Intl. Trans. in Op. Res. 14 (2007) 373–394

INTERNATIONAL TRANSACTIONS IN OPERATIONAL RESEARCH

MBNQA criteria in education: assigning weights from a Malaysian perspective and proposition for an alternative evaluation scheme

Rafikul Islam

Department of Business Administration, Kulliyyah of Economics and Management Sciences, International Islamic University Malaysia, Jalan Gombak, 53100 Kuala Lumpur, Malaysia
E-mail: rislam@iiu.edu.my

Received 29 June 2005; received in revised form 25 November 2005; accepted 11 January 2006

Abstract

In order to improve quality and productivity among American companies, the Malcolm Baldrige National Quality Award (MBNQA) was launched by the National Institute of Standards and Technology (NIST) under the US Department of Commerce in 1987. Over the years, the award has proved to be effective in improving companies' market share, customer satisfaction, employee morale, and also profitability. MBNQA has been a 'role model' in developing a national quality award for many other countries in the world. Furthermore, for organizational self-assessment, the criteria framework of MBNQA has been in use throughout the world. Presently, the award is offered in three categories: Business, Education, and Health care. The present paper is concerned with MBNQA in Education. NIST has developed a comprehensive set of criteria to be fulfilled in order to be eligible to win the award. However, in the existing literature, it is not clear how the weights are assigned to the criteria and subcriteria. The present paper uses the analytic hierarchy process (AHP) to reassign criteria weights from a Malaysian perspective. Furthermore, the paper points out the fallacy of the present evaluation scheme and proposes an alternative one based upon the absolute measurement procedure of AHP. The modified scheme is expected to enhance the fairness of the evaluation of the award aspirants. The paper also shares the experiences gathered in data collection using AHP.

Keywords: Malcolm Baldrige National Quality Award; total quality management; analytic hierarchy process; evaluation

1. Introduction

The Malcolm Baldrige National Quality Award (MBNQA) was launched by the US Department of Commerce to enhance competitiveness among American companies. The objectives of launching the award are as follows: (1) to recognize the companies who are doing an excellent job

Published by Blackwell Publishing, 9600 Garsington Road, Oxford, OX4 2DQ, UK and 350 Main St, Malden, MA 02148, USA.

in quality management, (2) to increase awareness of quality as an important element in competitiveness, (3) to share information on successful quality strategies and on the benefits derived from implementation of these strategies, and (4) to promote understanding of the requirements for quality excellence. The National Institute of Standard and Technology (NIST), an agency of the US Department of Commerce, manages the program and the American Society for Quality (ASQ) assists NIST. The board of examiners consists of noted quality experts who are selected from across the United States. For 2005, the board has about 540 members of which 10 serve as Judges and approximately 110 serve as Senior Examiners and the remainder serve as Examiners. During the period 1988–1998, MBNQA was awarded to only three types of business companies, namely Manufacturing, Service, and Small Business. In 1999, the Health Care and Education sectors were added to the categories of award. Up to three awards may be given annually in each of the five areas. However, if the performance is not up to the desired level, it may be the case that no award is given to any one or multiple areas (Babicz, 2002).The present paper pertains to MBNQA in the Education sector. Two major components of MBNQA in all the three categories are the criteria framework and the selection procedure. We address both components in this paper. The criteria set for Education (henceforth referred to as criteria) are built upon a set of core values and concepts: visionary leadership, learning-centered education, organizational and personal learning, valuing faculty, staff, and partners, agility, focus on the future, managing for innovation, management by facts, social responsibility, focus on results and creating value, and systems perspective (Baldrige National Quality Program, 2005a). The core values and concepts are embodied in seven categories of criteria: leadership, strategic planning, student, stakeholder and market focus, measurement, analysis and knowledge management, faculty and staff focus, process management, and organizational performance results. Each of these categories is subdivided into a number of items. At present, altogether, there are 19 items, each focusing on a major requirement. Items consist of one or more areas to address. Organizations applying for the Baldrige award are required to address their responses to the specific requirements in those areas. For completeness, a brief description of the criteria is provided in the next section.

2. MBNQA criteria in Education

The Baldrige criteria provide a system perspective (shown in Fig. 1) to manage organizations leading to performance excellence. The system perspective includes leaders' focus on strategic directions and students and other various stakeholders. Furthermore, senior leaders monitor, respond to, and manage performance based on organizational performance results. System perspective also includes various linkages among key processes and aligning the resources to improve overall performance and satisfy students and stakeholders' needs. MBNQA criteria are revised periodically. The initial set of criteria in 1988 had 62 items with 278 areas to address. By 1991 the criteria set had only 32 items and 99 areas to address. The 1995 criteria were revised to 24 items and 54 areas to address. The revision carried out in 1997 has produced a reasonable stability to the criteria set by having 20 items with 30 areas to address that are necessary to compete in today's marketplace, improve the linkage between process and results. In 2005, the criteria set has 19 items with 32 areas to address. In the following, we discuss the seven major categories of criteria adopted in 2005.

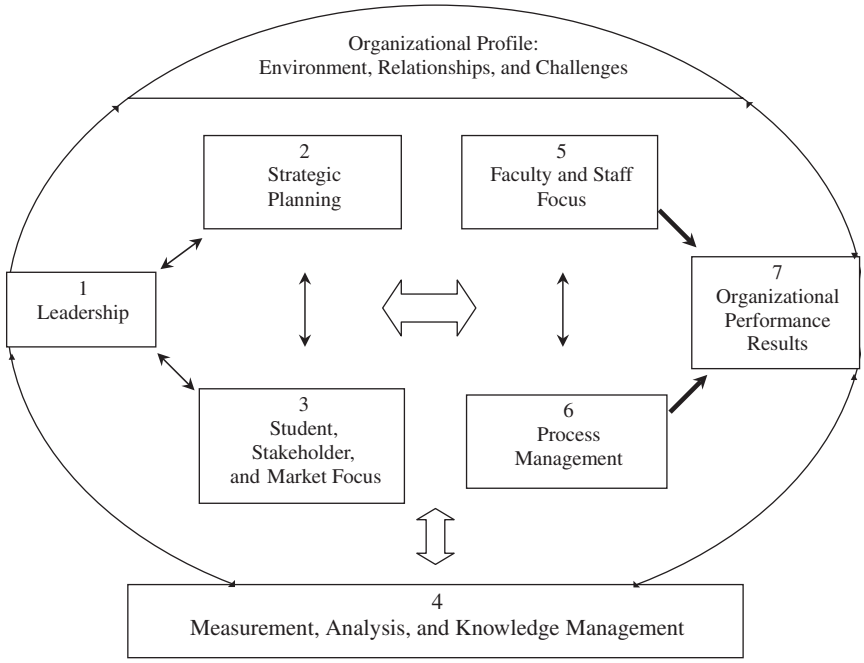

Fig. 1. Interrelationships among various categories of criteria of the Malcolm Baldrige National Quality Award.

