the institution test the effects of higher incentives for the four types of events

most effective for Autonomy and then measure the results by replicating the method presented in this paper.

Career fairs and volunteering were also very effective for Autonomy, which was unexpected. Both career fairs and volunteering may reflect Autonomy by the independence and initiative needed for participation. Therefore, in retrospect the positive results for these two types of events might have been anticipated.

This suggests that the institution consider increasing the number of career fairs; these are held relatively infrequently. It also suggests that in addition to increasing MAP miles as an incentive, the institution consider increasing support for the formation and sustainability of student clubs and volunteer activities. This can be achieved by initiatives to facilitate access to professional organizations, involving more faculty members, and raising the monetary support for the formation and sustenance of professional clubs. In addition, enhanced tutoring and peer advising processes would seem appropriate measures to increase the sense of professionalism for the dimension of Autonomy.

5.2.2. Self-Regulation and Societal Benefits

Table 11 shows the events having a high, medium, and low effectiveness for Self-Regulation and Societal Benefits, both in the same homogeneous subset, with the same categorization of events as defined earlier. Over the period of this study, the measures of professionalism for these two dimensions have consistently declined and are a matter of concern.

Self-Regulation means that an individual is able to integrate the judgment of others into their own self-assessment. Career fairs were expected to contribute to this awareness most because participants at career fairs will often be evaluated “on the spot” by company representatives who determine whether or not they might be suitable for a particular company or career. The expectation for the effectiveness of career fairs was borne out in the analysis.

For student clubs, however, results ran counter to the expectations that these would be among the most effective for Self-Regulation. Many of the responses from student club events that did have higher scores for Self-Regulation were from events held for a fraternal order and from a club named “Students in Business”. Participants described how others at these events were “pointing me in the right direction,” and showing me that “asking questions is very important.” These are indications of an awareness of Self-Regulation. It suggests the institution might consider increasing the sense of this dimension by a higher level of support for these clubs. Student clubs were also found effective for Autonomy, indicating that more support might simultaneously benefit both dimensions.

Another type event with results counter to expectations was the meet the firm events, which were found to contribute little to

Table 10

Correlation of cosine scores with MAP miles, $N = 1621$.

	Association	Self-Regulation	Autonomy	Dedication	Societal Benefits
Correlation with MAP miles	−0.0763	−0.0395	0.1350	−0.0268	0.0315

Table 11

MAP events with high, medium, and low levels of effectiveness for Self-Regulation and Societal Benefits.

	High	Medium	Low
Self-Regulation	Career fairs Student clubs	Workshops, information sessions, seminars, master classes, senior exec forums, CM forums	Meet the firm Company tours Volunteering
Societal Benefits	Career fairs Seminars CM forums	Master classes, volunteering, student clubs, senior executive forums, workshops, information sessions	Meet the firm Company tours

Self-Regulation. Reviewing the meet the firm events that were held, many involved company representatives who gave presentations of their companies. This differed from career fairs, at which company representatives were more focused on qualifying candidates. Self-Regulation involves accepting the judgment of other professionals, which seems to explain the disparity between these two types of events. This suggests the institution might restructure the content of meet the firm events to include more participant interaction with company representatives to bring a focus more specifically on the Self-Regulation dimension of professionalism.

The high effectiveness of career fairs and seminars for Societal Benefits was a surprise. After examining participants' responses from career fairs, this result seemed accounted for by company representatives who, in many cases, had discussed their company's commitments to social responsibility and diversity in the workplace, possibly as inducements for attendees to seek employment with their firms. This result was neither intended nor anticipated, and while positive, we do not believe it is guaranteed to continue in future career fairs. We still suggest that the institution increase the number of career fairs, but for the purpose of increasing the sense of Self-Regulation and Autonomy instead. We suggest the institution should expect that any increase in the sense of Social Responsibility associated with career fairs to be incidental.