2.1. Leadership

This category addresses how senior leaders guide and sustain the organization, set and communicate the organization's vision, values, and performance expectations. It focuses on senior leaders' actions to create and sustain a high-performance organization and an environment that is conducive for learning, student development, and achievement. This also examines how the organization fulfills its public responsibilities, especially how the senior leaders and faculty and staff encourage and practice good citizenship.

After compiling a list of common characteristics of winning companies, NIST concludes that the leaders from these successful companies are highly visible to all employees. They support and recognize the quality efforts made at every level (Leach, 1994).

2.2. Strategic planning

This category stresses that learning-centered education, long-term organizational sustainability, and competitive environment are the key strategic issues that need to be integral parts of the organization's overall planning. The category examines how the organization sets strategic directions and develops strategic objectives to guide and strengthen the performance of the entire organization. This category also examines how the organization converts the strategic objectives

into action plans and how the organization deploys the whole set of strategic objectives and action plans to all levels of the organization.

2.3. Student, stakeholder, and market focus

This category addresses how the organization determines the requirements, expectations, and preferences of the students and stakeholders with the focus on meeting their needs, delighting students and stakeholders, and building loyalty. This category stresses relationships as an important part of an overall listening, learning, and performance excellence strategy. Therefore, the category examines how the organization builds relationships with students and stakeholders and determines the key factors that attract students and lead to student and stakeholder satisfaction.

2.4. Measurement, analysis, and knowledge management

The aim of measurement and analysis is to guide the organization's process management toward the achievement of key organizational performance results and strategic objectives. The category examines how the organization selects, gathers, analyzes, manages, and improves its data, information, and knowledge assets. This category also addresses knowledge management and all basic performance-related information and comparative information, as well as how such information is analyzed and used to optimize organizational performance.

2.5. Faculty and staff focus

This addresses key human resource practices – those directed toward creating and maintaining a high-performance workplace with a strong focus on students' learning through faculty and staff's well-being and their satisfaction. In particular, the category examines the organization's compensation package, career progression, faculty and staff performance management, recognition, faculty and staff's continuing education and training. The category also examines the organization's working environment, faculty and staff support climate, and how the organization determines faculty and staff satisfaction, with the aim of fostering the well-being, satisfaction, and motivation of all faculty and staff while recognizing their diverse needs.

2.6. Process management

This category examines the organization's learning-centered processes for educational programs and offerings and students' services, with the aim of creating value for students and other key stakeholders. This category also examines the organization's support processes and operational planning with respect to financial management and planning for the continuity of operations, with the aim of improving overall operational performance.

2.7. Organizational performance results

This category examines the organization's actual performance in key areas: students' learning, student and stakeholder satisfaction, overall budgetary, financial, and market performance, faculty and staff performance, leadership, and social responsibility results. In other words, this category seeks real-time information on the organization's performance in the above areas. However, the MBNQA does not prescribe any specific quality tools, techniques, technology, or systems to achieve the results in the above areas. It is the organizations which need to identify/develop the required tools and techniques.

Despite the lack of any work on determining weights for the above categories of criteria, a substantial amount of the work has already been done on various other aspects of the MBNQA. We provide a brief account of some of the previous work in the next section.

3. Literature review

Ever since the launch, MBNQA has received considerable support from many sectors and the program started on a positive note. This has been encapsulated by Garvin's (1991, p. 80) comment:

> In just 4 years, Malcolm Baldrige National Quality Award has become the most important catalyst for transforming American business. More than any other initiative, public or private, it has reshaped managers' 'thinking and behavior'. The Baldrige Award not only codifies the principles of quality management in clear and accessible language, it also goes further: it provides companies with a comprehensive framework for assessing their progress toward the new paradigm of management and such commonly acknowledged goals as customer satisfaction and increased employee involvement.

Numerous organizations across the world have been applying the MBNQA criteria framework as a self-assessment tool (Steeples, 1993). According to the MBNQA official website (www.baldrige.nist.gov), approximately 2 million copies of the criteria booklet have been distributed all over the world since the publication of the first edition in 1988. Using the 1997 Baldrige criteria framework, Brown (1997) developed an instrument for use as a self-assessment tool. This instrument was designed to give the organization a rough idea of where the organization stands in the pursuit of excellence. BNQP (2005b, p. 17) states:

> Using the Baldrige Criteria for self-assessment and action planning makes you a winner even if you never apply for the Award. You win because you learn. You learn about what you do well and what you can do better. You win by using that learning to drive improvement and continuously achieve higher levels of performance.

Prybutok and Stafford (1997) have described a case study where MBNQA criteria were used for self-assessment at Baylor Health Care System (BHCS). The survey findings have helped BHCS to develop a few action plans to achieve its long-term objectives. One such action was to develop a $50 million information system to manage the organization.

For many companies, the major objective behind applying for MBNQA is to obtain feedback from the examiners. The feedback, which is frequently 30 or more pages long, is very helpful to the companies for further improvement. Kay Kendall, vice president of Corporate Quality at Brooks-PRI Automation Inc. (Chelmsford, MA) and judge of MBNQA, says (Babicz, 2002, p. 38):

> I really do encourage companies that are looking for that independent, holistic evaluation to consider applying. I just don't think that there's any other way of getting that kind of independent, in-depth assessment like the process. The caliber of the folks that you get at the national level as examiners and the business experience that they have is so strong that it's amazing how insightful that feedback is.

Mr Peterson, the former CEO of Ford Motor Company, a Baldrige Award applicant in 1989, says that, 'Participating in the Baldrige process was good for Ford, because the evaluation process offered opportunities for improvement, despite the fact that Ford did not win the prize' (Ghobadian and Woo, 1996, p. 43)

About 40 national-level quality programs worldwide have root in MBNQA. Perhaps it will be no exaggeration to say that the above fact alone suggests that the MBNQA criteria framework comprehensively captures the major dimensions of total quality management (TQM). In fact, Curkovic et al. (2000) have empirically investigated the correspondence between MBNQA criteria and TQM dimensions and the authors conclude that the MBNQA criteria framework does capture the core concepts of TQM.

In order to investigate whether quality management systems are related to organization results and customer satisfaction in hospitals, Goldstein and Schweikhart (2002) have examined the relationship among constructs in the Baldrige award for a health care organization. They conclude that by focusing on the content of the award criteria, hospitals can improve their overall performance. In a related study, Handfield and Ghosh (1995) find that leadership influences each of the first six criteria while process management and strategic planning significantly influence customer satisfaction.

Economic growth of a country or a state is a function of physical capital, human capital, natural resources, and state of technology. Economic development also depends on educational infrastructure, competitive business environment, and the degree of governmental support. It is expected that state-supported programs including quality awards will contribute positively to the state's economy. In this context, Fisher et al. (2001) empirically investigated the relationship between quality awards and economic development and they found a positive correlation between the two.

After the launch of MBNQA in the United States and having seen its impressive success, many other nations have started launching national quality awards. Puay et al. (1998) have compared nine such awards (three European, two North American, three Asia-Pacific, and one South American) on nine major criteria: leadership, impact on society, resource management, strategy and policy, human resource management, process quality, results, customer management and satisfaction, supplier/partner management and performance. Along the same lines, Ghobadian and Woo (1996) have discussed the characteristics, benefits, and shortcomings of four major quality awards: Deming Prize, MBNQA, European Quality Award, and Australian Quality Award.

Reimann and Herz (1996) have compared the MBNQA and the ISO 9000 quality management system. The authors are of the opinion that they differ fundamentally in focus, purpose, and

content. According to them, the focus of MBNQA is to enhance competitiveness, whereas the focus of ISO 9000 registration is conformity to practices specified in the registrant's own quality system. Most of the ISO 9000 requirements fall under the process management category of MBNQA. Overall, ISO 9000 registration requirements cover less than 10% of the Baldrige Award criteria (BNQP, 2005a).

Wilson and Collier (2000) used a survey instrument consisting of 101 questions to test the theory and causal performance linkages implied by the MBNQA. They concluded the following: (1) MBNQA theory supports that leadership drives the systems that produce results; (2) leadership is the most important driver of system performance; (3) leadership has no direct effect on financial results but influences overall performance; (4) information and analysis is statistically the second-most important category in MBNQA criteria; (5) process management is twice as important when predicting customer satisfaction than when predicting financial results. Prybutok et al. (2001) have investigated the relationships between leadership and other MBNQA criteria paying special attention to the information and analysis component.

Seymour (1996), in his two-volume book, detailed why and how US schools can use the MBNQA criteria framework to improve the overall quality of education. Johnson (1996) investigated the answer to the following questions on the application of Baldrige and State quality awards in the education sector: what is the purpose of those awards? What criteria are used to judge the applicants? What has been the impact to date on the educational community?

Despite all the above positive notes, criticisms of the MBNQA are also not rare. Collier (1992) has responded to several criticisms: MBNQA emphasizes process more than results or achievements, high cost of application, 'product mentality examiners don't know how to evaluate service firms', etc. The author concludes by saying (p. 94):

> The Malcolm Baldrige National Quality Award has raised the consciousness of the United States about quality performance. It helps explain the vital role quality performance plays in creating the world's standard of living and quality of life. Despite its critics, the MBNQA was established to achieve long-term national goals, and it is indeed achieving those goals.

Regarding the assignment of points to the criteria, Prybutok and Stafford (1997, p. 45) mention:

> A point allocation scheme emphasizes various categories and sub categories according to their relative contribution to the overall quality level of an organization's current quality practices and award points that reflect the organization's quality strengths and weaknesses in each of the seven categories.

Apart from the above reference, in the wide range of literature on MBNQA, we did not find any other work (even passing mention) regarding how the MBNQA criteria and subcriteria points are allocated. It seems that nobody has questioned the legitimacy of the allocated points. Details are also not available regarding the point allocation of other national quality awards. To the best of the author's knowledge, this is the first attempt to determine MBNQA criteria weights using some analytical technique on the basis of public opinion.

4. Research methodology

MBNQA is an American award for American organizations. The criteria framework of the award is quite comprehensive and it comprises most of the basic tenets of TQM. For this reason, MBNQA has been a 'role model' in developing national quality awards in many other countries. One such country is Malaysia, which launched the 'Prime Minister's Quality Award' (PMQA) in 1990. It is regarded as the most prestigious award in business in Malaysia. It is interesting to note that the criteria framework for PMQA is very similar to that for MBNQA. Although PMQA was launched in 1990, Malaysia is yet to start the national quality award in the education sector. In Malaysia, it is widely known that the Federal Government intends to make the country a regional center of educational excellence. In line with this objective, it will be highly fitting that the government takes the initiative to launch a national quality award for the education sector. If the government makes the decision to launch the award and adopts the MBNQA criteria framework, the next task for the government (or appropriate body from within the government) is to modify the weights assigned to the criteria set. In the multi-criteria decision making (MCDM) literature, it is well known that assigning weights to the criteria set is a 'local phenomenon'. This means that the weights will be different for different decision makers. This is logical and a matter of common sense. In view of this, we can confidently state that the MBNQA criteria weights will not be the same in the Malaysian setting. The present work has been carried out to obtain weights for the MBNQA criteria in education in the Malaysian context.

To determine the weights, we used the analytic hierarchy process (AHP) (Saaty, 1980). The AHP determines weights of a set of factors by comparing them pairwise and it uses its own (1/9, 9) ratio scale judgments. The description of the scale is provided in Saaty (1980, p. 54). The MBNQA criteria framework has seven categories and each category has a number of subcategories. Altogether eight pairwise comparison matrices (PCMs) were formed – one for the categories and the remaining seven for the subcategories under each of the seven categories. Instead of using any traditional questionnaire, we formed eight empty PCMs that comprise only the headings in the first rows and first columns, and the responses were collected on a personal contact basis.

We managed to contact 39 academicians from three Malaysian universities: International Islamic University Malaysia (IIUM), Multimedia University, and Universiti Malaya. For all the respondents, prior appointments were arranged before going to meet them. The breakdown of the respondents are as follows: professors – 5, associate professors – 15, assistant professors – 10, and lecturers – 9.

The majority of the respondents came from IIUM, the author's university. IIUM respondents comprise (in addition to others) the following persons:

- Director of the Quality Assurance Unit
- Deputy Director of the Quality Assurance Unit
- Dean of the Faculty of Economics and Management Sciences (KENMS)
- Deputy Dean (Academic affairs) of KENMS
- Deputy Dean (PG) of KENMS
- Director of the Management Center
- Deputy Director of the Management Center
- Head of the MBA and the Master of Management programs at the Management Center
- Heads of the three departments: Economics, Accounting, and Business Administration

5. Data collection and analysis

Data collection from every respondent started with an explanation of the (1/9, 9) ratio-scale, which was to be used in completing the PCMs. The explanation was repeated for all the respondents. Exhibit 1 shows a set of PCMs obtained from one of the 39 respondents. Acronyms of the criteria and subcriteria are used in all the PCMs in Exhibit 1 (full forms of the criteria and subcriteria are available in Table 1).