Seminars were found to be very effective for Societal Benefits. By reviewing the topics from individual seminars, it became clear that a range of issues was addressed, but most were not related to Societal Benefits, such as seminars for "IRS Info Sessions". But for Societal Benefits we found that the distribution of scores was highly skewed; the coefficient of skewness was 2.52. These led us to suspect that a few specific seminars happened to have a focus leading to more awareness of Societal Benefits. This suspicion was confirmed by finding that responses scoring highest for this dimension were predominately from "Faculty Brown-Bag Lunches." We found the topics discussed by the faculty related directly to Societal Benefits, and consequently found respondents who wrote that "we do have a social responsibility," and that "it is important to recognize the needs of the local community," and that doing so "is also good business and creates a stable community". This indicates that a majority of seminars may not be useful to increase the awareness of Societal Benefits, but that seminar topics focused on this dimension can be quite effective.

CM forums involve large and formal presentations made by top executives and often address topics related to Societal Benefits, and the effectiveness of those events was anticipated. This was readily apparent from responses emphasizing "how crucial it is for companies [to] incorporate environmental sustainability," and the importance that a company "meets the needs of present generations without compromising future generations".

These findings suggest that both seminars and CM forums can be most effective at increasing an awareness of the Societal Benefits dimension of professionalism, but that the effectiveness will be highly topic-dependent. They suggest that the inclusion of ethical and social responsibility issues by experts and senior executives will increase the participants' awareness of Societal Benefits. The institution might increase the number of seminars focusing on these issues, or request them to be integrated with executive-related events.

5.2.3. Association and Dedication

Association and Dedication are the dimensions for which MAP has been the most effective: over the period of this study, the measures for these have been consistently the highest of all dimensions, and even appear to have increased. Those events with high, medium, and low effectiveness are shown in Table 12.

The types of events anticipated to improve the sense of Association most, information sessions, meet the firm events, and senior executive forums, were all moderately successful. However, a quite unexpected result was that workshops and volunteer activities were among the most effective of all. The dimension of Association involves reinforcing one's identity as a professional. Upon reviewing responses, it became clear that workshops involving job and career preparation were requiring participants to think through issues related to their professional identity, offering an explanation for this result.

For example, several responses with high measures of Association were from a business etiquette workshop and these described how the workshop would help with becoming a "better fit into the company" and recognize the importance of considering "cultural and potential fit" when applying for a job. Participants in a corporate recruiting workshop explained that it would help them think about how to respond to a request often heard in interviews, which is to "Tell me about yourself." Answering this well requires forethought about one's own professional identity in an organization, i.e. Association.

Volunteer activities were also effective for Association, which was unexpected. To examine further, responses revealed that, like seminars for Societal Benefits, specific volunteer activities were most effective. Responses from volunteer activities for non-profit organizations such as the National Foundation for Teaching Entrepreneurship scored particularly highly. Respondents indicated that through their involvement they were defining a sense of professional identity, as demonstrated by participants writing about their desire to "work with nonprofit organizations" and to become "involved with entrepreneurship".

In addition to Association, volunteer activities were also found to be relatively effective for Societal Benefits, and the responses with high scores for volunteer work in non-profit organizations may be the link between the two. It suggests that volunteer work in non-profits may have synergistic effects for both dimensions, and that the institution look for more opportunities to bring non-profit volunteering into MAP. After doing so, the institution might evaluate whether or not the understanding of professionalism in both of these dimensions did in fact increase.

Based on these results we speculate that internships would stimulate a sense of Association in much the same way as volunteer work would for non-profit firms. Internships are not included as MAP events at present, but the results suggest that the institution consider a mechanism to incorporate them into MAP and through later reassessment determine if they do in fact contribute to Association.

The expectations for the most effective events for Dedication were generally met; CM forums were highly effective and master classes relatively so. However, senior executive forums did not meet expectations. This may be accounted for by the forums'

Table 12
MAP events with high, medium, and low levels of effectiveness for Association and Dedication.