The weights of the criteria and sub-criteria from each individual's pairwise comparison matrices were not determined, rather the responses from all the respondents were aggregated using the geometric mean rule (Basak and Saaty, 1993). To compute geometric mean, the values are multiplied first and then a root equal to the number of individuals who provided the values is taken. For example, the geometric mean of all the 39 values for the comparison Leadership and Strategic Planning is

$$\sqrt[39]{6 \times 3 \times \ldots \times 1/2} = 1.92.$$

Table 1
Malcolm Baldrige National Quality Award (MBNQA) criteria, their existing and proposed points (Malaysian context)

Criteria	MBNQA points	Proposed points
Leadership (L)	120	177
Senior leadership (SL)	70	104
Governance and Social Responsibility (GSR)	50	73
Strategic planning (SP)	85	134
Strategy development (SDev)	40	45
Strategy deployment (SDep)	45	89
Students, stakeholders, and market focus (SSMF)	85	105
Student, stakeholder, and market knowledge (SSMK)	40	39
Student and stakeholder relationship and satisfaction (SSRS)	45	66
Measurement, analysis, and knowledge management (MAKM)	90	72
Measurement, analysis, and review of organizational performance (MAROP)	45	30
Information and knowledge management (IKM)	45	42
Faculty and staff focus (FSF)	85	204
Work system (WS)	35	49
Faculty and staff learning and motivation (FSLM)	25	74
Faculty and staff well-being and satisfaction (FSWS)	25	81
Process management (PM)	85	87
Learning-centered processes (LCP)	45	51
Support processes and operational planning (SPOP)	40	36
Organizational performance results (OPR)	450	220
Students learning results (SLR)	100	50
Students and stakeholders focused results (SSFS)	70	39
Budgetary, financial, and market results (BFMR)	70	19
Faculty and staff results (FSR)	70	48
Organizational effectiveness results (OER)	70	35
Governance and social responsibility results (GSRR)	70	30

	L	SP	SSMF	MAKM	FSF	PM	OPR
L	1	6	1	7	1	7	1/3
SP		1	1/5	3	1/5	1	1/6
SSMF			1	5	1	4	1
MAKM				1	1/7	1	1/8
FSF					1	6	1/4
PM						1	1/7
OPR							1

L	SL	GSR
SL	1	4
GSR		1

SP	SDev	SDep
SDev	1	1/5
SDep		1

SSMF	SSMK	SSRS
SSMK	1	1/3
SSRS		1

MAKM	MAROP	IKM
MAROP	1	1/3
IKM		1

FSF	WS	FSLM	FSWS
WS	1	1/5	1
FSLM		1	1
FSWS			1

PM	LCP	SPOP
LCP	1	1/3
SPOP		1

OPR	SLR	SSFR	BFMR	FSR	OER	GSRR
SLR	1	3	6	7	5	1
SSFR		1	5	1	1	1/4
BFMR			1	1/7	1/7	1/8
FSR				1	1/4	1/5
OER					1	1/3
GSRR						1

Exhibit 1. A set of PCMs obtained from one of the respondents.

Similarly, the geometric means of all other comparisons shown in Exhibit 1 were calculated and Team Expert Choice (Saaty and Forman, 2000) was used for this purpose. The aggregated PCMs and the computed weights for each matrix are shown in Exhibit 2.

The weights shown in the last columns of the PCMs in Exhibit 2 are all fractional. But these can easily be converted to integral weights, which are shown in Table 1.

For comparison purposes, MBNQA weights are also shown in Table 1. From the proposed set of weights, the observations in the following section are clear.

6. Some observations in the new set of weights

6.1. Leadership

The category has received significantly higher weight compared to MBNQA. This is apparently because the respondents think that if the leadership is 'ok', then everything else will be 'ok' in the organization. One respondent openly said, 'I'll place leadership above all'. Furthermore, senior leadership has received more points compared to social responsibility as it does in the MBNQA framework.

	L	SP	SSMF	MAKM	FSF	PM	OPR	Wts
L	1	1.92	1.63	2.33	0.99	1.87	0.52	0.177
SP		1	1.46	1.96	0.77	1.43	0.68	0.134
SSMF			1	1.32	0.53	1.35	0.53	0.105
MAKM				1	0.32	0.85	0.36	0.072
FSF					1	2.98	0.88	0.204
PM						1	0.49	0.087
OPR							1	0.220

L	SL	GSR	Wts
SL	1	1.41	0.585
GSR		1	0.415

SP	SDev	SDep	wts
SDev	1	0.50	0.333
SDep		1	0.667

SSMF	SSMK	SSRS	Wts
SSMK	1	0.59	0.370
SSRS		1	0.630

MAKM	MAROP	IKM	Wts
MAROP	1	0.70	0.417
IKM		1	0.503

FSF	WS	FSLM	FSWS	Wts
WS	1	1/5	1	0.242
FSLM		1	1	0.362
FSWS			1	0.395

PM	LCP	SPOP	Wts
LCP	1	1.40	0.583
SPOP		1	0.417

OPR	SLR	SSFR	BFMR	FSR	OER	GSRR	Wts
SLR	1	1.13	2.65	1.34	1.10	1.82	0.226
SSFR		1	2.17	0.77	1.02	1.25	0.175
BFMR			1	0.37	0.56	0.71	0.086
FSR				1	1.70	1.57	0.220
OER					1	1.05	0.159
GSRR						1	0.134

Exhibit 2. The set of synthesized PCMs.

6.2. *Strategic planning*

This has also received higher points compared to MBNQA. However, unlike MBNQA, strategy deployment has received significantly higher points compared to strategy development. Out of 39 respondents, only seven said that strategy development was more important than strategy deployment, nine said they were equally important, and the rest, i.e., 23 respondents said that strategy deployment was more important than strategy development. One respondent commented, 'Many times, we develop many things but we don't really deploy and we don't make them work.' In fact, in a survey of a broad cross section of CEOs, the Malcolm Baldrige Foundation learned that CEOs believed deploying strategy is three times more difficult than developing strategy (BNQP, 2005a).

6.3. *Students, stakeholders, and market focus*

This also has received higher weight compared to MBNQA. Furthermore, 'Student and Stakeholder Relationship and Satisfaction' has received higher points than in MBNQA.

6.4. Measurement, analysis, and knowledge management

This has received lesser weightage here compared to MBNQA. This is because the respondents have assigned higher weights to all the previous three categories. Hence, as the total points are constant, some category is expected to receive lesser weight. However, it is to be noted that the difference between the two sets of points (proposed and MBNQA) is not significantly large.

6.5. Faculty and staff focus

This has received the highest amount of increase in terms of points. An organization's success depends upon the diverse backgrounds, knowledge, skills, creativity, and motivation of all its faculty and staff. Overall, the respondents strongly felt that if the staff's well-being and satisfaction were taken care of then it would be easier to achieve organizational performance results, especially student learning results.

6.6. Process management

This category along with its subcategories, namely, 'learning centered processes' and 'support processes and operational planning', have received almost the same weight as in MBNQA.

6.7. Organizational performance results

For this category, MBNQA points are 450 (almost half of the total points), whereas the proposed points total 220. This shows that the respondents have placed lesser weight on this category compared with MBNQA. However, this does not mean that all the respondents have done so. Table 2 presents the weightings of each category judged to be the most important by the respondents. Note that not only is 'organizational performance results' judged to be the most important category by a majority of the respondents, but also its range of weights have higher values for both upper and lower limits than the other three categories shown in Table 2.

Table 2
Respondents' most important category

Category	No. of respondents that have placed highest weighting	Range of weights (for Col. 2 respondents only)
Leadership	10	(0.262, 0.427)
Strategic planning	7	(0.245, 0.371)
Faculty and staff focus	6	(0.260, 0.379)
Organizational performance results	14	(0.318, 0.526)

	L	SP	SSMF	MAKM	FSF	PM	OPR
L	1	(1/9, 9)	(1/7, 9)	(1/8, 8)	(1/7, 9)	(1/7, 9)	(1/9, 6)
SP		1	(1/9, 9)	(1/8, 8)	(1/9, 8)	(1/6, 8)	(1/8, 8)
SSMF			1	(1/6, 9)	(1/8, 8)	(1/8, 8)	(1/9, 8)
MAKM				1	(1/9, 7)	(1/7, 7)	(1/9, 7)
FSF					1	(1/7, 9)	(1/9, 7)
PM						1	(1/9, 7)
OPR							1

L	SL	GSR
SL	1	(1/7, 7)
GSR		1

SP	SDev	SDep
SDev	1	(1/9, 7)
SDep		1

SSMF	SSMK	SSRS
SSMK	1	(1/8, 7)
SSRS		1

MAKM	MAROP	IKM
MAROP	1	(1/8, 7)
IKM		1

FSF	WS	FSLM	FSWS
WS	1	(1/8,7)	(1/8, 8)
FSLM		1	(1/8, 8)
FSWS			1

PM	LCP	SPOP
LCP	1	(1/7, 8)
SPOP		1

	SLR	SSFR	BFMR	FSR	OER	GSRR
SLR	1	(1/8, 7)	(1/6, 8)	(1/6, 7)	(1/6, 7)	(1/9, 9)
SSFR		1	(1/5, 8)	(1/8, 7)	(1/7, 7)	(1/9, 8)
BFMR			1	(1/7, 5)	(1/7, 6)	(1/9, 7)
FSR				1	(1/7, 8)	(1/9, 8)
OER					1	(1/9, 8)
GSRR						1

Exhibit 3. Interval PCMs considering the lowest and highest numbers for all the queries.

Overall, the respondents vary considerably in assigning weights to the MBNQA category and sub-categories. This is clear from the interval pairwise comparison matrices shown in Exhibit 3. The intervals for all the entries in all the matrices are formed by taking the lowest and highest numbers entered by all the 39 respondents.

7. Observations in AHP data collection

A number of issues were observed while collecting AHP data for the present work. These are discussed below.

7.1 Using narrow range of numbers

Some respondents used a very narrow range of numbers in articulating the responses by repeating one particular number from the $(1/9, 9)$ ratio scale many times. Observe the following matrix, where the respondent (one assistant professor) used only three numbers, namely 1/2, 2 and 3, in the range $(1/2, 3)$ out of the total 17 numbers in the range $(1/9, 9)$.

OPR	SLR	SSFR	BFMR	FSR	OER	GSRR
SLR	1	2	1/2	2	1/2	1/3
SSFR		1	3	2	1/2	2
BFMR			1	3	3	3
FSR				1	3	3
OER					1	3
GSRR						1

7.2 Using too many 1s

Note the following PCM filled up by one assistant professor:

OPR	SLR	SSFR	BFMR	FSR	OER	GSRR
SLR	1	5	1	1	1	1
SSFR		1	1/5	1	1	1
BFMR			1	1	1	5
FSR				1	1	5
OER					1	7
GSRR						1

It is to be noted that like previous observations, the respondent has used only three numbers, namely 1, 5, and 7, repeating '1' 10 times. In fact, the same respondent has used '1' 27 times out of the total 44 in the course of filling up all the eight pairwise comparison metrics. In particular, he has used 1s for all the entries in the matrices constructed for 'leadership', 'strategic planning', 'measurement, analysis and knowledge management', 'faculty and staff focus', and 'process management'. The matter should not be viewed as a one-off occurrence; the same thing was observed for a number of respondents.

7.3 Repeating one particular number many times

Note the following PCM filled up by a lecturer:

	L	SP	SSMF	MAKM	FSF	PM	OPR
L	1	8	1/8	7	8	8	1/8
SP		1	8	8	8	8	8
SSMF			1	8	8	8	8
MAKM				1	1/7	1/7	7
FSF					1	1/7	1/7
PM						1	7
OPR							1

Here the observation is similar to the first observation, i.e., using a narrow range of numbers. The respondent has used '8' 12 times out of the 21 comparisons.

7.4 Using 'high-level' numbers

Note the matrix in the foregoing observation and also the following matrix filled in by the same lecturer:

OPR	SLR	SSFR	BFMR	FSR	OER	GSRR
SLR	1	1/8	6	1/6	6	8
SSFR		1	8	7	7	6
BFMR			1	1/6	1/7	1/7
FSR				1	6	6
OER					1	1/7
GSRR						1

In both the matrices, the respondent has used only the numbers 6, 7, and 8 (including their reciprocals). For the remaining six matrices also, he has used only these three numbers. That is, he has not used the lower level numbers 1, 2, 3, 4, and 5 at all.

7.5 Using the extreme numbers

Note the following PCM filled up by one assistant professor:

	L	SP	SSMF	MAKM	FSF	PM	OPR
L	1	1/9	1/7	1/7	8	5	1/6
SP		1	9	7	8	8	8
SSMF			1	9	1/7	4	6
MAKM				1	1/7	7	7
FSF					1	7	1/7
PM						1	1/7
OPR							1

It is to be noted that basically the respondent is inclined to use numbers from the extremes, i.e., 1/9 and 9. The respondent is not considering the moderate numbers except for one comparison (SSMF, PM).