	High	Medium	Low
Association	Workshops Volunteering	Seminars, career fairs, <i>information sessions, meet the firm</i> , CM forums, student clubs,	<i>Company tours</i>
Dedication	<i>Company tours</i> CM forums	CM forums, workshops, volunteering, <i>master classes</i> , information sessions, career fairs, <i>seminars</i> , student clubs	Meet the firm <i>Senior executive forums</i>

topics; as discussed earlier, the agendas of senior executive forums were often related to Societal Benefit and these forums were more effective for that dimension. This suggests that if the format and content of senior executive forums remains the same that the institution should expect those forums to be effective for the Societal Benefits dimension but not for Dedication.

Company tours were expected to bring forth a sense of Association through first-hand exposure to a firm, and also a sense of Dedication because of the investment of time and effort required for participation. However, for each of these two dimensions company tours had opposite levels of effectiveness: highly effective for Dedication, and counter to expectations, the least effective of all events for Association.

In fact, for the four dimensions other than Dedication, company tours were among the least effective of all events. This suggests that the institution evaluate company tours with an eye towards either modifying their format, or possibly decreasing their number in favor of offerings demonstrated to be more effective such as workshops, student clubs, seminars, and forums as described above.

By using LSA for analyzing free-form text responses, we were able to develop valid and validate measures of professionalism and to assess the effectiveness of MAP at a granular level for the five dimensions of professionalism. The results of analysis showed that the effectiveness of different events met some of our initial expectations and held some surprises. By examining individual responses, we found plausible reasons why this was so.

This assessment provides suggestions for the institution to consider for targeting specific dimensions. By using this method for subsequent reassessment, the institution can evaluate the impact of the suggestions it chooses to implement. This can lead to a cycle of continuous improvement, moving from assessment, recommendations for change, implementation, and re-assessment as the institution moves towards fulfilling MAP's primary goal of increasing the sense of professionalism for participants in the program.

6. Summary and further research

This paper presents a study and method for evaluating the effectiveness of a professional development program and participants' understanding of five key dimensions professionalism. Those five dimensions are from Hall's classification of professionalism (Hall, 1968) which identifies five attitudinal dimensions: Association, Societal Benefits, Self-Regulation, Dedication, and Autonomy. Each of these five was measured from free-form text responses entered online by program participants.

The program assessed in this research is the Management Achievement Program (MAP), a program with the purpose of developing the professionalism of a university's undergraduate business students through an extensive series of events and activities. The assessment was conducted from an analysis of 1621 entered online as reflections participants' experiences at individual MAP events. The method for measurement and assessment we demonstrate is based on latent semantic analysis, a technique for comparing natural language text documents and measuring the degree of professionalism for MAP responses. This method makes the assessment of tacit knowledge and the assessment of an intangible learning known to be difficult to measure possible. It is less time and labor intensive than several existing methods, does not require training to administer, and can be consistently replicated.

The semantic analysis from this research reveals three significant findings about the program's effectiveness. First, there are clear indications of the dimensions for which the program has been the most, and for which it has been the least, effective. Second, it discerns distinct trends in these dimensions over the 2 year period of this study. Third, it uncovers which types of events and activities

have been most and least effective for this program. From insights developed from these findings, several recommendations are offered to the institution to improve MAP's effectiveness.

The analysis and assessment shows that MAP has been most effective for the dimensions of Association and Dedication followed by the dimensions of Self-Regulation and Societal Benefits. They also reveal Autonomy to be where MAP has been least effective. The results indicate that over the 2 year period of the study, MAP's effectiveness for Association and Dedication have been consistently high, while its effectiveness for Self-Regulation and Societal Benefits appears to have been declining somewhat. Autonomy has consistently remained the dimension for which MAP has been least effective.

Individual types of events were also measured and assessed for their effectiveness, and based on the insights gained several specific recommendations were offered to the institution. The method and assessment demonstrated in this research show how it is possible to use a semantic approach to assess a professional development program, identify its strengths and weaknesses, and enable program administrators to target specific areas for improvement. After the administrators choose to put suggestions into place, the method demonstrated in this research can be re-applied and used to "close the loop" in a cycle of continuous improvement.

A limitation of this research is that this study was conducted for only one professional development program aimed at undergraduate students. Further research could replicate the method presented in this paper for comparison with other programs. Studies using Hall's scale to measure professionalism have been most successful in North America, and at least one study found conflicting results when it was applied in the UK (Swales, 2003). Extending this study to other cultures and types of programs would also be useful further research.

Beyond assessing the results of changes the institution makes from the suggestions and recommendations in this study, further research is planned to map individual documents to other performance indicators such as attendance patterns and grade point average records. This will help to obtain a more comprehensive view of MAP's effectiveness in light of participant profiles, and to make further recommendations towards reaching the goals of this professional development program.

References

- ABET (2007). *Criteria for accrediting applied science programs*. Applied Science Accreditation Commission.
- Baker, M. J. (2009). Personal professional development. *Marketing Review*, 9, 99–113.
- Beard, D., Schwieger, D., & Surendran, K. (2008). Integrating soft skills assessment through university, college, and programmatic efforts at an AACSB Accredited Institution. *Journal of Information Systems Education*, 19, 229–240.
- Boland, H. G., & Tempel, E. R. (2004). Measuring professionalism. *New Directions for Philanthropic Fundraising*, 43, 5–20.
- Briedis, D. (2002). Developing effective assessment of student professional outcomes. *International Journal of Engineering Education*, 18, 208–216.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81–105.
- Chen, L. (2008). Job satisfaction among information system (IS) personnel. *Computers in Human Behavior*, 24, 105–118.
- Christie, R. S. (2009). Corporate social responsibility & SH&E regulatory compliance. *Professional Safety*, 54, 20–21.
- Coussement, K., & Van den Poel, D. (2008). Integrating the voice of customers through call center emails into a decision support system for churn prediction. *Information & Management*, 45, 164–174.
- Dikli, S. (2006). An overview of automated scoring of essays. *The Journal of Technology, Learning, and Assessment*, 5.
- Ducrot, J., Miller, S., & Goodman, P. (2008). Learning outcomes for a business information systems undergraduate program. *Communications of the Association for Information Systems*, 23, 95–122.
- Duwairi, R. M. (2006). A framework for the computerized assessment of university student essays. *Computers in Human Behavior*, 22, 381–388.
- Eva, K. W., & Regehr, G. (2008). 'I'll never play professional football' and other fallacies of self-assessment. *Journal of Continuing Education in the Health Professions*, 28, 14–19.