7.6 Revising elements in a PCM

The following PCM from one lecturer may be noted:

	L	SP	SSMF	MAKM	FSF	PM	OPR
L	1	1/2	1/6	2	1/2	1/5	1/8
SP		1	1/5	1/2	1/3	1/3	1/8
SSMF			1	2	2	1/2	1/5
MAKM				1	1/2	1/5	1/7
FSF					1	1/2	1/5
PM						1	1/7
OPR							1

Initially, the respondent stated '8' as the importance level for 'leadership' (L) over 'student, stakeholder and market focus' (SSMF). This means that 'L' is 'very strongly more' important compared to 'SSMF'. On the other hand, from the comparison (L, SP), we see that 'SP' is more important than 'L'. From the above two comparisons, we can conclude that SP is 'extremely more' important compared to 'SSMF'. However, the respondent specified '1/5' for the comparison, signifying gross inconsistency in the judgments. Although inconsistency is allowed in AHP, gross inconsistency like this is not at all desired. The matter was pointed out to the respondent and then she changed the comparison for (L, SSMF) from 8 to 1/6; apparently, the respondent made a 'mistake' in her previous judgment. This revision was possible because data were collected personally. If the AHP data are collected through normal questionnaire, then perhaps the inconsistency will remain. But it can be minimized by requesting the respondents to take extra care about the consistency matter. The author has also seen that some respondents revise their previous judgments themselves (without pointing out the matter). It happened to at least one professor who initially rated leadership as 5 over strategic planning. But later when she started rating strategic planning higher over some others, she herself revised the comparison (L, SP) from 5 to just 1.

7.7 Using 1s

It has been mentioned above that the data were collected personally. Initially, we explained the (1/9, 9) scale and how to use it. When we saw that one of the respondents was not using 1s (apparently, because the respondent thought that '1' could not be used), we reminded them that the factors could also be equally important and in the AHP scale, there was a scope to enter 1. After this reminder, we observed that the same respondent was using 1s for many later comparisons. However, this should not be viewed as an instance of influencing the respondent, as all respondents were totally free to articulate their responses.

7.8 Time taken to fill in all the eight PCMs

As expected, we noted that the respondents were taking different amounts of time to fill in the PCMs. Some respondents were very prompt in articulating their responses; they responded almost immediately after asking the AHP pairwise comparison question. On the other hand, others were more thoughtful in their judgments. Some respondents completed all the matrices in less than 15 min, while some others took more than 30 min. It was also observed that the respondents who take more time to fill in the matrices were more consistent in their judgments compared to the respondents who took less time.

7.9 Articulating the judgments using verbal phrases

Some respondents were found to be more comfortable in using verbal phrases like 'moderately more' or 'strongly more' compared to using numerical ratings like 3 or 5.

8. Proposing a new evaluation procedure for MBNQA

MBNQA applications are lengthy; the length of a typical application ranges from 50 to 75 pages. In 2004 alone, NIST received 60 applications of which 17 were in the education sector. In the first phase of the evaluation process, all applications undergo rigorous examination that involves about 300 h of review by a panel of experts using the scoring guidelines presented in Table 3. Categories 1 through 6 of the criteria set are evaluated using the process approach. Process refers to the methods the organization uses and improves to address the item requirements. Furthermore, four factors, namely Approach, Deployment, Learning, and Integration (A–D–L–I), are used to evaluate a process. The items in category 7 are evaluated based on results. A score taken from one of the several ranges from Table 3 is assigned to each item. Note that the examiner can only assign a score that is a multiple of 5 and the score is assigned using a holistic approach on the basis of the above four factors. In short, in assigning a score, firstly the examiner has to select an appropriate range and then find out the actual score (which is a multiple of 5, as mentioned above) belonging to the range. The Baldrige award scoring guideline (BNQP, 2005a, p. 54) says:

> Assigning the actual score within the chosen range requires evaluating whether the item response is clear to the statements in the next higher or next lower scoring range.

The following two points can be noted from the scoring guidelines:

- Firstly, the examiner has to choose the grade to be assigned to the applicant for every item. Next, he/she has to identify the particular score in the corresponding range.
- In any range of scores, only the score that is a multiple of 5 can be considered. For instance, in the range 10–25, only 10, 15, 20, or 25 can be assigned, leaving no room

Table 3
Malcolm Baldrige National Quality Award (MBNQA) applications scoring guidelines

Score	Interpretation
0% or 5%	Very poor
10%, 15%, 20%, or 25%	Poor
30%, 35%, 40%, or 45%	Average
50%, 55%, 60%, or 65%	Good
70%, 75%, 80%, or 85%	Very good
90%, 95%, or 100%	Excellent

for consideration of any other score in the range. It is unclear why an examiner cannot choose a score like 13 or 23.

The scoring scheme proposed in the present article does not require assigning a numerical score at all, what it needs is to specify the performance of the organization on a certain item using a letter grade (E-excellent, G-good, A-average, S-satisfactory, P-poor). The numerical equivalents of the letter grades are predetermined using the absolute measurement procedure of AHP (Saaty, 1990). We describe the steps of the absolute measurement of AHP in the following:

Step 1: Identify the criteria, sub-criteria, and institutes[1] (to be evaluated) for evaluation and place them into the AHP hierarchy.

Step 2: Calculate the weights of the decision criteria by the relative measurement of AHP, i.e., construct the pairwise comparison matrix for all the criteria and compute the normalized principal right eigenvector of the matrix. This vector gives the weights of the criteria. Divide the criteria into sub-criteria (where applicable) and calculate the weights of these sub-criteria in the same manner. Multiply these weights by the weights of the parent criteria.

Step 3: Divide each sub-criterion into several intensities or grades. Set priorities on the intensities by comparing them pairwise under each sub-criterion. Multiply these priorities by the priority of the parent sub-criterion.

If p_i, $i = 1, 2, \ldots, m$, is the weight of the ith main criterion, q_{ij}, $i = 1, 2, \ldots, m$, $j = 1, 2, \ldots, n$, is the weight of the jth sub-criterion of the ith criterion, then the global weight r_{kg} of the kth intensity, $k = 1, 2, \ldots, s$, with respect to the jth sub-criterion of the ith criterion is

$$r_{kg} = p_i \times q_{ij} \times r_k \tag{1}$$

where r_k is the local weight of the kth intensity.

Step 4: Take one institute at a time and measure its performance under each sub-criterion. Add the global priorities of the intensities for the institute. Repeat the process for all the institutes.

[1]Elementary and secondary schools, technical schools, professional schools, colleges, universities, etc.