- Evetts, J. (2003). The sociological analysis of professionalism. *International Sociology*, 18, 395.
- Foltz, P. W. (1996). Latent semantic analysis for text-based research. *Behavior Research Methods*, 28, 197–202.
- Foltz, P. W., & Dumais, S. T. (1992). Personalized information delivery: An analysis of information filtering methods. *Communications of the ACM*.
- Fouch, S. R. (2004). Developing functional and personal competencies through an interactive tax research case study. *Journal of Accounting Education*, 22, 275–287.
- Freidson, E. (2001). *Professionalism: The third logic*. Chicago, IL: The University of Chicago Press.
- Friedman, A., & Phillips, M. (2004). Continuing professional development: Developing a vision. *Journal of Education and Work*, 17, 361–376.
- Fuller, U., & Keim, B. (2008). Assessing students' practice of professional values. *ACM SIGCSE Bulletin*, 40, 88.
- Gardiner, L. R., Corbitt, G., & Adams, S. J. (2009). Program assessment: Getting to a practical how-to model. *The Journal of Education for Business*, 85, 139–144.
- Gioia, D. A. (2002). Business education's role in the crisis of corporate confidence. *The Academy of Management Executive* (1993), 16, 142–144.
- Gioia, D. A., & Corley, K. G. (2002). Being good versus looking good: Business school rankings and the Circean transformation from substance to image. *Academy of Management Learning & Education*, 1, 107–120.
- Glennie, J. A. (2002). Sarbanes-Oxley will dramatically alter relationships with tax service providers. *The Tax Executive*, 54, 362–364.
- Hall, R. H. (1968). Professionalization and bureaucratization. *American Sociological Review*, 33, 92–104.
- Hall, S. P., & Bryant, V. C. (2008). Assessment of professional skills for students in computing and engineering programs. *Journal of Computing Sciences in Colleges*, 23, 146–152.
- Johnson, A. (2010). When the dust settles compliance is left to cope. *ABA Banking Journal*, 102, 6.
- Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement*.
- Karseth, B., & Nerland, M. (2007). Building professionalism in a knowledge society: Examining discourses of knowledge in four professional associations. *Journal of Education and Work*, 20, 21.
- Kelley, C., Tong, P., & Choi, B. (2010). A review of assessment of student learning programs at AACSB schools: A Dean's perspective. *Journal of Education for Business*, 85, 299–306.
- Kolb, A. Y., & Kolb, D. (2009). Experiential learning theory. In S. Armstrong & C. V. Fukami (Eds.), *The SAGE handbook of management learning, education, and development*. London: SAGE Publications, Ltd.
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77, 1121–1134.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104, 211–240.
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). Introduction to latent semantic analysis. *Discourse Processes*, 25, 259–284.
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, 25, 259–284.
- Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge, UK: Cambridge University Press.
- Mayotte, S. (2010). Online assessment of problem solving skills. *Computers in Human Behavior*, 26, 1253–1258.
- Mihalcea, R., Corley, C., & Strapparava, C. (2006). Corpus-based and knowledge-based measures of text semantic similarity. *Proceedings of the National Conference on Artificial Intelligence* (p. 775).
- Oddi, L. F. (1986). Development and validation of an instrument to identify self-directed continuing learners. *Adult Education Quarterly*, 36, 97–107.
- Pfeffer, J., & Fong, C. T. (2002). The end of business schools? Less success than meets the eye. *Academy of Management Learning & Education*, 1, 78–95.
- Pilato, G., Pirrone, R., & Rizzo, R. (2008). A Kst-based system for student tutoring. *Applied Artificial Intelligence*, 22, 283–308.
- Porter, M. F. (1993). An algorithm for suffix stripping. *Program: Electronic Library and Information Systems*, 14, 130–137.
- Pringle, C., & Michel, M. (2007). Assessment practices in AACSB-accredited business schools. *The Journal of Education for Business*, 82, 202–211.
- Shuman, L. J., Besterfield-Sacre, M., & McGourty, J. (2005). The ABET 'professional skills'—Can they be taught? Can they be assessed? *Journal of Engineering Education*, 94, 41–55.
- Siller, T. J., Rosales, A., Haines, J., & Benally, A. (2009). Development of undergraduate students' professional skills. *Journal of Professional Issues in Engineering Education and Practice*, 135, 102.
- Snizek, W. E. (1972). Hall's professionalism scale: An empirical reassessment. *American Sociological Review*, 37, 109.
- Swales, S. (2003). Professionalism: Evolution and measurement. *The Service Industries Journal*, 23, 130–149.
- Trank, C. Q., & Rynes, S. L. (2003). Who moved our cheese? Reclaiming professionalism in business education. *Academy of Management Learning & Education*, 2, 189–205.
- Wailes, S. (2003). Professionalism: Evolution and measurement. *The Service Industries Journal*, 23, 130–149.
- Wallace, J. E. (2001). Explaining why lawyers want to leave the practice of law. *Sociology of Crime Law and Deviance*, 3, 117–145.
- Wallace, J. E., & Kay, F. M. (2008). The professionalism of practising law: A comparison across work contexts. *Journal of Organizational Behavior*, 29, 1021–1047.
- Watson, T. (2002). Professions and professionalism: Should we jump off the bandwagon, better to study where it is going. *International Studies of Management and Organization*, 32, 93–105.