The AHP hierarchy of the proposed evaluation exercise is shown in Fig. 2a and b. The PCM for the intensities (E-Excellent, G-Good, A-Average, S-Satisfactory, P-Poor) is shown in the following:

	E	G	A	S	P	Weights
E	1	3	5	6	7	0.494
G		1	3	5	7	0.268
A			1	3	5	0.133
S				1	3	0.067
P					1	0.037

a

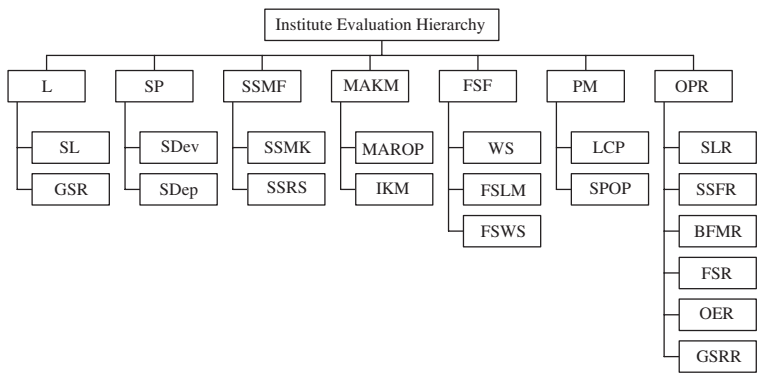

Fig. 2a. The hierarchy of the criteria of the evaluation process.

b

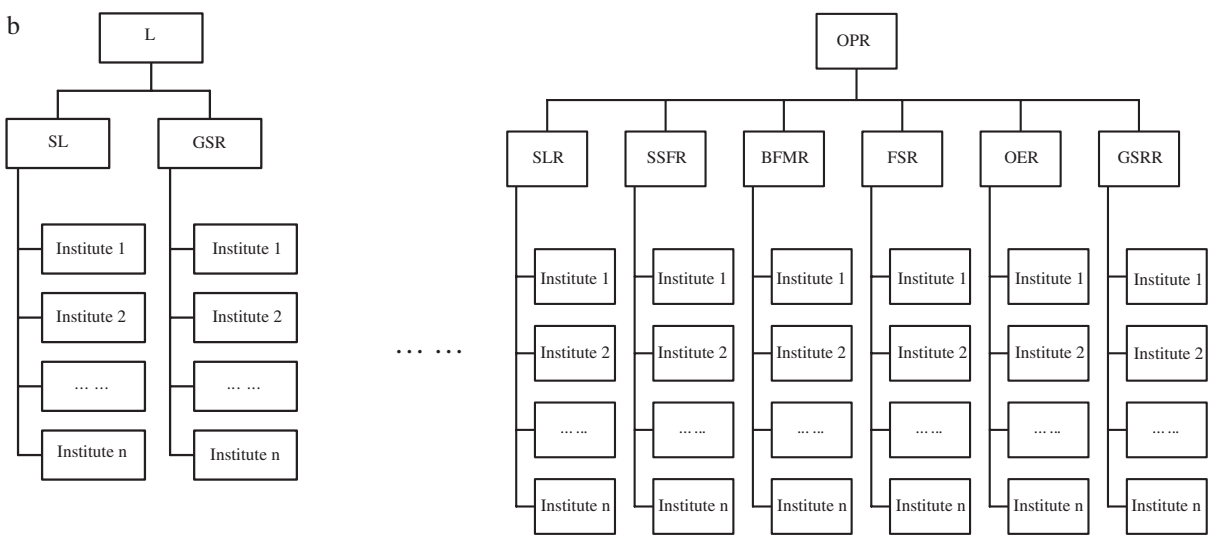

Fig. 2b. The partial hierarchy consisting of the institutes.

Table 4
Global weights of the intensities

	Excellent	Good	Average	Satisfactory	Poor
SL	0.051	0.028	0.014	0.007	0.004
GSR	0.036	0.020	0.010	0.005	0.003
SDev	0.019	0.013	0.006	0.004	0.002
SDep	0.038	0.027	0.012	0.007	0.005
SSMK	0.018	0.012	0.005	0.003	0.002
SSRS	0.030	0.020	0.008	0.005	0.003
MAROP	0.013	0.008	0.005	0.003	0.002
IKM	0.018	0.011	0.007	0.004	0.003
WS	0.023	0.015	0.006	0.004	0.002
FSLM	0.034	0.022	0.009	0.006	0.003
FSWS	0.037	0.024	0.010	0.006	0.003
LCP	0.025	0.014	0.007	0.003	0.002
SPOP	0.018	0.010	0.005	0.002	0.002
SLR	0.024	0.014	0.007	0.004	0.002
SSFR	0.018	0.011	0.005	0.003	0.001
BFMR	0.009	0.005	0.003	0.001	0.001
FSR	0.023	0.013	0.007	0.003	0.002
OER	0.017	0.010	0.005	0.002	0.001
GSRR	0.014	0.008	0.004	0.002	0.001

SL, senior leadership; GSR, governance and social responsibility; SDev, strategy development; SDep, strategy deployment; SSMK, student, stakeholder, and market knowledge; SSRS, student and stakeholder relationship and satisfaction; MAROP, measurement, analysis, and review of organizational performance; IKM, information and knowledge management; WS, work system; FSLM, faculty and staff learning and motivation; FSWS, faculty and staff well-being and satisfaction; LCP, learning-centered processes; SPOP, support processes and operational planning; SLR, students learning results; SSFR, students and stakeholders focused results; BFMR, budgetary, financial and market results; FSR, faculty and staff results; OER, organizational effectiveness results; GSRR, governance and social responsibility results.

Using formula (1), the global weights of the intensities for all the sub-criteria are determined and these are shown in Table 4.

Now, let us consider five institutions from Malaysia and evaluate them on every sub-criterion. Table 5 shows hypothetical data on performance of these institutions.

The fractional numbers in Columns 2 through 20 of Table 5 are numerical equivalents of the corresponding letter grades. The last column shows the overall weights of the institutions. The ranking of the institutions can easily be derived from the overall weights. After necessary adjustments in the second phase of the evaluation process, higher ranked institutions can be selected for site visits (third phase). In the final phase, judges select the winner on the basis of the recommendations of the senior examiners who visited the sites.

9. Conclusions

Normally, evaluation of a set of alternatives is carried out on the basis of a number of criteria. For evaluation, in addition to the identification of criteria, what is even more important is to assign an appropriate set of weights to the criteria. Obviously, a variation in the criteria weights impacts the ranking of the alternatives. In this article, the issue of assigning weights to the MBNQA criteria

Table 5
Five institutions and their performances on various subcriteria

	SL	GSR	SDev	SDep	SSMK	SSRS	MAROP	IKM	WS	FSLM
Inst 1	G	S	G	S	S	S	G	S	S	G
	0.028	0.010	0.013	0.012	0.005	0.008	0.008	0.007	0.006	0.022
Inst 2	E	G	G	G	S	G	S	S	G	G
	0.051	0.020	0.013	0.027	0.005	0.020	0.005	0.007	0.015	0.022
Inst 3	G	S	G	S	S	G	S	S	G	G
	0.028	0.010	0.013	0.012	0.005	0.020	0.005	0.007	0.015	0.022
Inst 4	S	G	G	G	G	S	E	E	E	G
	0.014	0.020	0.013	0.027	0.012	0.008	0.013	0.018	0.023	0.022
Inst 5	G	E	E	G	G	P	G	S	G	G
	0.028	0.036	0.019	0.027	0.012	0.005	0.008	0.007	0.015	0.022

	FSWS	LCP	SPOP	SLR	SSFR	BFMR	FSR	OER	GSRR	Wts.
Inst 1	G	G	E	G	G	G	S	S	G	0.225
	0.024	0.014	0.018	0.014	0.011	0.005	0.007	0.005	0.008	
Inst 2	S	G	S	E	S	S	G	E	G	0.284
	0.010	0.014	0.005	0.024	0.005	0.003	0.013	0.017	0.008	
Inst 3	G	S	G	G	E	E	G	S	S	0.241
	0.024	0.007	0.010	0.014	0.018	0.009	0.013	0.005	0.004	
Inst 4	G	E	S	G	G	G	E	G	P	0.289
	0.024	0.025	0.005	0.014	0.011	0.005	0.023	0.010	0.002	
Inst 5	S	E	G	E	E	S	G	G	S	0.296
	0.010	0.025	0.010	0.024	0.018	0.003	0.013	0.010	0.004	

SL, senior leadership; GSR, governance and social responsibility; SDev, strategy development; SDep, strategy deployment; SSMK, student, stakeholder, and market knowledge; SSRS, student and stakeholder relationship and satisfaction; MAROP, measurement, analysis, and review of organizational performance; IKM, information and knowledge management; WS, work system; FSLM, faculty and staff learning and motivation; FSWS, faculty and staff well-being and satisfaction; LCP, learning-centered processes; SPOP, support processes and operational planning; SLR, students learning results; SSFR, students and stakeholders focused results; BFMR, budgetary, financial, and market results; FSR, faculty and staff results; OER, organizational effectiveness results; GSRR, governance and social responsibility results.

has been addressed. The MBNQA framework that encapsulates the essentials of TQM has been used as a 'role model' in establishing many national quality awards throughout the world. In general, for any multi-criteria problem, the set of criteria weights varies from decision maker to decision maker. Depending upon the objective of the decision maker, the variation in the weights could be small or large. In this article, we have empirically determined the weights of the MBNQA criteria set in Education from a Malaysian perspective using an analytic technique. It is expected that the findings will be used as guidelines to develop the Malaysian quality award in Education. Furthermore, the procedure can be extended to assign weights to the criteria for any other nation that intends to launch a quality award in Education. However, the present study is limited to a small sample size.

The paper also proposes a novel evaluation scheme for MBNQA applications. It is expected that the new approach, if applied, will enhance further the quality of the evaluation exercise. The working of the proposed procedure has been shown using a hypothetical example. But the application can easily be extended to a real setting.

Acknowledgement

The author is grateful to the anonymous referees for their insightful comments. The author also wishes to thank all the respondents for their feedback in the data collection exercise.

References

Babicz, G., 2002. Assessing the Baldrige award. *Quality*, November, 36–39.

Baldrige National Quality Program 2005a. *Education Criteria for Performance Excellence*. National Institute of Standards and Technology, US Department of Commerce. Available at www.baldrige.nist.gov.

Baldrige National Quality Program 2005b. *Getting started*. National Institute of Standards and Technology, US Department of Commerce. Available at www.baldrige.nist.gov.

Basak, I., Saaty, T.L., 1993. Group decision making using the analytic hierarchy process. *Mathematical and Computer Modelling* 17, 101–109.

Brown, M.G., 1997. Measuring up against the 1997 Baldrige criteria. *Journal of Quality and Participation* 20, 4, 22–28.

Collier, D.A., 1992. Service, please: the Malcolm Baldrige National Quality Award. *Business Horizons*, July–August, 88–95.

Curkovic, S., Melnyk, S., Calantone, R., Handfield, R., 2000. Validating the Malcolm Baldrige National Quality Award framework through structural equation modeling. *International Journal of Production Research* 38, 4, 765–791.

Fisher, C., Dauterive, J., Barfield, J., 2001. Economic impacts of quality awards: Does offering an award bring returns to the state? *Total Quality Management* 12, 7&8, 981–987.

Garvin, D.A., 1991. How the Baldrige award really works. *Harvard Business Review*, November–December, 80–93.

Ghobadian, A., Woo, H.S., 1996. Characteristics, benefits and shortcomings of four major quality awards. *International Journal of Quality and Reliability Management* 13, 2, 10–44.

Goldstein, S.M., Schweikhart, S.B., 2002. Empirical support for the Baldrige award framework in US hospitals. *Health Care Management Review* 27, 1, 62–75.

Handfield, R.B., Ghosh, S., 1995. An empirical test of linkages between the Baldrige criteria and financial performance. *Proceedings of the Decision Sciences Institute Annual Meetings,* 1713–1715.

Johnson, H.H., 1996. The Baldrige and state quality awards for education. *The Journal for Quality and Participation* 19, 1, 88–92.

Leach, K.E., 1994. Lessons from the Baldrige. *Industry Week*, September 19, 42.

Prybutok, V.R., Stafford, M.R., 1997. Using Baldrige criteria for self-assessment. *Marketing Health Services*, 45–48.

Prybutok, V.R., Richards, R.M., Cutshall, R., 2001. The significance of information and analysis as a component of a leadership model based on Malcolm Baldrige National Quality Award criteria. *Journal of Computer Information Systems* 41, 4, 52–56.

Puay, S.H., Tan, K.C., Xic, M., Goh, T.N., 1998. A comparative study of nine national quality awards. *The TQM Magazine* 10, 1, 30–39.

Reimann, C.W., Herz, H.S., 1996. The Baldrige award and ISO 9000 registration compared. *Journal of Quality and Participation* 19, 1, 12–19.

Saaty, T.L., 1980. *The Analytic Hierarchy Process*. McGraw-Hill, New York.

Saaty, T.L., 1990. How to make a decision: The analytic hierarchy process. *European Journal of Operational Research* 48, 9–26.

Saaty, T.L., Forman, E.H., 2000. *Expert Choice–Advanced Decision Support Systems Software*. Expert Choice Inc., Pittsburgh.

Seymour, D., 1996. *High Performing Colleges: The Malcolm Baldrige National Quality Award as a Framework for Improving Higher Education*. Prescott Publishing, Marvills, OH.

Steeples, M.M., 1993. *The Corporate Guide to the Malcolm Baldrige National Quality Award*. Business one-Irwin, Homewood, IL.

Wilson, D.D., Collier, D.A., 2000. An empirical investigation of the Malcolm Baldrige National Quality Award causal model. *Decision Science* 32, 2, 361–390.